Springer Proceedings in Mathematics & Statistics

Alberto Pinto
David Zilberman *Editors*

Modeling, Dynamics, Optimization and Bioeconomics IV

DGS VI JOLATE, Madrid, Spain, May 2018, and ICABR, Berkeley, USA, May–June 2017—Selected Contributions

Springer

Springer Proceedings in Mathematics & Statistics

Volume 365

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at http://www.springer.com/series/10533

Alberto Pinto · David Zilberman
Editors

Modeling, Dynamics, Optimization and Bioeconomics IV

DGS VI JOLATE, Madrid, Spain, May 2018, and ICABR, Berkeley, USA, May–June 2017—Selected Contributions

Springer

Editors
Alberto Pinto
Department of Mathematics
and LIAAD—INESC TEC
Faculty of Science
University of Porto
Porto, Portugal

David Zilberman
Department of Agricultural and Resource
Economics
University of California
Berkeley, CA, USA

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-030-78162-0 ISBN 978-3-030-78163-7 (eBook)
https://doi.org/10.1007/978-3-030-78163-7

Mathematics Subject Classification: 37-XX, 91-XX, 60-XX, 62-XX

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

A foreword attempts to address the reader to the contents of the rest of the book. In this sense, although it appears first, it is usually written last. As a consequence of the case of this book, most of its chapters were written before the appearance of the COVID-19, but the foreword is being written while even suffering the third wave of this terrible pandemic.

Scientists, including mathematicians and economists, have some responsibilities in the challenges that we are facing in the second and third decades of the 21st century, and it should appear with extreme clarity during this time of the COVID-19 pandemic that we cannot even solve our very specific problems. Global problems require more and more multidisciplinary efforts and solutions. This book is a good test of such approaches and efforts. There are many areas where we should be able to produce out of our comfortable area either in basic or applied research that enables advance and solutions for social problems. The contents of this book extend from health, climate change, online behavior of agents, social responsibility of firms, neural networks, demand estimation of new products in absence of prior information, or management of pensions schemes. Together with some basic research, this publication greatly contributes to some of the mentioned lines of research in a multidisciplinary way.

The book has its origins in two conferences held at a time at the Universidad Nacional de Educación a Distancia (UNED) during the month of May 2018. UNED hosted the *6th International Conference on Dynamics Games and Science 2018* (DGS-VI-2018) and the *19th Jornadas Latinoamericanas de Teora Económica* (Jolate-XIX). As usual, these activities present an opportunity for junior and senior researchers to join in a vibrant environment of networking, discussing, and exchanging ideas and experiences. These two acts featured plenary lectures by prominent keynote and thematic parallel sessions. Thanks are due to the Scientific and Organization Committees for both the academic and social contents of the conferences. The authors of this foreword, the Dean of the Faculty of Economics and Business at UNED (A. A. Álvarez-López) and the Head of the Department of Economic Theory and Mathematical Economics in the same Faculty (J. M. Labeaga), wish to express their acknowledgement to Alberto Pinto and Elvio Accinelli (co-organizers of the conferences), and also to David Zilberman (co-editor of this book with Alberto

Pinto). Many thanks to them, and also to the people of UNED, who worked hard so that everyone could enjoy the best environment.

Those conferences were a meeting point of scientists coming from different areas, including Data Analysis, Dynamical Systems, Game Theory, Mathematical Finance, Optimization, Industrial Organization, Stochastic Optimal Control, Accounting, Marketing, Management and Business Organization, and their applications to Economics, Engineering, Energy, Natural Resources and, in general, Social Sciences.

Papers of the book, written by leading researchers, reflect this diversity of interactions among different fields, with the common point of the applicability of mathematical and quantitative methods. They discuss topics ranging from management and business to finance and accounting, including marketing and social corporate responsibility; from demand estimation with stochastic models to multi-agent systems, including applications of Game Theory; from pure mathematics to neural networks, including dynamical systems; and from health to transportation. A vibrant mixture reflecting the spirit of the mentioned conferences.

At the time of writing this foreword we globally face the COVID-19 pandemic and climate change. Although the lockdown is contributing around the world to improve the environmental conditions, once the economy starts, the environment will again suffer the pressure. We are aware that each problem should be solved at a time and now it is urgently time of fighting against coronavirus. But the environment will need our help as citizens and researchers once this nightmare passes on. Everyone's attendance, each one from his or her own field and capacity, is more important than ever. This volume gives a very good instance of what is possible to achieve when people from different fields apply their knowledge, on a mathematical and quantitative basis, to relevant problems in today's society.

Madrid, Spain
January 2021

José M. Labeaga
Alberto A. Álvarez-López

Acknowledgements

We thank the authors of the chapters for sharing their vision with us in this book and we thank the anonymous referees.

We are grateful to José M. Labeaga and Alberto A. Álvarez-López for contributing the foreword of the book.

We thank the Executive Editor for Mathematics, Computational Science and Engineering at Springer-Verlag, Martin Peters, for his invaluable suggestions and advice throughout this project.

We thank João Paulo Almeida, Susan Jenkins, José Martins, Abdelrahim Mousa, Atefeh Afsar, Bruno Oliveira, Diogo Pinheiro, Filipe Martins, Miguel Arantes, Miguel Ferreira, Renato Soeiro, Ricard Trinchet Arnejo, Susana Pinheiro and Yusuf Aliyu Ahmad for their invaluable help in assembling this volume and for editorial assistance.

Alberto Adrego Pinto would like to thank LIAAD-INESC TEC and gratefully acknowledge the financial support received by the FCT—Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology)—within project UIDB/50014/2020 and ERDF (European Regional Development Fund) through the COMPETE Program (operational program for competitiveness) and by National Funds, through the FCT within Project "Modeling, Dynamics and Games" with reference PTDC/MAT-APL/31753/2017.

Contents

Immune Response Model Fitting to CD4$^+$ T Cell Data in *Lymphocytic Choriomeningitis Virus* LCMV infection

Atefeh Afsar, Filipe Martins, Bruno M. P. M. Oliveira, and Alberto A. Pinto

Abstract We make two fits of an ODE system with 5 equations that model immune response by CD4$^+$ T cells with the presence of regulatory T cells (Tregs). We fit the simulations to data regarding gp61 and NP309 epitopes from mice infected with *lymphocytic choriomeningitis virus* LCMV. We optimized parameters relating to: the T cell maximum growth rate; the T cell capacity; the T cell homeostatic level; and the ending time of the immune activation phase after infection. We quantitatively and qualitatively compare the obtained results with previous fits in the literature using different ODE models and we show that we are able to calibrate the model and obtain good fits describing the data.

Keywords T cells · Regulatory T cells (Tregs) · *Lymphocytic choriomeningitis virus* (LCMV) · Epitope gp61 · Epitope NP309 · Antigenic stimulation · Residuals · Fits

1 Introduction

The immune system has many components, one of the most important being T cells, a type of lymphocyte that develops in the thymus. The invasion by a pathogen will result in Antigen presenting cells (APC) [3] presenting the pathogen's characteristic peptides. T cells are activated by their specific antigen and start secreting cytokines, e.g. interleukine 2 (IL-2), in order to alert other cells of the immune system and to promote proliferation [29]. Occasionally, a mistake may happen and a clonotype of

A. Afsar · F. Martins (✉) · A. A. Pinto
LIAAD–INESC TEC, Department of Mathematics, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal
e-mail: luis.f.martins@inesctec.pt

A. A. Pinto
e-mail: aapinto@fc.up.pt

B. M. P. M. Oliveira
LIAAD–INESC TEC, Faculdade de Ciências da Nutrição e Alimentação da Universidade do Porto, Rua do Campo Alegre, 823, 4150-180 Porto, Portugal
e-mail: bmpmo@fcna.up.pt

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_1

T cells is stimulated by self-antigens, resulting in an auto-immune response. The role of Regulatory T cells (Tregs), a subset of T cells, is to prevent such auto-immunity through their immune suppressive action. In particular, Tregs inhibit cytokine secretion by T cells [6, 25, 26]. Mis-regulation of Tregs may result in the appearance of auto-immune diseases such as IPEX [24].

There has been as increasing interest in mathematical modelling of immune responses based on both in vivo or in vitro data. See the reviews by Zhu et al. [29, 30]. There are several mathematical modelling techniques that have been used to model immune responses by T cells. See Callard et al. [11] and Lythe et al. [18] for reviews. León et al. [16] used a hypergeometric distribution in a discrete model, while de Boer et al. [12], Burroughs et al. [6], Pinto et al. [23], Blyuss et al. [4] and Khailaie et al. [15] studied systems of ordinary differential equations (ODE). In this work we will use the ODE model in Pinto et al. [23] developed from the article from Burroughs et al. [6]. A 7 ODE model with two T cell clonotypes was used in Atefeh et al. [1] to fit the time dynamics data from Homann et al. [14].

Here, we aim to use the model with 5 ODE from Pinto et al. [23], that will allow us to make two separate fits to the data from Homann et al. [14], one fit for epitope gp61 and another fit for epitope NP309. In Sect. 2, we present the immune response model. Afterwards, in Sect. 3, we fit the model to the data. We discuss and analyse the fits in Sect. 4 and we finish with some conclusions in Sect. 5.

2 Immune Response Model

We consider the set of 5 ordinary differential equations from Sect. 3 in Pinto et al. [23] to model $CD4^+$ T cells and regulatory T cells. Tregs are activated at a rate a by self antigens from an inactive state denoted by R, to an active state denoted by R^*. Similarly, T cells are activated at a rate b by their specific antigen to the IL-2 secreting state T^* from the non-secreting state T. All cells proliferate when they adsorb IL-2 cytokines. We also consider an inflow of immune T cells into the tissue (T_{in}) and Tregs (R_{in}), which can represent T cell circulation or naïve T cells from the thymus.

The model consists of a set of five ordinary differential equations. We have an equation for each T cell population (inactive Tregs R, active Tregs R^*, non-secreting T cells T, secreting or activated T cells T^*) and interleukine 2 density I.

$$
\begin{aligned}
\frac{dR}{dt} &= (\varepsilon\rho I - \beta(R + R^* + T + T^*) - d_R)R + \hat{k}(R^* - aR) + R_{in},\\
\frac{dR^*}{dt} &= (\varepsilon\rho I - \beta(R + R^* + T + T^*) - d_{R^*})R^* - \hat{k}(R^* - aR),\\
\frac{dT}{dt} &= (\rho I - \beta(R + R^* + T + T^*) - d_T)T + k(T^* - bT + \gamma R^* T^*) + T_{in},\\
\frac{dT^*}{dt} &= (\rho I - \beta(R + R^* + T + T^*) - d_{T^*})T^* - k(T^* - bT + \gamma R^* T^*),
\end{aligned}
$$

$$\frac{dI}{dt} = \sigma(T^* - (\alpha(R + R^* + T + T^*) + \delta)I).$$

The parameter values used and obtained in the fits are presented in the Appendix, in Tables 1 and 2. As in Afsar et al. [1], we will consider death rates to be equal $d_T = d_R$ and $d_{R^*} = d_{T^*}$, and also equal relaxation rates $k = \hat{k}$. The inflow of T cells T_{in} can be different from the inflow of Tregs R_{in}. Further details of the model are presented in Burroughs et al. [5–10], Oliveira et al. [21, 22], Pinto et al. [23] and Yusuf et al. [28].

3 Simulations and Fits

In this section we describe our methods and do the fits of the model to two time series of the immune response by CD4$^+$ T cells to *lymphocytic choriomeningitis virus* - LCMV in mice obtained from laboratory experiments from Homann et al. [14]. Each of the time series that constitute the data regarding the concentration of T cells that responded to each of the two different LCMV epitopes studied, the gp61 (14 data points) and the NP309 (8 data points).

Previously published articles provided the estimates for most of the parameter values [2, 6, 20, 21, 27] and we considered that the parameters in Table 1 are fixed. The parameters to be fitted are in Table 2. Their admissible range was obtained from the literature [6, 19, 21]. These parameters are: the T cell maximum growth rate ρ/α; the T cell capacity $T^{cap} = \rho/(\alpha\beta)$; the homeostatic T cell level T^{hom} which is related to the T cell inflow T_{in} by $T_{in} \approx T^{hom} d_T$, assuming no antigenic stimulation of T cells $b = 0$ and small homeostatic values $T^{hom} + R^{hom} << T^{cap}$. See [23] for further details. Note that the parameters used in the optimization procedure are related to the T cells activated by their respective epitope, while the parameters related with Tregs and the parameters related with cytokines were fixed.

Following Afsar et al. [1], we assumed that the T cells were stimulated by the pathogen between the inoculation time $t = 0$ and a final time t_{end}. Hence, we considered a specific step function of time for the antigen simulation parameter b: a high value of the antigenic stimulation of T cells $b = 1000$ for times within $[0, t_{end}]$, and a very small value, $b = 10^{-5}$, outside that interval, see Table 1. The ending times t_{end} of high antigen stimulation intensity b corresponding to the exposition to a pathogen were also considered in the optimization procedure to fit the model to the data. See Table 2 for the fits for epitopes gp61 and NP309.

The initial condition (at time $t = 0$) corresponds to the homeostatic controlled state, with a low concentration of T cells. This state was achieved by numerically integrating the model for a long time, for a small value of antigenic stimulation of T cells, $b = 10^{-5}$.

Our objective was to minimize the sum of squares of the residuals in a logarithmic scale:

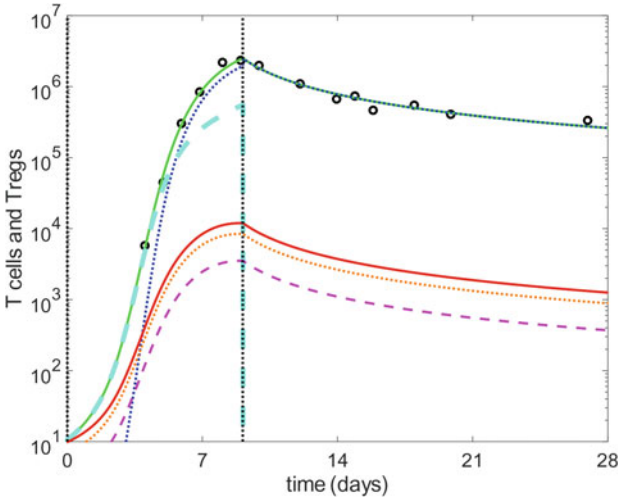


Fig. 1 Fit of the ODE model to the laboratory measurements of the CD4$^+$ T cells concentration for the gp61 epitope. The vertical axis has logarithmic scale. Green line: $T + T^*$; Blue dots: T; Cyan dashes: T^*; Red line: $R + R^*$; Orange dots: R; Violet dashes: R^*. The vertical dashed line marks the ending time $t_{end} = 9.102$ of the immune response phase. $res^2_{gp61} = 0.294$ and $MNSQ_{gp61} = 0.0294$. The parameters are shown in Tables 1 and 2

$$res^2 = \sum_{i=1}^{\#\text{data points}} \left(\log x(t_i) - \log(T(t_i) + T(t_i)^*)\right)^2$$

where #data points is the number of data points, $x(t_i)$ are the concentrations of T cells from Homann et al. [14] and t_i are the observation times, and $T(t_i) + T(t_i)^*$ represents the total T cell concentration from the fit at time t_i. The observation times range from 0 to 28 days after infection. Starting from several initial random parameter estimates, we used an iterative procedure to obtain the optimized values of the parameters. The optimized parameters for the best fit are presented in Table 2 of the appendix. The fit to the data is presented graphically in Figs. 1 and 2.

The software used for the simulations of the model was GNU Octave—*version 5.2.0*. The routine used for the numerical integration of the ODE system was `lsode`. The routine used for the optimization procedure of minimizing the residuals was `fminsearch` implementing the Nelder-Mead algorithm/downhill simplex method.

Starting from time $t = 0$ we observe in Figs. 1 and 2 that the concentration of T cells grows in a non-linear fashion, initially with the majority of T cells being secreting T cells T^*. Hence, as the IL-2 cytokine is being increasingly secreted (data not shown), we observe an increase in the growth rate of both T cells and Tregs. However, with higher concentration of Tregs, in particular, active Tregs R^*, there is an augmented relaxation rate of secreting T cells T^* to the non-secreting state T. At some point in time, around day 5 for epitope gp61 and near day 9 for epitope NP309, the concentration of secreting T cells is close to the concentration of non-secreting

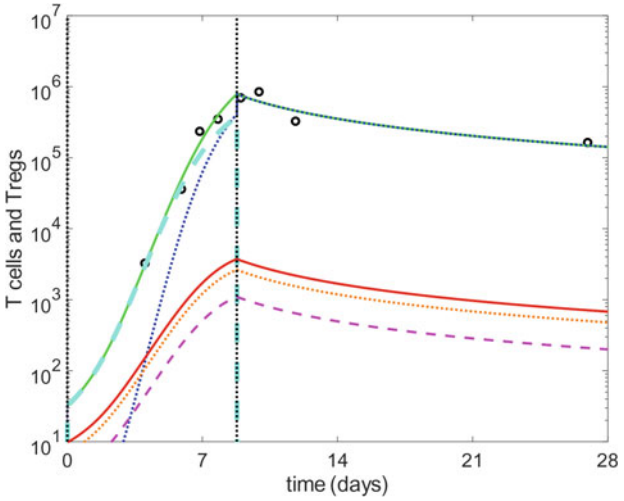


Fig. 2 Fit of the ODE model to the laboratory measurements of the CD4$^+$ T cells concentration for the NP309 epitope. The vertical axis has logarithmic scale. Green line: $T + T^*$; Blue dots: T; Cyan dashes: T^*; Red line: $R + R^*$; Orange dots: R; Violet dashes: R^*. The vertical dashed line marks the ending time $t_{end} = 8.782$ of the immune response phase. $res^2_{NP309} = 0.507$ and $MNSQ_{NP309} = 0.127$. The parameters are shown in Tables 1 and 2

T cells. After that point in time, the growth rate of T cells slows down until the end of the immune response phase. In the relaxation phase, we observe a reduction both in the concentration of T cells and Tregs due to the quick decrease of secreting T cells, and consequently the drop of IL-2 to negligible levels (data not shown), thereby stopping growth. Immediately after the end of the immune response phase, the decay in the concentrations of all cells is faster due to the action of the Fas-FasL death by apoptosis, modelled by the quadratic term with β (see more details in Afsar et al. [1] and Pinto et al. [23]). As the concentration of T cells becomes smaller, this term will have a reduced relevance in the decay and the linear death term will have a higher relative importance. We note that the concentration of inactive Tregs is higher than the concentration of active Tregs since their activation rate is relatively small $a < 1$.

For the gp61 epitope we obtained the residual sum of squares $res^2_{gp61} = 0.294$ while for the NP309 epitope we obtained $res^2_{NP309} = 0.507$ as presented in Figs. 1 and 2, respectively. Furthermore, we also present the residual mean square $MNSQ = res^2/df$ for each fit, where the number of degrees of freedom $df =$ #data points $-$#free parameters, and #free parameters $= 4$ is the number of free parameters. So we obtained, respectively, $MNSQ_{gp61} = 0.0294$ and $MNSQ_{NP309} = 0.127$.

4 Discussion

Afsar et al. [1] have previously simultaneously fitted the time series of the immune responses for epitopes gp61 and NP309 from Homann et al. [14] to a 7-ODE model. For each series, they have obtained the sum of residuals $res^2_{gp61} = 0.565$ and $res^2_{NP309} = 0.995$, and they have obtained the residual mean square $MNSQ = 0.0975$ for the simultaneous fit. Comparatively, we now obtain lower sum of residuals for both epitopes. Our residual mean square value for the epitope gp61 fit is lower than the value for the simultaneous fit from [1] while for the epitope NP309 fit it is slightly higher. This might be explained by the fact that on the simultaneous fit we have a total of $14 + 8 = 22$ data points, while having only 8 data points for the separate NP309 fit.

The same data have been analysed by de Boer et al. [13]. They have made two fits for each epitope, with a biphasic and a monophasic contraction phase respectively. For the gp61 epitope they obtained respectively $MNSQ_{gp61} = 0.06$ and $MNSQ_{gp61} = 0.17$. For the NP309 epitope they obtained respectively $MNSQ_{NP309} = 0.10$ and $MNSQ_{NP309} = 0.12$. When comparing our fits to theirs we see that for epitope gp61 fit we obtain lower residual mean square values than both their biphasic and monophasic models. For the epitope NP309 fit we obtain a higher residual mean square value, albeit very close to their value for the monophasic model.

The equations being used here are different from de Boer et al. [13]. Similarly to Afsar et al. [1], we use the same equation. We used the same equations over time and changed only one parameter in time: the antigenic stimulation of T cells b. This parameter was either in the "on" or "off" state, which is a simplification of a biologically more complex time evolution. The ending time of the immune activation phase t_{end} is a parameter in the optimization procedure and specific to each epitope. Moreover, contrary to [1] where $t_{ini} = 3.31$, here we have set the initial time of the immune activation phase $t_{ini} = 0$, which is the time when the disease is inoculated in the mice according to the experimental description in Homann et al. [14]. Opposite to de Boer et al. [13], we observe that the immune response activation seems to end earlier for epitope NP309 than for epitope gp61 ($t_{end} = 8.782$ and $t_{end} = 9.102$, respectively), but we agree with them that the peak time is quite uncertain and would require further research with different models and/or different and more frequent data around that point.

As in Afsar et al. [1], we assumed that the initial condition is the controlled homeostatic state. Regarding the homesostatic T cells level, in [1], the T^{hom} values for each T cell clonotype are 2409 and 787.74, while here we obtained $T^{hom} = 10.94$ and $T^{hom} = 32.79$ for gp61 and NP309, respectively. While the homeostatic T cells level apparently differ by two orders of magnitude, the difference is mostly explained by these values effectively working as a parameter determining the initial condition for the ODE system. Notice that around day 3, for both epitopes, our simulations have concentrations of T cells near 1000 (see Figs. 1 and 2). With respect to the value of the maximum growth rate of T cells, it is close to the one reported by de Boer et al. [13], although the "true" growth rate in our model is time dependent (see Afsar et al. [1] for more details). We estimated a higher maximum growth rate for the gp61

epitope than for the NP309 epitope, which is also the case in de Boer et al. [13]. Using the quadratic death term β allows for a smooth transition between the rapid apoptosis phase we observe around day 9 and the slower death rate we observe later, instead of two separate phases as presented in de Boer et al. [13]. We observe that this is also indirectly present in the optimization procedure since β influences T cell capacity $T^{cap} = \rho/(\alpha\beta)$.

Although the value we selected for the death rates of T cells is low, our model does not include the differentiation into memory cells, like in Afsar et al. [1], thus we used the initial 28 days of the data from Homann et al. [14]. Another limitation is to consider that the antigenic stimulation of T cells b is a step function of time instead of a smooth function, that would require more parameters. Since we have a small number of data points, the fixed parameters have the same values for both epitopes, which can be interpreted as that both clonotypes of T cells have similar characteristics. Our model has two strengths. Firstly, it has the same equations in all phases of time, the immune activation phase and the contraction phase, and without needing to consider biphasic regimes. Secondly, in our fit the free parameters were specific to the T cells responding to each epitope.

5 Conclusions

In conclusion, our model allows to achieve a good fit to the data of the immune responses from Homann et al. [14] to the gp61 and NP309 LCMV epitopes, having most of the parameters at the values described from the literature, and optimizing a set of 4 parameters related to the maximum T cell growth rate, the T cell capacity, the T cell homeostatic level, and the ending time of the immune activation phase. We discuss and interpret the results and compare our fits with other fits using different ODE models from the literature.

As future work, we could compute confidence intervals for the parameters and check for robustness of the estimated parameters using other data sets with more observations and more frequent T cell measurements over time.

Acknowledgements The authors would like to thank the financial support by FCT–Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) as part of project UID/EEA/50014/2019 and within project "Modelling, Dynamics and Games" with reference PTDC/MAT–APL/31753/2017. Atefeh Afsar would like to thank the financial support of FCT–Fundação para a Ciência e a Tecnologia–through a Ph.D. grant of the MAP–PDMA program with reference PD/BD/142886/2018.

Appendix

The parameters of our model and their default values as well as the fitted parameters are presented in the following two tables (Tables 1 and 2).

Table 1 Parameters values for our model of T cells and Tregs from [1, 2, 6, 13, 17, 21, 27]

Fixed parameters	Symbol	Value
T cells T, T^*		
Death rate ratio of active: inactive T cells	d_{T^*}/d_T	1
Death rate of inactive T cells (day^{-1})	d_T	10^{-3}
Secretion reversion[b] (day^{-1})	k	2.4
Antigen Stimulation of T cells	b	1000
Tregs R, R^*		
Growth rate ratio Tregs: T cells	ε	0.6
Relaxation rate (day^{-1})	$\hat{k}$	2.4
Death rate ratio of inactive Tregs: inactive T cells	d_R/d_T	1
Death rate relative ratio of Tregs: T cells	$\frac{d_{R^*}}{d_R}/\frac{d_{T^*}}{d_T}$	1
Tregs antigen stimulation level (day^{-1})	$a\hat{k}$	1
Homeostatic Treg level[c] (cells day^{-1})	R^{hom}	10
Secretion inhibition	γ	$10\ R_{hom}^{-1}$
Interleukin-2 (IL-2) Cytokine		
Max. cytokine concentration[d] (pM)	$1/\alpha$	200 pM
IL2 secretion rate (pM/day)	σ	144
Cytokine decay rate (day^{-1})	$\sigma\delta$	36

[b]This is in the absence of Tregs

[c]This is in terms of the homeostatic Treg level $R_{hom} \approx \frac{R_{in}}{d_R}\left(1 - \frac{d_R - d_{R^*}}{\hat{k}(a+1)+d_R}\right)$

[d]This is taken as 20 times the receptor affinity ($10pM$)

Table 2 Fitted parameters values for our model of T cells and Tregs, with ranges adopted from [1, 6, 13, 20, 21]

Parameter	Symbol	Range	Fitted value
gp61 epitope ($res^2 = 0.294$, $MNSQ = 0.0294$)			
T cell Maximum growth rate (day^{-1})	ρ/α	< 6	3.298
Capacity of T cells (10^7 cells)	$T^{cap} = \rho/(\alpha\beta)$	$0.3 - 3$	1.847
Homeostatic T cells level[a] (cells day^{-1})	T^{hom}	$< 10^5$	10.94
End of the immune activation phase (day)	t_{end}	$6 - 15$	9.102
NP309 epitope ($res^2_{NP309} = 0.507$, $MNSQ_{NP309} = 0.127$)			
T cell Maximum growth rate (day^{-1})	ρ/α	< 6	1.702
Capacity of T cells (10^7 cells)	$T^{cap} = \rho/(\alpha\beta)$	$0.3 - 3$	0.577
Homeostatic T cells level[a] (cells day^{-1})	T^{hom}	$< 10^5$	32.79
End of the immune activation phase (day)	t_{end}	$6 - 15$	8.782

[a]T cell inflow level is given by $T_{in} = T^{hom} d_T$

References

1. Atefeh, A., Martins, F., Oliveira, B.M.P.M., Pinto, A.: A fit of CD4$^+$ T cell immune response to an infection by lymphocytic choriomeningitis virus. Math. Biosci. Eng. **16**(6), 7009–7021 (2019)
2. Anderson, P.M., Sorenson, M.A.: Effects of route and formulation on clinical pharmacokinetics of interleukine-2. Clin. Pharmacokinet. **27**, 19–31 (1994)
3. Banchereau, J., Briere, F., Caux, C., Davoust, J., Lebecque, S., Liu, Y.-J., Pulendran, B., Palucka, K.: Immunobiology of dendritic cells. Annu. Rev. Immunol. **18**(1), 767–811 (2000)
4. Blyuss, K.B., Nicholson, L.B.: The role of tunable activation thresholds in the dynamics of autoimmunity. J. Theor. Biol. **308**, 45–55 (2012)
5. Burroughs, N.J., Ferreira, M., Martins, J., Oliveira, B.M.P.M.: Alberto A. Pinto, and N. Stollenwerk. Dynamics and biological thresholds. In: Pinto, A.A., Rand, D.A., Peixoto, M.M. (eds.) Dynamics, Games and Science I, DYNA 2008, in Honor of Maurício Peixoto and David Rand, vol. 1 of Springer Proceedings in Mathematics, pp. 183–191. Springer, Berlin Heidelberg (2011)
6. Burroughs, N.J., Oliveira, B.M.P.M., Pinto, A.A.: Regulatory T cell adjustment of quorum growth thresholds and the control of local immune responses. J. Theor. Biol. **241**, 134–141 (2006). July
7. Burroughs, N.J., Ferreira, M., Oliveira, B.M.P.M., Pinto, A.A.: Autoimmunity arising from bystander proliferation of T cells in an immune response model. Math. Comput. Model. **53**, 1389–1393 (2011)
8. Burroughs, N.J., Ferreira, M., Oliveira, B.M.P.M., Pinto, A.A.: A transcritical bifurcation in an immune response model. J. Differ. Equ. Appl. **17**(7), 1101–1106 (2011). July
9. Burroughs, N.J., Oliveira, B.M.P.M., Pinto, A.A., Ferreira, M.: Immune response dynamics. Math. Comput. Model. **53**, 1410–1419 (2011)
10. Burroughs, N.J., Oliveira, B.M.P.M., Pinto, A.A., Sequeira, H.J.T.: Sensibility of the quorum growth thresholds controlling local immune responses. Math. Comput. Model. **47**(7–8), 714–725 (2008). April
11. Callard, R.E., Yates, A.J.: Immunology and mathematics: crossing the divide. Immunology **115**, 21–33 (2005)
12. de Boer, R.J., Hogeweg, P.: Immunological discrimination between self and non-self by precursor depletion and memory accumulation. J. Theor. Biol. **124**(3), 343–369 (1987). February
13. de Boer, R.J., Homann, D., Perelson, A.S.: Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. J. Immunol. **171**, 3928–3935 (2003)
14. Homann, D., Teyton, L, Oldstone, M.B.A.: Differential regulation of antiviral T-cell immunity results in stable CD8+ but declining CD4+ T-cell memory. Nat. Med. 7(8), 913–919 (2001)
15. Khailaie, S., Bahrami, F., Janahmadi, M., Milanez-Almeida, P., Huehn, J., Meyer-Hermann, M.: A mathematical model of immune activation with a unified self-nonself concept. Front. Immunol. **4**, 474 (2013). December
16. León, K., Lage, A., Carneiro, J.: Tolerance and immunity in a mathematical model of T-cell mediated suppression. J. Theor. Biol. **225**, 107–126 (2003)
17. Lowenthal, J.W., Greene, W.C.: Contrasting interleukine 2 binding properties of the alpha (p55) and beta (p70) protein subunits of the human high-affinity interleukine 2 receptor. J. Exp. Med. **166**, 1155–1069 (1987)
18. Lythe, G., Molina-París, C.: Some deterministic and stochastic mathematical models of naïve T-cell homeostasis. Immunol. Rev. **285**(1), 206–217 (2018)
19. Michie, C.A., McLean, A., Alcock, C., Beverley, P.C.L.: Life-span of human lymphocyte subsets defined by CD45 isoforms. Nature **360**, 264–265 (1992)
20. Moskophidis, D., Battegay, M., Vandenbroek, M., Laine, E., Hoffmann-Rohrer, U., Zinkernagel, R.M.: Role of virus and host variables in virus persistence or immunopathological disease caused by a non-cytolytic virus. J. Gen. Virol. **76**, 381–391 (1995)

21. Oliveira, B.M.P.M., Figueiredo, I.P., Burroughs, N.J., Pinto, A.A.: Approximate equilibria for a T cell and Treg model. Appl. Math. Inf. Sci. **9**(5), 2221–2231 (2015)
22. Oliveira, B.M.P.M., Trinchet, R., Otero-Espinar, M. V., Pinto, A., Burroughs, N.: Modelling the suppression of autoimmunity after pathogen infection. Math. Methods Appl. Sci. 41(18), 8565–8570 (2018)
23. Pinto, A.A., Burroughs, N.J., Ferreira, F., Oliveira, B.M.P.M.: Dynamics of immunological models. Acta. Biotheor. **58**, 391–404 (2010)
24. Sakaguchi, S.: Naturally arising $CD4^+$ regulatory T cells for immunological self-tolerance and negative control of immune responses. Annu. Rev. Immunol. **22**, 531–562 (2004)
25. Shevach, E.M., McHugh, R.S., Piccirillo, C.A., Thornton, A.M.: Control of T-cell activation by $CD4^+$ $CD25^+$ suppressor T cells. Immunol. Rev. **182**, 58–67 (2001)
26. Thornton, A.M., Shevach, E.M.: $CD4^+$ $CD25^+$ immunoregulatory T cells suppress polyclonal T cell activation in vitro by inhibiting interleukine 2 production. J. Exp. Med. **188**(2), 287–296 (1998)
27. Veiga-Fernandes, H., Walter, U., Bourgeois, C., McLean, A., Rocha, B.: Response of naïve and memory $CD8^+$ T cells to antigen stimulation in vivo. Nat. Immunol. **1**, 47–53 (2000)
28. Yusuf, A.A., Figueiredo, I.P., Afsar, A., Burroughs, N.J., Oliveira, B.M.P.M., Pinto, A.A.: The effect of a linear tuning between the antigenic stimulations of $CD4^+$ T cells and $CD4^+$ Tregs. Mathematics 8(2), 293. (February 2020)
29. Zhu, J., Paul, W.E.: CD4 T cells: fates, functions, and faults. Blood **112**(5), 1557–1569 (2008)
30. Zhu, J., Yamane, H., Paul, W.E.: Differentiation of effector CD4 T cell populations. Annu. Rev. Immunol. **28**(1), 445–489 (2010)

Construction of a New Model to Investigate Breast Cancer Data

Umut Ağyüz, Vilda Purutçuoğlu, Eda Purutçuoğlu, and Yüksel Ürün

Abstract Modelling is a way to describe the elements of the network/system, their states and their interactions with other elements in order to understand the current state of knowledge of a system. Thereby, the mathematical models may predict the experiments which are difficult or impossible to do in the lab and can be used to discover indirect relationships between model's components. Hereby, the aim of this study is to develop a network structure for the breast cancer from the analyses of different datasets which include the data of the luminal type at the stage 1–3 breast cancer diagnosed in total 377 patients and related to the PI3KCD signalling pathway. Accordingly, in the analyses, the relations of the 65 oncogenes are revealed by a true network in a binary format. Then, we construct the quasi breast cancer networks by using different parametric and non-parametric models, namely, Gaussian graphical model, copula Gaussian graphical model and multivariate adaptive regression splines with/without interaction terms. In the computations, we evaluate the performance of all suggested mathematical models via F-measure and accuracy measure criteria. We consider that the outcomes can be useful for the selection of the best fitted model in the construction of the breast cancer gene-gene interaction networks.

Keywords Breast cancer · Multivariate adaptive regression splines model · Copula Gaussian graphical model · Generalized additive models · Systems biology

U. Ağyüz
GENZ Biotechnology, Bilkent Cyberpark, 06800 Ankara, Turkey
e-mail: umut@genzbio.com

V. Purutçuoğlu (✉)
Department of Statistics, METU, Middle East Technical University, 06800 Ankara, Turkey
e-mail: vpurutcu@metu.edu.tr

E. Purutçuoğlu
Department of Social Work, Ankara University, Ankara, Turkey
e-mail: purutcu@agri.ankara.edu.tr

Y. Ürün
Department of Medical Oncology, Ankara University School of Medicine, Ankara, Turkey
e-mail: yuksel.urun@ankara.edu.tr

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_2

1 Introduction

As the biological systems have very high dimensions and complex structures, the understanding of their activations becomes more difficult if their properties are only detected from the analyses in labs. Therefore, various mathematical models are suggested to describe these systems. By this way, we can check the current state of knowledge about the underlying structures, predict the experiments which are difficult or impossible to do in the lab and discover indirect relationships between models' components [31]. These mathematical models are based on distinct assumptions in order to decrease the complexity and to follow the core activations. So, they can be mainly classified under three branches, namely, Boolean models, steady-state models and stochastic models. The former modeling type can detect only the on/off positions of the systems by checking the kick-off of each gene in the flow of activation. If the gene indicates activations after the on/off position, it is denoted by 1/0, respectively. Due to its simplicity, this model is generally preferred to gather information about not well-known systems [7]. On the other hand, the steady-state models are the most common modelling type since the majority of the available data about the biological systems is suitable for this sort of models. The ordinary differential equations models [7], the generalized additive models [16], probabilistic Boolean models [32] are some examples of this type of modelling. Finally, the stochastic models aim to explain the random nature of the systems by taking into account the molecular changes of the proteins [7]. But since such detailed information can be hardly obtained for complex systems and there are sparse measurements which are suitable for this type of models, their applications are limited. Hence, in this study, we work on the steady-state type of models due to their common usages and focus on specific type of network diseases, selected as the breast cancer.

The metastases to various tissues in patients with early breast cancer is still diagnosed despite the advanced management and prognosis of the disease. The location of metastases may be a determinant for the length of survival after the recurrence [19, 24]. Accordingly, about the half of the patients with metastatic breast cancer develops hepatic metastases. However, the molecular determinants of tissues which are specific metastatic markers have not yet been specified. It is found that a better command of knowledge about the molecular markers of the organ specific tropism will affect the decision about the adjuvant therapy and the treatment of advanced disease. Nevertheless, there are also certain studies in the literature which indicate particularly the effect of some proteins in breast cancers. For instance, Kimbung et al. [20] find that the CLDN2 protein is frequently over-expressed in the breast cancer liver metastases, and in addition, conclusively demonstrate that the primary tumors from patients who are diagnosed with hepatic recurrences also frequently express high levels of the claudin-2 protein [20]. They reveal the evidence that claudin-2 is a potential prognostic factor for predicting the likelihood of a breast tumor to relapse specifically in the liver and is furthermore a general predictor of the early breast cancer recurrences.

Furthermore, Turashvili et al. [29] examine 30 samples, which are composed of the normal ductal and lobular cells from 10 patients, IDC cells (invasive ductal cancer) from 5 patients, ILC cells from 5 patients and microdissected from cryosections of 10 mastectomy specimens from postmenopausal patients gather from the Affymetrix U133 Plus 2.0 arrays. In their works, they show that the genes involved in the extracellular matrix-tumor cell interactions, such as Collagen type I, III, V, XI, Fibronectin 1 and Versican, are found to be increased in tumor cells. Moreover, it is found that the cytoskeleton proteins such as Type I [8, 18, 30] and Type II [19] keratins are decreased in tumor cells. Additionally, the expression of seven differentially expressed genes (namely, CDH1, EMP1, DDR1, DVL1, KRT5, KRT6 and KRT17) is verified by the immunohistochemistry on the tissue microarrays. The expression of ASPN mRNA is validated by in situ hybridization on frozen sections, and CTHRC1, ASPN and COL3A1 are tested by PCR. Hence, the combined pairwise comparison and the alterations of gene expressions reveal that the ductal and lobular carcinomas have several genes in common, however, they can be discriminated both at the gene and the protein levels. Later, LaBreche et al. [23] describe a human breast tumor predictor based on the gene expression of the tumor-bearing transgenic mouse blood [23]. The list of 4276 Affymetrix Mouse Genome 430 2.0 probe identifiers is translated into 2595 orthologous human probes (Affymetrix HU133 Plus 2) and is also used to filter the human dataset of this study, which is normalized employing principal components of the Affymetrix control probe sets. This strategy overcomes many of the limitations of earlier studies by using the model system to reduce noise and to identify transcripts associated with the presence of a breast tumor over other potentially confounding factors. This may serve as a proof-of-concept for using an animal model to develop a blood-based diagnostic, and it can establish an experimental framework for identifying predictors of solid tumors, not only in the context of the breast cancer, but also, in other types of cancers. Besides, Kreike et al. [22] use gene expression profiling to study invasive breast carcinomas from patients. In their studies, by using the 18K cDNA microarrays, the gene expression profiles are obtained from 50 patients who undergo BCT (Breast conserving therapy). In the analyses of these 50 patients, 19 develop a local recurrence; the remaining 31 patients are selected as controls as they are free of the local recurrence at least 11 years after the treatment. For 9 of 19 patients, it is seen that the local recurrence is also available for the gene expression profiling. Then, unsupervised and supervised methods of classification are used to separate patients in groups corresponding to the disease outcome and to study the overall gene expression pattern of primary tumors and their recurrences. Finally, a systematic evaluation of the effects of the tissue preservation and the prolonged cold ischemia on the RNA quality as well as the expression of single genes and multigene signatures across 17 primary breast tumors that represent a range of disease stages and surgical conditions are performed [17]. According to the results, it is observed that the sample preservation in RNA later improves the RNA yield and quality, whereas, the cold ischemia increases the RNA fragmentation as measured by the $3'/5'$ expression ratio of control genes. However, expression levels of single genes and multigene signatures that are of the diagnostic

the relevances in the breast cancer are mostly unaffected by the sample preservation method or the prolonged cold ischemic duration.

Hereby, in this study, in order to both see the specific effect of proteins and construct a quasi protein-protein interaction network for the breast cancer, we use different datasets which include the data of the luminal type stage 1–3 breast cancer diagnosed in total 377 patients. We suggest distinct mathematical modeling approaches in order to identify the predictors of breast tumors. Whereas, in these analyses, we cannot model the associated biological network with the realistic size and the complexity over time periods due to the computational limitation in inference of the model parameters. Hereby, we ignore certain details of the state of the system and find the gene relevance networks which are based on the covariance graph models.

Accordingly, in the construction of networks, we perform various models which are particularly designed for high dimensional and highly nonlinear as well as correlated structures. We apply the Gaussian graphical model [10, 11, 32] and the copula Gaussian graphical model [9] with two different types of the Bayesian inference algorithms [25, 27] among parametric networks' modeling. On the other hand, we apply the multivariate adaptive regression splines, shortly MARS, [1–3, 14], which are specifically designed for the biological networks among non-parametric approaches. In the analyses, we use the underlying approaches to construct networks via the significant genes in the breast cancer. These genes are gathered from the KEGG database by detecting the microarray studies about breast cancers and then by selecting the differentially significant genes from these microarray analyses. Later, we generate quasi systems whose interactions can be also validated from the oncogenous researches. Finally, we evaluate the performance of our mathematical models under two major accuracy measures and also investigate the new biological findings from the analyses.

Hereby, in the organization of the study, we shortly describe the mathematical models in Sect. 2. In Sect. 3, we present the description of the data collection and the selected genes as well as the results of the analyses with discussion. Finally, in Sect. 4, we summarize the outputs and present the future work.

2 Mathematical Models

In the construction of biological networks, there are various modeling types which are dependent on different assumptions. For instance, the selection of an appropriate model can change whether your system has stochastic activations from its nature, or it can be better described via deterministic activations, or there is a limited information about the system and its major flow of activations can be understood via the on/off position of each gene. Hence, we can choose stochastic, steady-state or boolean types of networks, in order. Among these alternates, the steady-state types of models are more common as majority of the available data, such as microarray datasets [6], can be better explained by these sorts of models. In this study, while constructing the quasi breast cancer networks via oncogenous genes, we apply five microarray datasets and

implement five different modeling choices. These models are specifically designed for high dimensional and sparse biological systems.

The first model which we suggest is the *Gaussian graphical model* (GGM). This model is based on the multivariate normality assumption of the states Y, i.e. observations, when the genes have the mean μ and covariance Σ. Here, Σ is a $p \times p$-dimensional matrix for a p number of total genes when each gene has n number of observations. So $Y = (Y_1, Y_2, \ldots, Y_p)$ and $Y \sim N_p(\mu, \Sigma)$. In GGM, under the lasso regression

$$Y_p = \beta Y_{-p} + \varepsilon \tag{1}$$

when β refers to the regression coefficient, ε denotes the random errors and finally Y_p shows the state of the pth gene. Accordingly, Y_{-p} is the state of the remaining genes except the pth gene. So in this model, each gene is explained via its regression of other genes in the system which implies a very high dependency between the predictors, i.e., genes, from the definition of the mathematical model. Our aim in this model is to infer the regression coefficient which can be formulated via the inverse of the covariance matrix Σ. This matrix is called the precision matrix Θ and the relation between β and Θ can be defined by

$$\beta = -\frac{\Theta_{-p,p}}{\Theta_{pp}}. \tag{2}$$

In this expression $\Theta_{-p,p}$ indicates the precision when the pth state is excluded and $\Theta_{p,p}$ denotes the precision computed under totally p genes. Under the normal distribution, the zero precision implies the independence between the associated pair of random variables and accordingly, under the multivariate normality, it implies the conditional independence between the gene-pair. Thereby, by using the simplicity of the conditional independency under the multivariate normal distribution, Θ is estimated in place of β in GGM due to their proportional relations. Thus, in the construction of the network via GGM, any non-zero entry in Θ is interpreted as the functional relationship between genes and is shown by an indirected edge between genes. On the other hand, the zero entry in Θ stands for the conditionally no relationship between the pair, given that all other genes and is presented without edge in the graphical representation of the system. Thus, in inference of this highly dependent and linear model, we use either some iterative methods based on Bayesian approaches, such as reversible jump Markov chain Monte Carlo (RJMCMC) [12, 13, 27] and birth-and-death MCMC (BDMCMC) [25], or some alternative penalized likelihood algorithms based on the frequentist theory [33], such as the graphical approach, also called the glasso method [15], or some nonparametric methods based on optimizations, such as the neighborhood selection [26], adaptive lasso and the fused lasso. Indeed in the applications of the RJMCMC and BDMCMC algorithms, the description of the multivariate normality assumption of GGM is slightly revised in the sense that this joint distribution is partitioned via the Gaussian copula [9, 28]. This new representation of GGM is known as the copula Gaussian graphical model

(CGGM). In the estimation of model parameters in CGGM, which is still the precision matrix Θ similar to GGM, the high dimensional Θ is inferred via the Cholesky decomposition which enables us to construct the model for every dimension due to its simplicity in matrix decomposition. This idea is used in the RJMCMC algorithm. Whereas, because of the implementation of the Bayesian steps in RJMCMC, the computational time of this approach cannot be efficient for the R programming language although the accuracy of the inference is high [12, 27]. On the other hand, BDMCMC, as an MCMCM method, is more efficient than RJCMCM, however, its accuracy is low [4, 12, 27]. In this study, we apply both GGM and CGGM with two alternative estimation methods, in the construction of the breast cancer networks as the steady-state and parametric modeling approaches.

On the other side, we also apply nonparametric version of these parametric representations above. In this study, we perform the multivariate adaptive regression splines (MARS) model among alternates. MARS [14] is a complex regression model where the predictors in the model are highly dependent on each other and their relations are nonlinear. Basically, this model belongs to the generalized additive model [16] and can be considered as the alternative approach of the ordinary differential equations [7] in the construction of the biological networks. The simple mathematical description of the model is

$$y = \beta_0 + \sum_{m=1}^{M} \beta_m H_m(x) + \varepsilon \tag{3}$$

where β_0 presents the intercept term and β_m refers to the regression coefficient for the random variables $x = (x_1, x_2, \ldots, x_n)'$ and a response variable y. Moreover, M shows the total number of parameters and ε is the random error term. Finally, H_m denotes the spline basis function (BFs) as can be seen below.

$$H_m(y) = \prod_{k=1}^{K_m} [\max(s_{k,m}(x_{v(k,m)}) - t_{k,m}, 0)] \tag{4}$$

in which K_m implies the number of truncated linear functions multiplied in the mth BF and $s_{k,m}$ is the input variable corresponding to the kth truncated linear function in the mth BF. In the calculation via MARS, there are two-stage algorithms, namely, forward and backward stage. In the forward stage, the most complex model is generated by taking into account all interactions' effects. In the backward stage, the less effective regressors are eliminated by using the generalized cross validation criterion so that the optimal simplest model can be selected. In the application of this model in the biological network, the full model in Eq. 3 is reduced to Eq. 1 by discarding all interactions' effects and is implemented for each gene in the system separately, and then by binding the individual estimated links for each gene is combined under a single precision as well as an adjacency matrix [2, 3]. This simple version of MARS, also called loop-based MARS [2], is called the MARS without interaction effect model in this study. Then, as the extension of this model, we also accept the second-

order interaction effects [1] since this sort of interactions can be used to describe the feed-forward loop motifs [5] in the networks. This version of MARS is named as MARS with interaction terms in our analyses. In the following part, we present the application of these models in real datasets.

3 Application

3.1 Data Description

In modeling of the breast cancer networks, we initial select 65 hot spot oncogenes from the breast cancer pathway hsa05524 in the KEGG database (https://www.genome.jp/kegg-bin/show-pathway?hsa05224). These genes are mentioned in a number of GWAS studies about the breast cancer [17, 20, 22, 23, 29]. Then, we use the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and choose five raw datasets which include the underlying hot spot oncogenes. The chosen datasets are named as GSE46141 [20], GSE4913 [29], GSE27567 [23], GSE5764 [22] and GSE25011 [17]. These data include the measurements of the luminal type at stage 1–3 breast cancer diagnosed in a total of 377 patients. Below, we represent the details of each dataset.

1. **GSE46141**: The samples are obtained by the fine needle aspiration biopsy from metastatic lesions at different anatomical locations in metastatic breast cancer patients. Then, the samples are collected prior to the treatment. The studied metastasis samples (total sample size n=91) include the liver metastases ($n = 16$), bone metastases ($n = 5$), lung metastases ($n = 2$), lymph node metastases ($n = 39$), local metastases [breast metastases (n 11), skin metastases ($n = 17$) and ascites metastasis ($n = 1$). In this dataset, whole genome transcriptional analysis is performed using the Rosetta/Merck Human RSTA Custom Affymetrix 2.0 microarray [HuRSTA-2a520709] platform.
2. **GSE5764**: In this dataset, the samples include the normal ductal ($n = 10$), normal lobular ($n = 10$), invasive ductal carcinoma ($n = 5$) and invasive lobular carcinoma ($n = 5$) cells obtained by the laser capture microdissection system from 5 invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) patients. Whole genome transcriptional analysis is performed by using the Affymetrix Human Genome U133 Plus 2.0 Array platform and the expressional differences between normal and carcinoma cells as well as the expressional differences between IDC and ILC are investigated.
3. **GSE27567**: In this study, the peripheral blood cell samples (PBCs) from the MMTV/c-MYC transgenic female mice in the FVB/NJ (an albino, inbred laboratory mouse) strain background are used. Hereby, the tumor-free control mice ($n = 28$ while 14 of them in the validation set and the remaining 14 mices in the training set) and mices with advanced mammary tumors ($n = 65$ while 33 of them in the validation set and the remaining 32 of them in the training set)

are included in the study. Furthermore, whole genome transcriptional analysis is implemented by using the Affymetrix Mouse Genome 430 2.0 Array platform. In this analysis, a model based on the transcriptional profiling is produced to distinguish the tumor-free and tumor-bearing mice with 100% sensitivity and 100% specificity. Later, the data in mice are confirmed in the human samples. In this validation, a total of 161 samples including benign breast abnormalities ($n = 37$), ectopic cancers ($n = 22$), malignant breast cancers ($n = 57$), healthy controls ($n = 31$) and post-surgery breast cancers ($n = 15$) are included in the study and whole genome transcriptional analysis is conducted under the Affymetrix Human Genome U133 Plus 2.0 Array platform.

4. **GSE4913**: In this analysis, the primary tumors from 50 patients who received breast-conserving therapy (BCT) are included. Among these 50 patients, 9 patients experience the local recurrence of the tumor. These 9 tumors are also included in the study. Moreover, whole genome transcriptional analysis is applied by using the NKI-AVL Homo sapiens 18K cDNA microarray platform. Here, the gene expression profiles of recurrent and non-recurrent tumors are found as similar in such a way that the recurrent tumors are less differentiated and the ER-negativity is higher in these tumors. Additionally, the expression profiles of primary tumors and their locally recurrent counterpart are observed similarly too.
5. **GSE25011**: In this dataset, 11 breast cancer tissues are taken and most of the patients are at the stage IIA or IIB. Then, the impact of RNA-later, snap-freezing and tissue waiting time until the RNA-stabilization on the RNA integrity and microarray expression analysis are investigated in the analyzing part. Accordingly, 86 samples from 11 breast cancer tumors including different time points are used in the microarray gene expression study and finally, whole genome transcriptional analysis is performed by using the Affymetrix Human Genome U133A Array platform.

Once the raw measurements from the above datasets are taken, they are normalized via the inter-array normalization by using the quantile method [18] before the mathematical modeling. The selected gene names are listed as follows: ESR1, NCOA1, NCOA3, JUN, SP1, TNFSF11, FGF1, FGF18, FGFR1, IGF1R, EGF, EGFR, KIT, SHC1, GRB2, SOS1, SOS2, KRAS, NRAS, MAPK1, PIK3CA, PIK3CD, PIK3CB, PIK3R1, PIK3R3, PTEN, AKT2, AKT3, MTOR, RPS6KB2, JAG1, NOTCH1, NOTCH2, NOTCH3, NOTCH4, HES1, HEYL, HEY2, WNT2, WNT5A, WNT10B, WNT11, FZD1, FZD7, FZD2, FZD3, FZD4, FZD5, FZD6, LRP6, DVL3, DVL2, AXIN1, APC2, CTNNB1, CSNK1A1, TCF7, TCF7L2, LEF1, GADD45B, GADD45G, BAK1, CDK6, E2F3, BRCA1. We present the functional properties of these genes in Tables 1, 2, 3 and 4. Moreover, we present the quasi true network with those related genes in Fig. 1. In this graph, the pathway between growth factors, receptor tyrosine kinases, oncogenes and PI3KCD is demonstrated.

Finally, in order to prepare the true network for the selected 65 oncogenes, their relations are identified in the STRING database by using expressions and co-expression evidence links. In our analyses, we download all the relations and

Table 1 Gene names and functionalities of selected breast cancer genes (continued)

Gene symbol	Gene name	Function
AKT2	AKT Serine/Threonine Kinase 2	AKT2 is a kinase and downstream mediator of PI3K pathway and phosphorates many substrates
AKT3	AKT Serine/Threonine Kinase 3	AKT3 is a kinase and downstream mediator of the PI3K pathway. AKT3 phosphorates many substrates and is involved in the cell survival and the tumor formation
APC2	Adenomatosis Polyposis Coli 2	APC2 is closely related to the tumor suppressor APC. It is involved in the inhibition of the Wnt signalling through the degradation of beta-catenin and the stabilization of microtubule
AXIN1	Axis Inhibition Protein 1	It is a negative regulator of the Wnt signalling through the degradation of beta-catenin and likely a tumor suppressor. It interacts with APC and beta-catenin
BAK1	BCL2 Antagonist/Killer 1	BAK1 is localized in mitochondria and involved in the induction of the apoptosis through the release of cytochromes from the mitochondria. It interacts with the tumor suppressor P53
BRCA1	Breast Cancer 1, Early Onset	BRCA1 plays a role in the DNA repair and the genomic stability. BRCA1 mutations are associated with the hereditary breast and ovarian cancers
CDK6	Cyclin Dependent Kinase 6	CDK6 is involved in the cell-cycle regulation and the G1/S transition. It interacts with the tumor suppressor, pRB and over expresses in cancers
CSNK1A1	Casein Kinase 1 Alpha 1	It participates in the Wnt signalling and phosphorates beta-catenin
CTNNB1	Beta-Catenin	It is a key element of the Wnt signalling pathway and involved in the regulation of the cell adhesion.
DVL2	Dishevelled Segment Polarity Protein 2	DVL2 is involved in the Wnt signalling by binding to frizzled family members
DVL3	Dishevelled Segment Polarity Protein 3	DVL3 is involved in the Wnt signalling and the cell proliferation
E2F3	E2F Transcription Factor 3	E2F3 is a DNA binding transcription factor that is involved in the cell-cycle regulation and the DNA replication. It interacts with RB
EGF	Epidermal Growth Factor	EGF is a potent mitogenic growth factor that binds to EGFR. Its deregulation is implicated in several cancers
EGFR	Epidermal Growth Factor Receptor	EGFR is a receptor tyrosine kinase. It is involved in the cell proliferation. The amplifications and mutations in this gene are implicated in the progression of many cancers.
ESR1	Estrogen Receptor 1	ESR1 is a nuclear estrogen receptor and involved in the pathogenesis of breast cancers. It plays a role in the regulation of gene expressions and the cell proliferation

Table 2 Gene names and functionalities of selected breast cancer genes (continued)

Gene symbol	Gene name	Function
FGF1	Fibroblast Growth Factor 1	FGF1 is a ligand of FGFR1 and a potent mitogen. It is involved in tumor formations and invasions
FGF18	Fibroblast Growth Factor 18	FGF18 is involved in the cell proliferation, especially, in the liver and the intestine
FGFR1	Fibroblast Growth Factor Receptor 1	FGFR1 is a receptor for the fibroblast growth factors. Its downstream signaling is involved with PI3K and MAPK signallings
FZD1	Frizzled Class Receptor 1	FZD1 is a receptor for Wnt proteins and functions in the Wnt signalling
FZD2	Frizzled Class Receptor 2	FZD2 is a receptor for Wnt proteins and functions in the Wnt signalling
FZD3	Frizzled Class Receptor 3	FZD3 is a receptor for Wnt proteins and functions in the Wnt signalling
FZD4	Frizzled Class Receptor 4	FZD4 is a receptor for Wnt proteins and functions in the Wnt signalling
FZD5	Frizzled Class Receptor 5	FZD5 is a receptor for Wnt proteins and functions in Wnt signalling
FZD6	Frizzled Class Receptor 6	FZD6 is a receptor for Wnt proteins and functions in Wnt signalling
FZD7	Frizzled Class Receptor 7	FZD7 is a receptor for Wnt proteins and functions in the Wnt signalling
GADD45B	Growth Arrest And DNA Damage Inducible Beta	GADD45B is activated by the DNA damage. Then, it activates P38/JNK signalling through MEKK4. It is involved in the regulation of apoptosis
GADD45G	Growth Arrest And DNA Damage Inducible Gamma	GADD45G is activated by the DNA damage. Then, it activates P38/JNK signalling through MEKK4. It is involved in the regulation of apoptosis
GRB2	Growth Factor Receptor Bound Protein 2	GRB2 is an adaptor protein that binds to the cell surface growth factor receptors and transmits signals to the RAS signalling
HES1	Hes Family BHLH Transcription Factor 1	HES1 a transcriptional repressor of genes that require a bHLH protein for their transcription. It may have a role in responses to the DNA cross-link damage
HEYL	Hes Related Family BHLH Transcription Factor With YRPW Motif-Like	HEYL is a repressive transcription factor. It is a downstream effector of the notch signalling
HEY2	Hes Related Family BHLH Transcription Factor With YRPW Motif 2	HEY2 is a repressive transcription factor. It is a downstream effector of the notch signalling
IGF1R	Insulin Like Growth Factor 1 Receptor	IGF1R is a receptor tyrosine kinase for IGF1. It is involved in the cell survival, tumor transformations, PI3K and RAS-MAPK activations
JAG1	Jagged 1	JAG1 is a ligand notch1 receptor and involved in the notch signalling

Table 3 Gene names and functionalities of selected breast cancer genes (continued)

Gene symbol	Gene name	Function
JUN	Jun Proto-Oncogene AP-1 Transcription Factor Subunit	JUN is a transcription factor and involved in cancers
KIT	KIT Proto-Oncogene Receptor Tyrosine Kinase	KIT is a receptor of the stem cell factor. It activates AKT1 and RAS signallings. The mutations in this gene are associated with cancers
KRAS	KRAS Proto-Oncogene, GTPase	KRAS is a member of the RAS family. The mutations in RAS genes are frequently observed in cancers and the RAS signalling is critical in the promotion of the cell proliferation
LEF1	Lymphoid Enhancer Binding Factor 1	LEF1 is a transcriptional factor and involved in the Wnt signalling. It is associated with various cancers
LRP6	LDL Receptor Related Protein 6	LRP6 acts as a co-receptor in the Wnt signaling and participates in the Wnt signalling
MAPK1	Mitogen-Activated Protein Kinase 1	MAPK1 is a serine/threonine kinase involved in MAPK/ERK signalling. MAPK1 is downstream of the RAS signalling. It controls diverse biological processes and phosphorates many substrates
MTOR	Mammalian Target of Rapamycin	MTOR is a part of the PI3K/Akt pathway and involved in regulations of cell growths and metabolisms. Its deregulation is seen in many cancers
NCOA1	Nuclear Receptor Coactivator 1	NCOA1 is a nuclear receptor coactivator of steroid receptors such as ER
NCOA3	Nuclear Receptor Coactivator 3	NCOA3 is a nuclear receptor coactivator of steroid receptors such as ER
NOTCH1	Notch 1	Notch1 is a transmembrane receptor involved in the notch signalling. The notch signalling is important for the cell fate determination, cell-cell interactions and the cell proliferation. The notch1 receptor overexpression is observed in numerous cancer types
NOTCH2	Notch 2	Notch2 is a transmembrane receptor involved in the notch signalling. The notch signalling is important for the cell fate determination and cell-cell interactions
NOTCH3	Notch 3	Notch3 is a transmembrane receptor involved in Notch signalling. Notch signalling is important for cell fate determination and cell-cell interactions
NOTCH4	Notch 4	Notch4 is a transmembrane receptor involved in the notch signalling. The notch signalling is important for the cell fate determination and cell-cell interactions.
NRAS	NRAS Proto-Oncogene, GTPase	NRAS is a member of the RAS family. Mutations in the RAS genes are frequently observed in cancers and the RAS signalling is critical in the promotion of the cell proliferation

Table 4 Gene names and functionalities of selected breast cancer genes (continued)

Gene symbol	Gene name	Function
PIK3CA	Phosphatidylinositol-4, 5-Bisphosphate 3-Kinase Catalytic Subunit Alpha	PIK3CA is the catalytic subunit involved in the PI3K pathway which is implicated in cancers, especiallyi in breast cancers. PI3KCA phosphorates phosphatidylinositol molecules. It is involved in the cell proliferation, cell survivals and inhibitions of apoptosis
PIK3CB	Phosphatidylinositol-4, 5-Bisphosphate 3-Kinase Catalytic Subunit Beta	PIK3CB is the catalytic subunit involved in the PI3K pathway which is implicated in cancers, especially, in breast cancers. PI3KCB phosphorates phosphatidylinositol molecules. It is involved in the cell proliferation, cell survivals and inhibitions of apoptosis
PIK3CD	Phosphatidylinositol-4, 5-Bisphosphate 3-Kinase Catalytic Subunit Delta	PIK3CD is the catalytic subunit involved in the PI3K pathway. PI3KCD phosphorates phosphatidylinositol molecules. It is involved in immune responses
PIK3R1	Phosphoinositide-3-Kinase Regulatory Subunit 1	PIK3R1 is the regulatory subunit involved in the PI3K pathway which is implicated in cancers, especially, in breast cancers. The catalytic subunit of PI3K phosphorates phosphatidylinositol molecules. The PI3K pathway is involved in the cell proliferation, cell survivals and inhibitions of apoptosis
PIK3R3	Phosphoinositide-3-Kinase Regulatory Subunit 3	PIK3R3 is the regulatory subunit involved in PI3K pathway which is implicated in cancer especially in breast cancer. Catalytic subunit of PI3K phosphorates phosphatidylinositol molecules. PI3K pathway is involved in cell proliferation, cell survival and inhibition of apoptosis
PTEN	Phosphatase And Tensin Homolog	PTEN is a tumor suppressor with a high frequency of mutations in various cancers. PTEN dephosphorylation phosphatidylinositol molecules to counteract the PI3K pathway which is important in carcinogenesis
RPS6KB2	Ribosomal Protein S6 Kinase B2	RPS6KB2 phosphorylates the S6 ribosomal protein and eukaryotic translation initiation factor 4B to increase protein syntheses
SHC1	SHC Adaptor Protein 1	SHC1 couples the receptor tyrosine kinase activation to the RAS signalling by recruiting other adaptor proteins

(continued)

Table 4 (continued)

Gene symbol	Gene name	Function
SP1	Specificity Protein 1	SP1 is a zinc finger transcription factor that may act either as an activator or a repressor
SOS1	Son Of Sevenless Homolog 1	SOS1 is a guanine nucleotide exchange factor that is involved in the activation RAS
SOS2	Son Of Sevenless Homolog 2	SOS2 is a guanine nucleotide exchange factor that is involved in the activation RAS

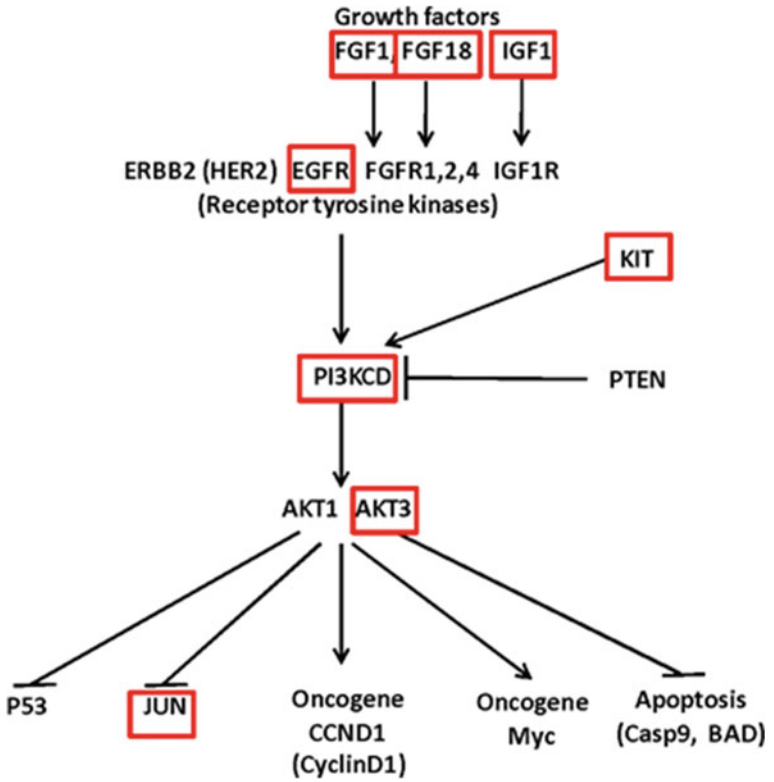


Fig. 1 PI3KCD centered signalling pathway. Here, factors that are implicated in the network of the model are indicated in a red box

transfer all the links amongst these 65 genes in the binary format so that the estimated adjacency matrices obtained from our network modelings can be comparable. During all these calculations, we use the R programming language and the codes of models are originally developed, apart from GGM (with huge package) and BDMCMC (with BDgraph package) calculations.

3.2 Results

In our analyses, we use four different models, as stated beforehand, for the construction of breast cancer networks. These models are GGM, CGGM, MARS without/with interaction models. Among these alternatives, we conduct the inference of CGGM via two approaches, RJMCMC and BDMCMC since the performance of the model is highly dependent on the selected estimation methods. Then, we compare the validation of the estimated networks by the following accuracy measures (Table 6).

Table 5 Gene names and functionalities of selected breast cancer genes (continued)

Gene symbol	Gene name	Function
TCF7	Transcription Factor 7	TCF7 is a transcription factor which is complexed with Beta-catenin and involved in the Wnt signalling. It is mainly expressed in T-cells
TCF7L2	Transcription Factor 7 Like 2	TCF7L2 is a transcription factor involved in the Wnt signalling
TNFSF11	TNF Superfamily Member 11	TNFSF11 is a cytokine of TNF Superfamily. It binds to TNFRSF11B and TNFRSF11A. It can activate AKT
WNT2	Wnt Family Member 2	WNT2 is a member of the Wnt family and acts as ligands for the Wnt signalling These proteins are implicated in oncogeneses
WNT5A	Wnt Family Member 5A	WNT5A is a member of the Wnt family and act as ligands for the Wnt signalling. These proteins are implicated in oncogenesss
WNT10B	Wnt Family Member 10B	WNT10B is a member of the Wnt family and acts as ligands for the Wnt signalling. These proteins are implicated in oncogeneses
WNT11	Wnt Family Member 11	WNT11 is a member of the Wnt family and acts as ligands for the Wnt signalling. These proteins are implicated in oncogeneses

Table 6 General confusion matrix

		Actual class	
		Positive	Negative
Predicted class	Positive	TP	FP
	Negative	FN	TN

$$\mathrm{F} = 2\,\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{5}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP+FP+TN+FN}}. \tag{6}$$

where

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}, \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}. \tag{8}$$

In these mathematical expressions, TP denotes the true positive which implies the number of correctly classified objects that have positive label, TN shows the true negative which presents the number of correctly classified objects having negative label. On the other hand, FP is the false positive which stands for the number of misclassified having negative label and finally, FN indicates the false negative which means the number of misclassified objects with positive label. The description of each notation is also represented by a confusion matrix as shown in Table 6 while actual and predicted classes imply the biologically validated and estimated interactions between any two genes, respectively.

Hereby, in Table 7 indicates the mean of F-measure and accuracy of each model based on our five breast cancer datasets. From the findings, it is seen that GGM is the best option among parametric models and MARS without interaction is more successful than other nonparametric models. Indeed, from our previous analyses [4], it has been shown that CGGM can capture the true links more successfully than GGM, particularly, when it is estimated via the RJMCMC approach. Whereas, its performance is affected very much from the number of the MCMC iterations which directly controls the convergence of the estimates. For large networks, the rate of convergence is low due to the complicated computations of the MCMC scheme of the RJMCMC algorithm. Moreover, its codes are run under the R programming language. Therefore, its accuracy measures are lower than GGM in this analysis. Therefore, we consider that for a fast computational time, GGM is more advantageous and CGGM

Table 7 Comparison of means of F-measures and accuracies for five datasets

Model	F-measure	Accuracy
GGM	**0.1703**	**0.9295**
CGGM via RJMCMC	0.1049	0.6985
CGGM via BDMCMC	0.0549	0.8369
MARS without interaction	**0.3801**	**0.9498**
MARS with interaction	0.1725	0.7070

Table 8 Comparison of F-measures and accuracies for every single dataset with the MARS without interaction model

Dataset	F-measure	Accuracy
GSE 46141	0.1113	0.8682
GSE 4913	0.3978	0.9568
GSE 5764	0.5731	0.9894
GSE 25011	0.0979	0.8581
GSE 27567	0.1118	0.8696

with RJMCMC may be good with high numbers of iterations. On the other hand, when we assess both parametric and nonparametric models together, the nonparametric model gives more accurate results via the MARS model. If we compare the outcomes of MARS without and with interaction models, for our five datasets, the simplest model (MARS without interaction) has better accuracies. However, this result may not be generalized for all datasets [4]. Because if the selected genes for the description of genes have very high interactions and generate hubs [5], MARS with interaction model more successfully estimates the systems [1]. On the contrary, if the topology of the system becomes random, the MARS without interaction model can give more accurate results. But, in both ways, it is observed that MARS has a better performance than other alternative models. In Table 8, we present the performance of each dataset under the MARS without interaction model. The tabulated values indicate that the estimated model has an almost perfect accuracy with moderate F-measures.

Hereby, if we interpret the estimated structure of the network via its quasi true network as seen in Fig. 1, it is observed that the PI3K pathway can be stimulated by growth factors and their counterpart receptor tyrosine kinases (RTKs). Indeed, the PI3KCD activating signal pathway has a network with various important cellular proliferative factors such as c-Myc and Cyclin D1. PI3KCD inhibits p53 and apoptosis. JUN is also related with apoptosis in the breast cancer and it is inhibited by the PI3K signalling pathway [8]. In the figure, the red boxes indicate the proteins found from our estimated network and show that there is a high correlation between our estimated network and the pathways which have high impact on the breast cancer. So, in the signalling network of our model, growth factors such as FGF1, FGF18 and IGF1, and an RTK, EGFR, are implicated. PI3KCD is the central in our network as

it is in the PI3K pathway. Moreover, PI3KCD phosphorates and activates AKT1 and AKT3, and AKT3 is also an important element of the network. Here, it's important to note that the PI3KCD and AKT3 mutations are frequent in the breast cancer. The PI3KCD mutation rate is 43 in the ER-positive breast cancer and the AKT3 mutation rate is 28 in the basal type breast cancer [8]. Moreover, phosphoinositide 3-kinases (PI3Ks) generate lipid second messengers that regulate a broad variety of cellular responses such as growth, cell cycle progression, differentiation, vesicular traffic and cell migration [30]. Additionally, the PI3K activity is critical in a wide variety of normal and pathological responses, including the immune regulation, metabolic control and the cancer. However, despite the importance of this signalling system, the regulation of the PI3K gene expression under normal and disease conditions is not clear, whereas, the phosphatidylinositol 3-kinase pathway is the most frequently mutated pathway in the breast cancer [21].

Accordingly, the consensus of the models shows that the candidate models are significantly strong to catch the PI3KCD signaling interactions from biological networks of the breast cancer. However, MARS without interaction gives the best significant results and suggests to be used for tissue biopsies from postmenapousal women who are diagnosed ductal or lobular breast cancers Table 8. In this study, the candidate models are obtained by making use of whole genome transcriptional data, and the PI3KCD and PI3K pathway elements become prominent in these models. In line with our models, the mutational analysis of clinical specimens of breast tumors reveal that there is a high frequency of mutations in factors central to the PI3K pathway such as PI3KCD, AKT3 and PTEN. In fact, PI3KCD is the most frequently mutated gene in luminal breast cancers [8]. It is important to note that there is a correlation between candidate models and mutational status of the breast cancer. This is clear that these models should be trained with more data to increase the significance. However, in this article, we demonstrate which models are strong candidates for hunting biological interactions from real biological data, which might be useful in clinics, and what kind of information should be expected from results of these analyses.

4 Conclusion

In this study, we have compared the performance of five models in the construction of the breast cancer networks based on five real datasets. In this analysis, we have generated the list of potential oncogenous genes and validated our estimated networks with the STRING database' results. From the assessment of models based on F-measure and accuracy criteria, it has been observed that the nonparametric model, MARS without interaction, is more successful than alternates. Moreover, our estimated network can validate the related literature and as seen from the graph, the analyses of five datasets are all related to the PI3KCD signaling pathway in the sense that this pathway is closely related to growth factors, receptor tyrosine kinases, oncogenes and apoptosis related genes.

As the future work, we consider defining a merging rule in order to combine different datasets so that we can use distinct data simultaneously for inference. Also, we have been preparing an R package which combines all these listed models in order to facilitate the application of these models via the researchers.

Acknowledgements The authors thank to the BAP project at Middle East Technical University (Project no: BAP-08-11-2017-035 for their support.

Conflict of Interest No conflicts of interests to be declared.

References

1. Ağraz, M., Purutçuoğlu, V.: Extended lasso-type MARS (LMARS) model in the description of biological network. J. Stat. Comput. Simul. **89**(1), 1–14 (2019)
2. Ayyildiz, E., Ağraz, M., Purutçuoğlu, V.: MARS as an alternative approach of Gaussian graphical model for biochemical networks. J. Appl. Stat. **44c**(16), 2858–2876 (2017)
3. Ayyıldız, E., Purutçuoğlu, V.: Generating various types of graphical models via MARS. In: Arslan, O. (ed.) Chapter in: Information Complexity and Statistical Modeling in High Dimensions with Applications. Springer (In print) (2019)
4. Bahçivancı, B., Purutçuooğlu, V., Purutçuoğlu, E., Ürün, Y.: Estimation of gynecological cancer networks via target proteins. J. Multidiscip. Eng. Sci. **5**(12), 9296–9302 (2018)
5. Barabási, A.L., Oltvai, Z.N.: Network biology: understanding the cells functional organization. Nat. Rev. Genet. **5**, 101–113 (2004)
6. Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., et al.: Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat. Genet. **29**(4), 365–371 (2001)
7. Bower, J.M., Bolouri, H.: Computational Modeling of Genetic and Biochemical Networks. MIT Press, Cambridge (2001)
8. Cancer Genome Atlas: Comprehensive molecular portraits of human breast tumours. Nature **490**(7418), 61–70 (2012)
9. Dobra, A., Lenkoski, A.: Copula Gaussian graphical models and their application to modeling functional disability data. Ann. Appl. Stat. **5**(2A), 969–993 (2010)
10. Dokuzoğlu, D., Purutçuoğlu, V.: Comprehensive analyses of Gaussian graphical model under different biological networks. Acta Phys. Pol. Ser. A **132**, 1106–1111 (2017)
11. Edwards, D.: Introduction to Graphical Modelling, 2nd edn. Springer Texts in Statistics (2000)
12. Farnoudkia, H., Purutçuoğlu, V.: Copula Gaussian Graphical Modelling of Biological Networks and Bayesian Inference of Model Parameters. Scientia Iranica (in press) (2019)
13. Farr, W.M., Mandel, I., Stevens, D.: An efficient interpolation technique for jump proposals in reversible-jump Markov chain Monte Carlo calculations. R. Soc. Open Sci. 2 (2015). https://doi.org/10.1098/rsos.150030
14. Friedman, J.H.: Multivariate adaptive regression splines. Ann. Stat. **19**(1), 1–67 (1991)
15. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. Biostatistics **9**, 432–441 (2008)
16. Hastie, T., Tibshirani, R., Friedman, J.H.: The Element of Statistical Learning. Springer, New York (2001)
17. Hatzis, C., Sun, H., Yao, H., Hubbard, R.E., Meric-Bernstam, F., Babiera, G.V., Wu, Y., Pusztai, L., Symmans, W.F.: Effects of tissue handling on RNA integrity and microarray measurements from resected breast cancers. J. Natl. Cancer Inst. **103**(24), 1871–1883 (2011)
18. Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., et al.: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4**, 249–264 (2003)

19. Imkampe, A., Bendall, S., Bates, T.: The significance of the site of recurrence to subsequent breast cancer survival. Eur. J. Surg. Oncol. **33**, 420–423 (2007)
20. Kimbung, S., Kovács, A., Bendahl, P.O., Malmström, P., Fernö, M., Hatschek, T., Hedenfalk, I.: Claudin2 is an independent negative prognostic factor in breast cancer and specifically predicts early liver recurrences. Mol. Oncol. **8**(1), 119–128 (2014)
21. Kok, K., Nock, G.E., Verrall, E.A.G., Mitchell, M.P., Hommes, D.W., et al.: Regulation of p110d PI 3-Kinase Gene Expression. PLoS ONE **4**(4), (2012). https://doi.org/10.1371/journal.pone.0005145
22. Kreike, B., Halfwerk, H., Kristel, P., Glas, A., Peterse, H., Bartelink, H., Van de Vijver, M.J.: Gene expression profiles of primary breast carcinomas from patients at high risk for local recurrence after breast-conserving therapy. Clin. Cancer Res. **12**(19), 5705–5712 (2006)
23. LaBreche, H.G., Nevins, J.R., Huang, E.: Integrating factor analysis and a transgenic mouse model to reveal a peripheral blood predictor of breast tumors. BMC Med. Genomics. **4**(61) (2011)
24. Largillier, R., Ferrero, J.M., Doyen, J., Barriere, J., Namer, M., Mari, V., Courdi, A., Hannoun-Levi, J.M., Ettore, F., Birtwisle-Peyrottes, I., Balu-Maestro, C., Marcy, P.Y., Raoust, I., Lallement, M., Chamorey, E.: Prognostic factors in 1,038 women with metastatic breast cancer. Ann. Oncol. **19**, 2012–2019 (2008)
25. Mohammadi, A., Wit, E.C.: BDgraph: Bayesian structure learning of graphs in R. Bayesian Anal. **10**, 109–138 (2015)
26. Meinshausen, N., Bühlmann, P.: High dimensional graphs and variable selection with the lasso. Ann. Stat. **34**, 1436–1462 (2006)
27. Purutçuoğlu, V., Farnoudkia, H.: Gibbs sampling in inference of copula Gaussian graphical model adapted to biological networks. Acta Phys. Pol. Ser A **132**, 1112–1117 (2017)
28. Trivedi, P.K., Zimmer, D.M.: Copula modeling: an introduction for practitioners. Found. Trends R Econom. **1**, 1–111 (2005)
29. Turashvili, G., Bouchal, J., Baumforth, K., Wei, W, Dziechciarkova, M., Ehrmann, J., Klein, J., Fridman, E., Skarda, J., Srovnal, J. et al.: Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. BMC Cancer. **7**, 55–75 (2007)
30. Vanhaesebroeck, B., Leevers, S.J., Ahmadi, K., Timms, J., Katso, R., et al.: Synthesis and function of 3-phosphorylated inositol lipids. Ann. Rev. Biochem. **70**, 535–602 (2001)
31. Wilkinson, D.J.: Stochastic Modelling for Systems Biology. Taylor and Francis, Boca Raton, FL (2006)
32. Whittaker, J.: Graphical Models in Applied Multivariate Statistics. Wiley, New York (1990)
33. Wolpert, R.L., Schmidler, S.C.: α-Stable limit laws for harmonic mean estimators of marginal likelihoods. Stat. Sinica **22**, 1233–1251 (2012)

Optimal Pension Fund Management Under Risk and Uncertainty: The Case Study of Poland

I. Baltas, M. Szczepański, L. Dopierala, K. Kolodziejczyk, Gerhard-Wilhelm Weber, and A. N. Yannacopoulos

Abstract During the last decade, and especially after the financial crisis, the problem of providing supplementary pensions to the retirees has attracted a lot of attention from official bodies, as well as private financial institutions, worldwide. In this effort, there are various possible solutions, one of which is provided by pension fund schemes. Essentially, a pension fund scheme constitutes an independent legal entity that represents accumulated wealth stemming from pooled contributions of its members. The aim of the proposed research is to study the problem of optimal management of defined contribution (DC) pension fund schemes within general, complex and (as much as possible) realistic frameworks. From both a theoretical and practical point of view, one of the most important issue regarding fund management is the construction of optimal investment portfolio, because the success of a DC plan crucially depends on the effective investment of the available funds. Even though this problem has been heavily studied in the relative literature, the vast majority of the available works focuses: (i) on simple stylized models which allow for a very general understanding and are mainly based on intentionally unrealistic assumptions in order to provide closed-form (and paradigmatic) solutions, and (ii)

I. Baltas (✉)
Department of Financial and Management Engineering, University of the Aegean, 41 Kountouriotou Str., 82100 Chios, Greece
e-mail: jmpaltas@aegean.gr

M. Szczepański · K. Kolodziejczyk · G.-W. Weber
Faculty of Engineering Management, Poznań University of Technology, J.Rychlewskiego 2, 60-965 Poznań, Poland

L. Dopierala
Faculty of Economics, Department of International Business, University of Gdansk, Armii Krajowej 119/121, 81-824 Sopot, Poland

A. N. Yannacopoulos
Department of Statistics, Athens University of Economics and Business, 76 Patission Str., 10434 Athens, Greece

I. Baltas · A. N. Yannacopoulos
Stochastic Modeling and Applications Laboratory, Athens University of Economics and Business, 76 Patission Str., 10434 Athens, Greece

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_3

on risk levels (unrealistic) rather than uncertainty (realistic). This chapter presents preliminary results/general ideas of our project and aims to provide a detailed and an (as much as possible) realistic framework that takes into account the exposure of the fund portfolio into several market risks as well as model uncertainty with respect to the evolution of several unknown market parameters that govern the behavior of the fund portfolio. Our research will be directed towards the new public and occupational pension schemes in Poland.

Keywords Optimal portfolio · Defined contribution · Pension reform · Public and occupational pension schemes in Poland · Stochastic optimal control · Stochastic games

1 Introduction

Saving for retirement—like any form of long-term savings—involves various risks. The problem of optimal management of assets accumulated in pension funds under risk and uncertainty is a matter of fundamental importance—both in the theoretical (theory of finance, stochastic analysis) and practical dimension. It is part of a wider problem of planning, saving and optimal investment of pension funds. Optimal investment issue concerns not only additional pension systems (occupational and individual private pension schemes). In many countries around the world, capital solutions were first introduced and then limited also in public pension systems. Population ageing[1] and financial sustainability concerns have created pressures on policy makers to introduce pension reforms at the end of twentieth and in first decade of twenty-first centuries. The primary goal of these reforms was to reduce the risk in pension systems, ensure their long-term stability and adequacy of benefits. Some of these reforms of public pension consisted of changing the parameters in the system (e.g., extending the required periods of payment of retirement contributions allowing for the payment of a pension, extension of the statutory retirement age), others were systemic (e.g., full or partial privatization of public pension systems in Latin America, most of the countries of Central and Eastern Europe and of former Soviet Union from 1981 until 2014).

This second solution was supported in the 1990s by the World Bank [World Bank 1994]. The introduction of a capital-financed pillar to public pension systems and development of addition pension schemes (occupational and individual pensions savings, usually supported by the State in form of tax incentives) was to ensure risk diversification. Traditional public pension systems have been—and most countries still are—financed using the Pay-As-You-Go (PAYG) method: contributions from current workers are used to pay benefits to current retirees, giving current workers

[1] The demographic aging of the population (due to the extension of the average life expectancy and the reduction of fertility) means a reduction in the number of working population in relation to pensioners. This is a challenge for the long-term financial stability of pension systems, public finances, as well as the labor market and the pace of economic growth.

"promises" in return of contributions (these "promises" have different legal weights in different countries). The realization of these "promises" or legal obligations will be financed from contributions or general taxes of the next working generation. In such an unfunded pension system, no pension fund investments are needed. In the literature on pension economics, attention is also drawn to the political risk related to PAYG financing: the "promises" currently being submitted, what amounts of future pensions may not be kept, and system parameters, e.g., premiums, retirement age, valorization (indexation) of retirement benefits in relation to the in-crease in inflation and average pay of benefits, are easily changed. In addition, the increase in pensions depends on the labor force and wage growth rate, which in long term is likely to slow down in an aging population. Capital financing of pension systems is based on the use of contributions from current workers to accumulate assets; these assets are employed in part or in full to pay benefits in the future. Pension systems can be partially financed with PAYG, and partially funded.

Full or partially funded pension schemes were to provide higher benefits from the same pension contributions (because the rate of investment in the financial market in the long run exceeds the rate of labor incomes) and reduce the politicization of pension systems. Of course, capital-financed, fully or partially funded pension systems are exposed to investment risk, which is transferred to the system participants. Proponents of the privatization of public pension systems, however, have assumed that the risk could be reduced by an appropriate selection of financial instruments with different risk levels and proper investment policy conducted by specialized financial institutions which are managing pension funds.

Partial privatization of public pension systems (the introduction of a multi-pillar system, partially financed by the PAYG method, and partly on a capital basis) has found application in pension reforms in post-socialist countries in the 1990s and early 21st century. Along with multi-pillar pension systems (consisting of public, occupational and individual pillars) and the diversification of financing methods and related risks in the public, basic pension system first seemed to be an attractive, efficient, more safe and generally rational and profitable solution for all stakeholders (workers and future retirees, financial market institutions, private financial services providers, and the State).

This "experimentation" in social security systems proved to be unsuccessful. Sixty percent of countries that had privatized public mandatory pensions having reversed the privatization until 2018 [50]. One of the reasons was the realization of investment risk in pension funds as a result of the global financial and economic crisis of 2007–2009. That crisis has caused a fall in the value of assets of pension funds, with a particularly drastic impact on funds with the formula of a defined contribution. The privatization of public pension systems has not contributed to the adequacy of pension benefits; on the contrary—they have been reduced [28, 50]. The true beneficiaries of the privatization of public pension systems were private financial service providers managing pension funds, who generated high profits from various types of fees for asset management.

At this point it is worth recalling that the main types of retirement plans, in terms of determining the amount of the expected benefit are defined as defined benefit (DB)

and defined contribution (DC) schemes. Their design causes the polarization of risk distribution associated with pension provision. Most of the risks in DB plan are borne by the organizer, while in DC plan they are carried by the participant. Hybrid plans offer the possibility of risk-sharing between the stakeholders of a pension plan. Much less popular are hybrid solutions, especially in the context of occupational pension schemes associated with providing the expected level of pension. The investment risk occurs both in the accumulation and in the deccumulation (consumption) phase of pension funds (see, e.g., Lin et al. [44], Szczepański [57] and references therein).

While the investment risk in public pension systems has diminished due to the total or partial reduction of privately funded pension funds in public pension systems (reversing pension privatization), it remains still high in additional DC (private and occupational) pension systems. This wealth is to be invested over a long period of time (usually from 20 to 40 years) in order to provide its members with retirement benefits (in the form of periodic pension payments or a one-off payment). Employers, as well as employees (e.g., countries or other official bodies), that are part of the fund, periodically pay contributions before retirement (according to certain rules) which are appropriately invested in the financial market, in order for retirees to receive benefits at the time of retirement.

In general, as already mentioned, there are two completely different methods to design a pension fund scheme: (i) DC plan, and (ii) DB plan. According to a DC plan, every member of the fund contributes a fixed proportion of her income (before retirement), which are collected in an individual investment account and the benefits to be received (after retirement) consist of a fraction of the true fund value. Thus, they solely depend on the investment performance of the fund portfolio during its lifetime. In other words, for the DC case, the benefits to be received are not known before they are really obtained, i.e., that is, they are random variables. On the other hand, according to a DB plan, the benefits are usually initially fixed while the contributions are dynamically adjusted in order to keep the fund in balance. Today, the DC pension plans are far more popular, mostly due to the rapid dynamic development of financial markets, especially during the last twenty years, which has provided investors with a selection of attractive and advanced financial products. In some countries (including the USA and Great Britain) hybrid retirement programs are used, which are a combination of elements of DC and DB programs. However, they have a limited scope, they are used only in occupational pension schemes, and not in public pension systems.

Since the success of a DC plan crucially depends on the effective investment of the available funds due to contributions, the optimal management of the fund reduces to the problem of optimal portfolio selection from an available collection of financial assets. A prominent feature of this problem is encapsulated by the fact that the underlying financial variables that govern the evolution of the portfolio's value at each instant of time (like, e.g., volatility of assets returns, interest rates, etc.) are of stochastic nature, which is a typical characteristic of financial markets. To be more precise, it is well-known and clearly evidenced in the relative literature, that both volatility of asset returns (cf., e.g., Andersen and Bollerslev [3], Andersen et al. [38], Bates [8], Engle [30], Fama and French [31], Jones [41], Chen et al. [20] and

references therein) and interest rates (cf., e.g., Brennan and Schwartz [15], Brown and Schaefer [17], Deelstra [25], Markellos and Psychoyios [46] and references therein) are not constants but display a stochastic (random) behavior. Hence, in order to provide a realistic framework for the proposed research, it is of utmost importance to appropriately capture this behavior. In this attempt, we will resort to Stochastic Analysis, in particular, to Stochastic Modeling, Optimization and Optimal Control techniques.

Our chapter presents preliminary results/ideas of a broader research project. The aim of the proposed research is to study the problem of optimal management of DC pension fund schemes within general, stochastic and, hence (as much as possible), realistic frameworks in both continuous and discrete time frameworks. Even though the problem of optimal management of pension funds (in both DC case and DB case) has been extensively studied (especially in the continuous time framework) in the related literature (cf., e.g., Battocchio and Menoncin [9], Bodie et al. [11], Boulier et al. [12], Deelstra et al. [26], Di Giancinto et al. [34], Guan and Liang [36, 37], Zhang et al. [60] and references offered there), the vast majority of the available related literature focuses: (i) on simple stylized models including a very limited pool of risk factors (typically focusing on volatility risk or interest rates risk) and adopting some rather unrealistic modeling assumptions (e.g., volatility being constant through time, etc.) in order to provide closed-form solutions, (ii) on the extremely restricting and unrealistic assumption that the underlying probability model for the risk factors is known with certainty, i.e., focus on investment risk and ignoring the effects of uncertainty, which has been acknowledged by many authors (cf., e.g., Anderson et al. [2, 3], Hansen and Sargent [38] and references offer there) as the key factor when it comes to realistic modeling. Hence, it is apparent that these models are not sufficient for a detailed and realistic quantitative treatment of the problem under consideration.

In contrast to the existing literature, our proposal aims to provide a detailed and an (as much as possible) realistic framework that takes into account the exposure of the fund portfolio not only into several market risks (e.g., interest rate risk, volatility risk, etc.) but also to uncertainty with respect to the evolution of several unknown market parameters that characterize the behavior of the fund portfolio. To be more precise, due to the stochastic character of the benefits of a DC plan, in combination with the fact that the investment horizon of such schemes is usually large, the return from such an investment is highly exposed to both microeconomic and macroeconomic factors that affect the behavior of the underlying financial variables (e.g., bonds, stocks, various financial derivatives, etc.) which compose the fund portfolio. Therefore, it is obvious that the proposed problem is a highly complicated one, that requires a very delicate approach in order to exploit its qualitative and quantitative nature. More specifically, we are going to investigate the following main questions:

- What is the optimal investment strategy for a DC pension fund scheme if we consider the stochastic nature of all the (known) underlying variables? Namely, among others, we will consider the case of: (a) stochastic interest rates, (b) stochastic volatility, (c) stochastic salaries, and (d) combination of the above.

- What is the sensitivity of the optimal investment strategy for the DC pension fund scheme with respect to the fund manager's risk "appetite"?
- Given the fact that many of the parameters involved are subject to uncertainty, what is the optimal investment decision for a DC pension fund portfolio within the worst-case scenario for the underlying financial market?

The above questions are of crucial importance and must be effectively addressed in order to make the best decision possible at each instant of time for the fund portfolio. In fact, due to their importance, each one of them constitutes a research area on its own and dictates for an elegant approach (from both a quantitative and qualitative point of view). From a general standpoint, the first two questions constitute a natural extension of the available relative literature and are mainly risk-oriented. That is, we are mainly interested in providing the very best decision possible, at each instant of time, by solely focusing on the exposure of the fund portfolio to the several underlying market risks, like, e.g., volatility risk, interest rate risk, inflation risk, etc. To be more precise, (Q1) focuses on the stochastic character of the financial variables that govern the evolution of the fund portfolio, with the most prominent ones, being, volatility of asset returns and interest rates. Even though this problem has been studied in the relative literature (mainly, in the continuous time framework) our proposal aims to move a step further by adopting a discrete time framework, in order to provide an as much as realistic setting. Furthermore, we will also consider combinations of the above stochastic factors, in order to provide practitioners with an (as much as possible) realistic model that can be adopted in a wide variety of cases. This stage also requires the calibration of the results according to real financial market data and appropriate testing. (Q2) has its origins to enterprise risk management and aims to touch the problem of decision making by incorporating decision maker's preferences towards risk, a subject that has been highlighted by many authors (cf., e.g., Breen et al. [14], Maringer [45] and references therein) as being of crucial importance within an expected Von Neumann-Morgenstern utility maximization framework. As most of the problems studied within the pension fund management framework are basically utility maximization problems, it is of great interest to examine the sensitivity of the results with respect to various utility functions, i.e., with respect to various attitudes towards risk. Finally, (Q3) (which, to the best of our knowledge, has only been partially studied in the relative literature, see, e.g., Sun et al. [56]) aims to study the above problems from a completely different point of view, by introducing uncertainty (in the Knightian sense) rather than risk in the model at hand. This approach has its origins in game theory and aims to place the first stones towards realistic risk management.

An outline of this chapter is as follows. In Sect. 2, we present the general framework in order to introduce model uncertainty aspects to a stochastic control problem. Although this setting is well known, the presentation adopted here is DC pension fund management—oriented and will lay the ground for our future research. In Sect. 3, as a illustration, we present a simple stochastic model for the optimal management of DC pension funds under risk and uncertainty. Finally, in Sect. 4, we focus on the

design of a pension fund scheme, from a risk management point of view, presenting a detailed study concerning the case of Poland.

2 Optimal Management of Defined Contribution Pension Funds Under Model Uncertainty

Stochastic optimal control is an important branch of Mathematics that has found interesting applications in a wide variety of fields, such as, mathematical economics, mathematical finance, actuarial mathematics and modern portfolio risk management (see, among others, Akume and Weber [1], Baltas and Yannacopoulos [6], Browne [18], Duarte et al. [29], Savku et al. [55] and references therein). More precisely, its use has contributed a lot in mathematical modeling in Finance and has led to a deep understanding of the interplay of various sources of risk in optimal portfolio management (see, e.g., Merton [49]). However, despite its importance and wide range of applicability, within this framework it is tacitly assumed that the decision maker has complete faith in the model she faces, in the sense that the exact probability law of the stochastic risk factors in the underlying model, is precisely known. As it turns out (see, e.g., Anderson et al. [2, 3], Hansen and Sargent [38] and references therein) this assumption is far from being realistic, as often more than one possible models can be plausible representations for the system at hand, compatible with the observed data.

A very convenient way to introduce model uncertainty aspects to a stochastic control problem, is to let the decision maker distrust the model she faces. Even though this idea is rather simple, it is quite effective and constitutes the cornerstone of realistic, robust modeling. To be more precise, let us consider a fund portfolio manager who, at each instant of time, is responsible for optimally distributing the accumulated wealth (that stems from pooled contributions of the DC fund members) among several risky assets (e.g., bonds, stocks, derivatives, etc.), according, of course, to certain investment rules. The fund manager is responsible for making all the necessary investment decisions that will yield the highest return, or in mathematical terms, to drive the fund portfolio (whose wealth is subject to random fluctuations modeled by a number of risk factors) to a desired optimal state. In the language of stochastic control, the fund manager chooses the control process. In this framework, optimality is usually meant in the sense of maximizing the portfolio returns (of course other equivalent goals are also possible, e.g., maximizing the expected utility of its terminal wealth for some appropriately defined utility function, or minimizing a penalized distance of the portfolio's terminal wealth from a predefined goal). In order to introduce model uncertainty aspects to this problem, we let the fund portfolio manager, who controls the evolution of the system by selecting the associated control process (that is, the proportion of the fund's wealth to be distributed among the several risky assets), to question its validity as an appropriate model to describe the future states of the world. The basic philosophy of this specific form of uncertainty

is the structural assumption that there exists a "true" benchmark probability model (which is represented by a probability measure $\mathbb{P}$) related to the exact law of the process that introduces stochasticity in the system (e.g., a driving multi-dimensional Brownian motion), the fund manager is unaware of, and a probability model (which is represented by a probability measure $\mathbb{Q}$, not necessarily coinciding with $\mathbb{P}$) which is the fund manager's "idea" about how this "true" law in fact looks like. Clearly, the optimal decision rule depends on the underlying model, of which, the fund manager is uncertain (this risk, is often referred to, as model risk). Within this framework, the fund manager faces again the initial optimal control problem (e.g., a mean variance portfolio optimization problem) but this problem is now considered over the worst possible scenario, that is, by using the probability model that may create the most unfavorable scenario for the problem at hand.

From a mathematical standpoint, the above situation raises many challenges and dictates for a delicate approach in order to be effectively treated. In this vein, the classical techniques of stochastic optimal control are augmented with model selection techniques, resulting to what is widely known as robust optimal control theory. Technically speaking, robust control has close connections with game theory. In fact, a robust control problem can always be restated as a two player (in its simplest form) zero sum stochastic differential game. Within the DC fund management framework adopted in our research, the first player is the fund manager (the decision maker) and the other one is, a fictitious malevolent player, called Nature. Of course, this setting can be easily relaxed to cover many different situations of great interest (e.g., to consider competition between two different fund management companies). Under a measure change framework, the fund manager chooses the control process (that is, the investment policy) so as to drive the underlying system to a desired state. On the other hand, Nature, antagonistically chooses the probability model in order to create the most unfavorable scenario for the fund manager. Therefore, the two players engage in a game, whose Nash equilibrium corresponds to the optimal robust decision. In the context of continuous time diffusion processes and restricting ourselves to a class of measures which have exponential density with respect to the Lebesgue measure, a direct application of the celebrated Girsanov's theorem leads to the reformulation of the above game, as a stochastic differential game, amenable to the powerful tools of dynamic programming.

In general, by employing the classical techniques of dynamic programming, the value function of a stochastic differential game is associated with a second order partial differential equation, known as the Bellman-Isaacs (BI) equation. This is a highly nonlinear equation (for most problems of interest) and, as expected, classical (smooth) solutions are found only in some special cases. In fact, as it turns out, the most natural concept of a solution for this equation, is the notion of viscosity solutions (for more information on this subject, the interested reader is referred to Crandall et al. [22], Fleming and Souganidis [32] and references therein). Within this framework, even though the derivation of the optimal controls of the two players (which may act as a useful benchmark for the fund manager when looking to decide for the optimal investment strategy) is not possible, we can shed light into the underlying mechanisms that govern the evolution of the pension fund stochastic system. On the

other hand, under the additional assumption of smoothness of the value function of the game, it is possible to provide some closed form solutions for the BI equation (and as a result, to derive, in closed form, the optimal strategies for both players). This approach has been heavily used with success by many authors and within a wide range of robust control applications (different from the one proposed here); see, e.g., Baltas and Yannacopoulos [6, 7], Brock et al. [16], Branger et al. [13], Flor and Larsen [33], Kara et al. [42], Mataramvura and Øksendal [47], Pinar [52], Rieder and Wopperer [53], Zawisza [58], and references therein.

Concerning the application of robust control techniques to the problem of optimal management of DC pension funds, there exists only a limited body of work (see, e.g., among others, Sun et al. [56]). Our research aims to fill this gap by: (a) adopting a discrete-time framework, (b) making more realistic assumptions concerning the structural characteristics of the associated robust control problem, most of which stem from the new Polish pension system, and (c) provide a detailed numerical study for the problem at hand. It is our strong belief, that our results, apart from the mathematical interest, will also be very useful for pension fund managers, from both an investment and a risk management point of view.

3 A Simple Model

In the present section, as an illustration of the above theoretical discussion, we present a simple model for the optimal management of DC pension funds under uncertainty. First of all, we need to define the underlying probability space that will lay the ground for the definition of our stochastic financial market. In this vein, let us consider the filtered probability basis $(\Omega, \mathbb{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ that satisfies the usual hypotheses, where $\mathscr{F}_t = \sigma(W_1(s), W_2(s);\ s \leq t)$ is the natural filtration induced by the standard independent Brownian motions $(W_1(t);\ t \geq 0)$ and $(W_2(t);\ t \geq 0)$.

3.1 The Financial Market

We adopt a continuous time model for the financial market on the fixed finite time horizon $[0, T]$, with $T \in (0, \infty)$, consisting of the following investment opportunities:

- A zero coupon bond with maturity $T > 0$ and dynamics described by

$$\begin{aligned} \frac{dP(t,T)}{P(t,T)} &= (r + \alpha\theta)dt + \alpha dW_2(t), \\ P(0,T) &> 0, \end{aligned} \tag{1}$$

where $P(t, T)$ denotes the price of the bond at time $t \in [0, T]$. Here, $r > 0$ and $\alpha > 0$ stand, respectively, for the interest rate and the volatility of bond prices (and are assumed, for simplicity reasons, to be constants—case of flat interest rate). Moreover, $\alpha\theta$ (for some $\theta > 0$) stands for the excess return on the bond.

- Another risky asset (e.g., a financial index or stock) which evolves according to the stochastic differential equation

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu dt + \sigma dW_1(t), \\ S(0) &= S_0 > 0, \end{aligned} \tag{2}$$

where $S(t)$ denotes the price of the index at time $t \in [0, T]$. Here $\mu > r > 0$ stands for the appreciation rate of the stock prices and $\sigma > 0$ stands for the volatility of the stock prices.

- A risk free asset (bank account) with unit price $B(t)$ at time $t \in [0, T]$ and dynamics described by the ordinary differential equation

$$\begin{aligned} dB(t) &= rB(t)dt, \\ B(0) &= 1. \end{aligned} \tag{3}$$

It has to be pointed out, that as the number of traded assets on the market (zero-coupon bond and stock) equals the number of sources of noise (the $(\mathbb{F}, \mathbb{P})$-Brownian motions $(W_1(t);\ t \geq 0)$ and $(W_2(t);\ t \geq 0)$), the market is complete. As a result, we have placed ourselves within a very convenient framework in order to demonstrate our robust approach to the problem of optimal management of DC pension funds.

3.2 The Stochastic Salary

Salaries of the contributors are in general stochastic, in the sense that it is not possible today to know with certainty their level after such a large time interval (due to e.g., impossible to predict external both macroeconomic and micro economic factors). As a result, in the present work we consider the stochastic process $(L(t);\ t \geq 0)$ that denotes the average salary at time $t \in [0, T]$. Furthermore, we assume that this process evolves according to the stochastic differential equation

$$\begin{aligned} \frac{dL(t)}{L(t)} &= \mu_L(r)dt + k_2 dW_1(t) + k_3 dW_2(t), \\ L(0) &= l_0 > 0, \end{aligned} \tag{4}$$

where $l_0 \in \mathbb{R}_+$ denotes the initial average salary level. In the above equation, $\mu_L(r)$ may be considered as the expected instantaneous growth rate of the average salaries and is considered to be a function of the interest rate (see, e.g., Bat-

tocchio and Menoncin [9], Zhang et al. [60] and references therein), whereas the terms $\int_0^t k_2 dW_1(s)$ and $\int_0^t k_2 dW_2(s)$ may be considered as the fluctuations around this growth rate (two sources of uncertainty stemming from the stochastic interest rates and the stochastic volatility). More precisely, $k_1, k_2 \in \mathbb{R}$ are appropriate constants (scaling factors), that aim to capture the effect that stochastic interest rates and stochastic volatility have on the evolution of the average salary.

3.3 Contributions

According to the defined contribution pension scheme that is employed in the present section, employees that become part of the pension fund under consideration (at time $t = 0$) have to pay contributions. These contributions, according to various pension funds schemes (e.g. the new Polish pension scheme), are defined as proportion of their salary (in fact, this is an assumption that has been heavily used with success in the relative literature; see e.g., Korn et al. [43], and references therein). In what follows, we deal with a specific class of employees (for now—on future works we will consider many different classes) that share the same structural characteristics (these characteristics might be, for example, profession, years of experience, education level, etc.). We also let the contributions to be paid continuously. In this vein, the term $qL(t)$ denotes the aggregate contributions up to time $t \in [0, T]$, where q stands for the average contribution rate and, as already stated before, $L(t)$ stands for the average salary of the class under consideration. A natural restriction is to let $0 < q < 1$: The inequality $q > 0$ means that every member has to contribute something to become part of the fund, while, on the other hand, the inequality $q < 1$ means that the maximum average contribution is less that the average salary.

3.4 Stochastic Differential Equation for the Fund's Wealth

We envision a fund manager, who, at time $t = 0$, is endowed with some initial wealth $x > 0$ and whose actions cannot affect the market prices. The portfolio process $\pi(t) = \pi(t, \omega) : [0, T] \times \Omega \to \Pi_1 \subset \mathbb{R}$ denotes the proportion of of the fund's wealth $X(t)$ invested in the stock and the process $b(t) = b(t, \omega) : [0, T] \times \Omega \to \Pi_2 \subset \mathbb{R}$ denotes the proportion of of the fund's wealth $X(t)$ invested in the zero coupon bond at time $t \in [0, T]$. The remaining proportion $(1 - \pi(t) - b(t))X(t)$ is invested in the remaining asset (bank account). Here, Π_1, Π_2 are fixed closed and convex subsets of $\mathbb{R}$; typically compact. As a result, the wealth process corresponding to the strategy $(\pi(t), b(t))$, is defined as the solution of the following stochastic differential equation

$$dX(t) = \pi(t)X(t)\frac{dS(t)}{S(t)} + b(t)X(t)\frac{dP(t,T)}{P(t,T)} + (1-\pi(t)-b(t))\,X(t)\frac{dB(t)}{B(t)} + qL(t)dt.$$

Therefore, and referring to the initial wealth as x:

$$\begin{aligned} dX(t) &= ([r+\pi(t)(\mu-r)+\alpha\theta b(t)]\,X(t)+qL(t))\,dt \\ &\quad + \sigma\pi(t)X(t)dW_1(t) + \alpha b(t)X(t)dW_2(t), \\ X(0) &= x > 0. \end{aligned} \tag{5}$$

Definition 1 Let $\mathbb{F}$ be a general filtration. We denote by $\mathcal{A}(\mathbb{F}; T)$ the class of admissible strategies $(\pi(t), b(t))$ that satisfy the following conditions:

(i) $\pi(t), b(t)$ are progressively measurable mappings with respect to the filtration $\mathbb{F}$;
(ii) $0 \le \pi(t) \le 1$ and $0 \le b(t) \le 1$;
(iii) $\mathbb{E}\left[\int_0^T (\sigma\pi(t))^2\,dt\right] < \infty$ and $\mathbb{E}\left[\int_0^T (\alpha b(t))^2\,dt\right] < \infty$, $\mathbb{P}$-a.s.;
(iv) The SDE (5) admits a unique strong solution, denoted by $X(t)$.

3.5 Model Uncertainty Concerns

The aim of the present section is to place the problem of optimal DC pension fund management within a model uncertainty framework. In fact, this is the reason we chose to work within such a simple framework, as it allows to focus more on model uncertainty aspects. Of course, this work can be extended in a handful of ways. In order to fulfill our goal, we curry on our program by assuming that the portfolio fund manager is uncertain as to the true nature of the stochastic processes W_1 and W_2 that drive uncertainty into the state Eq. (5) that describes fund's wealth at time $t > 0$, in the sense that the exact law for W_1 and W_2 is not known. In an attempt to adopt a measure theoretic framework, we furthermore assume that there exists an unknown drift process $u(t)$ related to the Brownian motion W_1 and an unknown drift process $\lambda(t)$ related to the Brownian motion W_2. Of course, this is not the most general way to introduce uncertainty to our model (or, better stated, concerning the stochastic processes W_1 and W_2); however, it provides a simple, well-known and heavily studied framework to effectively study the problem at hand (see, e.g., Baltas and Yannacopoulos [5]). This is equivalent to state (thanks to Girsanov's theorem) that there exists a "true" probability measure related to the true law of the processes W_1 and W_2, the fund manager is unaware of, and a probability measure Q, which is the manager's idea of what the exact law of W_1 and W_2 looks like.

In the present section, we assume that the fund manager faces an expected utility maximization problem. However, this problem is considered under the probability measure Q (which is the manager's idea of the future states of the world). In other words, the fund manager seeks to solve the optimal control problem

$$\sup_{(\pi,b)\in\mathscr{A}^{\mathbb{F}}} \mathbb{E}_Q\Big[U(X(T))\Big|\mathscr{F}_t\Big],$$

for some appropriately defined utility function $U(\cdot)$. However, as the manager is uncertain about the validity of Q as the appropriate way to describe the future (states of the world) evolution of state equation (5), she seeks to adopt a more careful (i.e., robust) approach, that of seeking to minimize the worst possible scenario concerning the true description of the noise terms. In other words, she seeks to maximize the minimum possible value of the expected utility over all possible scenarios concerning the true state of the system, which is quantified as

$$\inf_{Q\in\mathscr{Q}} \mathbb{E}_Q\Big[U(X(T))\Big|\mathscr{F}_t\Big],$$

where $\mathscr{Q}$ is an appropriate class of probability measure. Putting all these together, the risk manager faces the robust stochastic optimal control problem

$$\sup_{(\pi,b)\in\mathscr{A}^{\mathbb{F}}} \inf_{Q\in\mathscr{Q}} \mathbb{E}_Q\Big[U(X(T))\Big|\mathscr{F}_t\Big], \tag{6}$$

subject ot the state process (5).

Definition 2 (The set $\mathscr{Q}$) The set of acceptable probability measures $\mathscr{Q}$ for the agent is a set enjoying the following two properties:

(i). We will only consider the class of measures $\mathscr{Q}$, such that considering the stochastic process W under the reference probability measure $\mathbb{P}$ and under the probability measure $\mathscr{Q}$ results to a change of drift to the Brownian motion W.

(ii). There is a maximum allowed deviation of the managers measure $\mathscr{Q}$ from the reference measure $\mathbb{P}$. In other words, the manager is not allowed to freely choose between various probability models as every departure will be penalized by an appropriately defined penalty function, a special case of which is the Kullback-Leibler relative entropy $\mathscr{H}(\mathbb{P}|\mathscr{Q})$.

The above two conditions specify the set $\mathscr{Q}$. Here, we briefly discuss the mathematical details surrounding the model uncertainty framework adopted in the present section. For more information, the interested reader is referred to Anderson et.al [2, 3], Hansen and Sargent [38] and references therein. To be more precise, let us consider the progressively measurable and square integrable stochastic processes $(u,\lambda) := \big((u(t),\lambda(t))\,;\ t\in[0,T]\big)$ taking values in some compact and convex set $\mathscr{Y} := \mathscr{Y}_1\times\mathscr{Y}_2 \subset \mathbb{R}^2$. Each one of these processes may be considered as an unknown drift process related to the Brownian motions $(W_1(t);\ t\geq 0)$ and $(W_2(t);\ t\geq 0)$. We define the probability measure Q on on $\mathscr{F}_T$ as $dQ/d\mathbb{P} := \mathscr{E}\big(\int_0 u(t)dW_1(t) + \int_0 \lambda(t)dW_2(t)\big)_T$, where $\mathscr{E}\big(M\big)_t := \exp\big(M(t) - \langle M\rangle_t/2\big)$ denotes the Doléans-Dade exponential of a continuous local martingale $M(t)$. Under the additional assumption

that the processes (u, λ) satisfy Novikov's integrability condition, it follows from the celebrated Girsanov's theorem that the processes $\widetilde{W}_1(t) = W_1(t) - \int_0^t u(s)ds$ and $\widetilde{W}_2(t) = W_2(t) - \int_0^t \lambda(s)ds$ are $(\mathbb{F}, Q)$-Brownian motions. In this vein, we have the following result.

In what follows, we define the new state variable $Y(t)$ as

$$Y(t) = \frac{\widetilde{X}(t)}{\widetilde{L}(t)}, \tag{7}$$

where $\widetilde{X}(t)$ and $\widetilde{L}(t)$ denote the fund's wealth process and the stochastic salary process, under the equivalent probability measure. This new variable is often referred to, as relative wealth (see, e.g., Zhang and Rong [59]). In fact, this new variable stands for a measure of attractiveness of the pension fund scheme, as it allows to compare with the true current financial situation of the fund members.

Proposition 1 *The relative wealth process under the equivalent probability measure $Q \in \mathcal{Q}$ is denoted as $Y(t)$ and is defined as the solution of the following stochastic differential equation:*

$$\begin{aligned} dY(t) = &\Big[r + (\mu - r - \sigma k_2)\pi(t) + (\theta - k_3)\alpha b(t) + (\sigma\pi(t) - k_2)u(t) \\ &+ (\alpha b(t) - k_3)\lambda(t) - \left(\mu_L(r) - k_2^2 - k_3^2\right)\Big]Y(t)dt + qdt \\ &+ (\sigma\pi(t) - k_2)Y(t)d\widetilde{W}_1(t) + (\alpha b(t) - k_3)Y(t)d\widetilde{W}_2(t), \end{aligned} \tag{8}$$

with initial condition $Y(0) = y > 0$.

Proof The proof follows immediately by substituting the semimartignale decompositions $\widetilde{W}_1(t) = W_1(t) - \int_0^t u(s)ds$ and $\widetilde{W}_2(t) = W_2(t) - \int_0^t \lambda(s)ds$ in Eqs. (4) and (5) and by a straightforward application of the quotient Itô rule on (7).

As a result, from now on we consider the robust control problem

$$\begin{aligned} &\sup_{(\pi,b)\in\mathcal{A}^{\mathbb{F}}} \inf_{Q\in\mathcal{Q}} J(t, y) \\ &= \sup_{(\pi,b)\in\mathcal{A}^{\mathbb{F}}} \inf_{(u,\lambda)\in\mathcal{U}} \mathbb{E}_Q\left[U\left(Y(T)\right) + \frac{1}{2\beta}\int_t^T \left(u^2(s) + \lambda^2(s)\right)ds\right], \end{aligned} \tag{9}$$

subject to the state dynamics (8). The intuition behind the robust control problem (9) is that we do not allow the fund manager to freely choose between different probability measures, that is, we prevent the fund manager from considering models that deviate "too much" from the reference model. This is accomplished by constraining this choice by an appropriate penalty function, namely, the Kullback-Leibler divergence, which, in our case is defined as $\mathbb{E}_Q\left(\frac{1}{2}\int_0^T \left(u^2(t) + \lambda^2(t)\right)dt\right)$. Moreover, this penalty function is weighted by the term $1/\beta$. The positive constant β is referred

to as the preference for robustness parameter, and serves as a measure to quantify the preference for robustness of the fund manager. In fact, concerning the possible allowed values for this parameter, there exist two interesting limiting cases:

- $\beta \to 0$. In this case, the fund manager fully trusts the model she is offered and seeks no robustness. In this case, the robust stochastic optimal control problem (9) reduces to the (simple) stochastic optimal control problem:

$$\sup_{(\pi,b)\in\mathscr{A}^{\mathbb{F}}} \mathbb{E}_{\mathbb{P}}\Big[U(Y(T))\Big|\mathscr{F}_t\Big], \tag{10}$$

  subject to the state dynamics

$$\begin{aligned} dY(t) = &\Big[r + (\mu - r - \sigma k_2)\pi(t) + (\theta - k_3)\alpha b(t) - \mu_L(r) + k_2^2 + k_3^2\Big]Y(t)dt \\ &+ q\,dt + (\sigma\pi(t) - k_2)Y(t)dW_1(t) + (\alpha b(t) - k_3)Y(t)dW_2(t), \end{aligned} \tag{11}$$

  with initial condition $Y(0) = y > 0$. This problem can be easily addressed by adopting dynamic programming techniques.
- $\beta \to \infty$. In this case, the fund manager has no faith in the model she faces and seeks alternative models with larger entropy. However, it has to be pointed out that this case is not easily treatable as the inner minimization problem becomes undefined (since it looses its convex character). For more information on this subject, see, e.g., Baltas et al. [5] for a detailed study of a related robust control problem, in this limit.

Remark 1 The robust optimal control problem (9) is in the form of a two player, zero-sum stochastic differential game. The first player is the fund manager (player I) and the second player (player II) is a fictitious adversarial agent, commonly referred to, as Nature. Player I, is endowed with some initial wealth and decides the optimal proportion of the fund's wealth to be invested in the stock and bond markets, as well as, the bank account. On the other hand, player II antagonistically chooses the probability measure $Q \in \mathscr{Q}$ (that is the probability model, as there exists a one to one relationship between a probability model and a probability measure) in order to create the worst possible scenario for the fund manager.

In the present section, we constructed a simple stochastic model that introduces uncertainty aspects to the problem of optimal management of DC pension funds. In order to solve the problem (9) subject to the dynamic constraint (8), one has to derive the associated Bellman-Isaacs (BI) equation and solve it, as the optimal controls are defined as functions of the derivative of this solution. However, as already mentioned in Sect. 2, at this stage a major obstacle arises, as it is not in general possible to find a smooth solution to the BI. This is an area that our future research aims to contribute by following: (a) a weak solution point of view, and (b) an algorithmic approach to numerically solve the BI. Of course, the model presented here can be extended in a variety of ways (e.g., consider the case of stochastic volatility, stochastic interest rates,

the effect of mortality, etc.) but each one of these extension comes with an associated complexity cost. Finally, from a quantitative point of view (and this is another future focus point of our group), the stochastic models adopted in the present section (and the ones that will be used in our future endeavors), will be calibrated to the specific case of the new Polish pension scheme, in order to make our results as realistic as possible.

4 Risk Distribution and Design of Pension Scheme: A Case Study of Poland

4.1 Design of Polish Pension Scheme from the Point of View of Risk Management

Poland introduced a comprehensive reform of its old-age pension system in 1999. The reform established a defined-contribution, multi-pillar system involving: a PAYG pillar based on notional (non-financial) defined contributions (NDC) and administered by the Social Insurance Institution (in Polish: Zaklad Ubezpieczeń Spolecznych; in short: ZUS), a mandatory funded pillar in which private pension funds manage individuals' contributions, and a voluntary third pillar consisting of company pension plans and other savings vehicles (cf. Table 1[2]).

The total pension contribution rate amounts to 19.52% of gross wages (by along pillar 1 and pillar 2). The contributions (premiums) are financed equally by both employer and employee. In fact, 16.60% of pension contributions are transferred to pillar 1 (being written down on the individual accounts and sub-accounts of those

Table 1 The architecture (design) of the three-pillar Polish pension system

Pillar 1	Pillar 2	Pillar 3
Mandatory	Mandatory/Voluntary	Voluntary
PAYG	Funded	Funded
Basic pension benefit	Basic pension benefit	Additional/supplementary pension benefit
Notional Defined Contribution (NDC)	Defined Contribution (DC)	Defined Contribution (DC)
Managed by public institution: Social Insurance Institution (ZUS)	Privately managed: open pension funds (OFE's) managed by Pension Fund Societies (PFSs)	Privately managed: individual and group (occupational) pension savings, managed by different financial institutions

[2] Open Pension Funds (in Polish: Otwarte Fundusze Emerytalne; in short: OFE) were introduced in 1999 and have been obligatory since 1999. As of 1 April 2014, they are voluntary. The role of the pillar 2 has been marginalized. Source: own elaboration.

insured) and 2.92%[3] goes to open pension funds (pillar 2), if the insured person is a member of an OFE. If not, the entire 19.52% are transferred to the pillar 1 (see Rutecka [54], p. 130). The notional interest rate is defined as 100% of the growth of the real covered wage bill, and no less than the price of inflation. The pillar 2 is a voluntary funded defined contribution (FDC) scheme. Contributions paid into the second pillar are indexed with the rate of return on pension fund investments. One of its main objectives in the economic dimension was the division of risk between the financial market and the labor market by introducing a three-pillar structure and, in particular the second capital funded pillar and OFEs operating within it (see Góra [35]). After retirement (in the deccumulation phase of a pension system), pension benefits are indexed annually by inflation with at least 20% of the real average wage growth. The pension formula is to a large extent similar to the first and the second pillar. Benefits are equal to the accumulated capital from contributions (plus indexation) divided by life expectancy, obtained from the observed unisex period mortality tables. Mortality tables are recalculated by Polish Central Statistical Office (in short: GUS) every year.

The assumptions of the systemic pension reform introduced in Poland in 1999 predicted the development of additional voluntary pension schemes ("the third pillar of the pension scheme"; see Table 1). The pillar 3 consists of voluntary and quasi-obligatory private pension plans:

- occupational pension plans (in Polish: "Pracownicze Programy Emerytalne"; in short: PPE),
- individual retirement accounts (in Polish: "Indywidualne Konta Emerytalne"; in short: IKE),
- individual retirement saving accounts (in Polish: "Indywidualne Konta Zabezpieczenia Em-erytalnego"; in short: IKZE).
- employee capital plans (in Polish: "Pracownicze Plany Kapitałowe"; in short: PPK)—new, quasi-obligatory occupational pension schemes, introduced in 2019; they have been introduced since the 1, July, 2019 first in big companies with 250 or more employees, than in small and medium companies in 2020, and in public sector in 2021.

In the years 1999–2004 (until the introduction of IKEs), the only form of saving for retirement, benefiting from certain (relatively modest) economic and fiscal incentives from the state, were PPEs. However, the current development of PPE in Poland has been very slow. Only a little bit more than 1,000 employers offer their employees the opportunity to participate in pension schemes (1 053 PPEs at the end of 2017, of which 645 in form of a contract with insurance company, 382 in the form of a contract with mutual investment fund and implemented with an employee pension fund, the so called "Pracownicze Fundusze Emerytalne"; in short: PFE). The number of participants at the end of 2017 was about 400 000 and total value of assets was 1.224,6 bln PLN (about 360 bln EUR). In this respect, Poland does not

[3] Initially, from 1999 to 2011, contributions to the 2nd pillar of the reformed public pension systems were much higher and were equal to 7.3% of gross wages.

compare favorably with other EU countries, including some former socialist states (e.g., Slovakia and the Czech Republic), where occupational pension schemes are more prevalent.

While pillar 1 (PAYG) is in the accumulation (savings) phase, the pension system is more sensitive to the risk of demographics which increases with the aging of the population, and the funded pillar in public system is subject to different (demographically uncorrelated) kinds of risk (including investment risk). Additional pension schemes (individual ones: IKE, IKZE, and occupational ones: PPE and PPK) with DC formula are exposed to investment risk in accumulation (savings) phase and to longevity risk in payout phase of the scheme (deccumulation of pension capital).

Due to different regulations concerning acceptable investment strategy and available financial instruments, the problem of optimal investment portfolio management must be analyzed differently in comparison to pension funds operating in the public pension system (OFE), and differently to additional pension systems which are individual (IKE, IKZE) and occupational pension schemes (PPE, PPK).

4.2 Analysis of Investment Policy and Risk Management in Open Pension Funds

A brief history of OFEs in Poland can be divided into two main stages. The first one took place in the years 1999–2014. During this period the contribution to OFEs was compulsory for every employee and covered by the general pension system. The second stage began in 2014 and continues to this day. Its essential feature is the optional nature of OFE. By default, every system participant pays a full mandatory pension contribution to ZUS, including a special ZUS sub-account, while participation in an OFE requires an opt-out decision.

At the beginning of this subsection, let us return to the end of the 1990s. At that time, the Polish parliament defined the principles of investment activity for OFE introducing the second mandatory fully funded pillar. In article 139, the following entry appears: The Fund places its assets in accordance with the provisions of this Act, striving to achieve the maximum level of security and profitability of investments made. Thus, already in 1997, it was emphasized that not only investment profitability is important, but also the risk closely related to the process of in-vesting the capital of future pensioners. Mazurek-Krasodomska [48] noted that stressing the security at the start of pillar 2 in Poland was justified by the peculiarity of OFEs, which consisted in the compulsory payment of contributions by future pensioners (see Mazurek-Krasodomska [48], p. 24).

In the above-mentioned Act, there were more rules that aimed at limiting the investment risk. These included provisions limiting the risks related to: OFEs operations, Pension Fund Societies (in Polish: “Powszechne Towarzystwa Emerytalne”; in short: PTE) operations, control over OFEs and PTEs, and investment policy. However, legal investment limits for OFEs changed signifi-cantly with the trans-

Table 2 Regulations limiting the investment risk. Source: own elaborations based on Act of 28 August 1997 and Jakubowski [39, 40]

Risk related to	Regulations
OFEs operations	Each PTE can manage only one OFE (exception: TFE takes over OFE management from another PTE or as a result of TFE's merger); watching over the security and legal compliance of transactions carried out by the OFE depository
PTEs operations	Legal and physical separation of pension fund from managing company; an obligation to pay into the Guarantee Fund, which ensures the interests of fund members
Control over OFEs and PTEs	Special supervision by Polish Financial Supervision Authority; Treasury as the last resort
Investment policy (until 2014)	No maximum limit for investments in treasury bonds and treasury bills; debt securities guaranteed or backed by the State Treasury or the National Bank of Poland and bonds issued by BGK (Bank Gospodarstwa Krajowego) were free of maximum limit; maximum investment limits for other income instruments; shares listed on a stock exchange could consist here 40% of OFE assets; maximum limit for foreign investments at the level of 5%; minimum required rate of return
Investment policy (since 2014)	No maximum level for equity allocations; in 2014 OFEs had to invest at least 75% of their assets in equity instruments; in 2015 the limit was 55%; in 2016 it was lowered to 35% and in 2017 to 15% in 2018; no minimum or maximum level for investments in shares of companies listed on regulated markets in Poland (since 2018); OFEs not allowed to invest in government bonds, treasury bills and other debt instruments issued or guaranteed by the State Treasury, National Bank of Poland, governments or central banks; maximum limit for foreign investment at the level of 30% (since 2016); elimination of the minimum required rate of return

fer of assets to ZUS in 2014 (cf. Table 2). Jakubowski [40] states that initially OFEs assets were managed like in balanced funds, now OFEs are managed just like equity funds (for more information, the interested reader is referred to Jakubowski [40], pp. 42–43).

As Czerwińska [23] observed, strategies implemented by OFE in the years 1999–2009 show the model of investment portfolio allocation −30% of shares and 65% of debt instruments (mainly government bonds). Thus, the shares of companies were only the second-largest category of investment instrument used by pension funds. In the indicated period, these shares accounted for 22% of the investment portfolio in 2008 (the lowest level) up to 35% in 2000 and 2007 (the highest level).

Czerwińska [23] explains that such a structure of a typical portfolio was determined to a decisive extent by: situation on the Polish financial market (shallow market with low liquidity), high sup-ply and high profitability of Treasury debt securities and unfavorable situations on the stock market in 2001–2002 and 2008, no restrictions

on investment in treasury securities, and regulations that limit both the concentration of fund capital in one company and activity on foreign financial markets.

However, as noted before, the conditions of OFEs activity changed very significantly in 2014. The Polish Financial Supervision Authority (in Polish: "Komisja Nadzoru Finansowego"; in short: KNF) has briefly summarized the main changes in one of its studies [Polish Financial Supervision Authority [4], pp. 13–16]. Among these changes, it should be state:

- reclassification of 51.5% of OFE members funds to the ZUS sub-account,
- enabling system participants to choose which institution will receive a part of the retirement contribution (ZUS or OFE),
- elimination of a minimum rate of return and reduction of fees,
- introduction of a so-called safety slider,
- strengthening pillar 3 (the possibility of additional savings on IKZE).

From the point of view of OFEs and TFEs, these changes meant completely new operating conditions and the need to adjust the investment policy. Firstly, a transfer of assets from OFEs to the ZUS cuts drastically the size of the pension market of Poland. Secondly, lower contribution paid to OFE[4] narrowed the capital inflow to these funds. Thirdly, introduction of freedom to pay contributions to OFEs led to a significant drop in the value of the contributions paid to pillar 2 (Jakubowski [40], pp. 44).

The evaluation of the results of OFEs investment activity is a very complex matter (cf. Chybalski [21]). For the purposes of this study, only the measure resulting directly from the 1997 Act and its subsequent amendments have been used. This indicator of OFE investment performance is the Weighted Average Rate of Return. Initially, it was calculated on the basis of changes in the value of OFE accounting unit in the 24-month period preceding the end of each quarter. Then, from the second quarter of 2004, the measurement period was extended to 36 months. In addition, it was decided to limit up to 15% of the weight that can be attributed to a single OFE (cf. Buchowiec [19], pp. 409–422). Figures 1 and 2 show the calculation of OFEs return rates made employing both of these methods. OFEs investment achievements turned out to be the largest in the first years of the new system operation. But in the long-term there was a clear downward trend.

[4] In this case, it should be clarified that the amount of the contribution transferred to OFE was actually reduced earlier because already in 2011. In the face of financial pressure caused by the slowdown of the Polish economy, the contribution to OFE was reduced from 7.3% to 2.3%. In the initial period the contribution was reduced from 7.3% to 2.3%. In subsequent years, a gradual increase to 3.5% was programmed. The remaining part of the contribution (from 7.3%, which was previously transferred to OFE) is recorded on the ZUS sub-account.

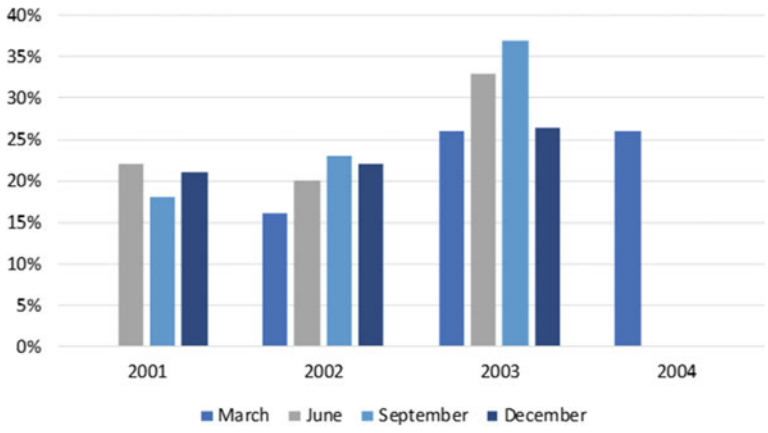


Fig. 1 OFEs rates of return for the 2-years period

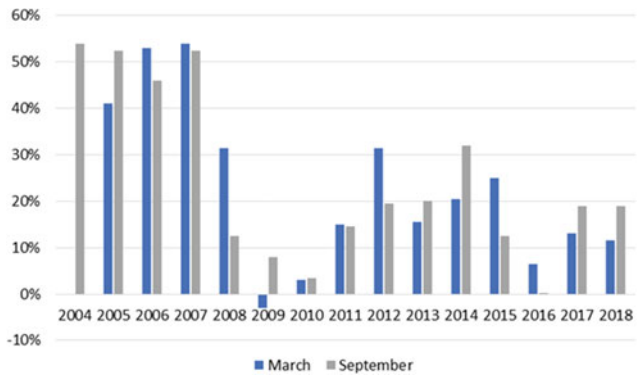


Fig. 2 OFEs rates of return for the 3-years period. *Source* data provided by the Polish Financial Supervision Authority

4.3 Risk in Occupational Pension Schemes

Defined contribution occupational pension schemes are often referred to as money purchase pension schemes, because in return for invested premiums, recorded in the personalized retirement accounts, the program participant will in future receive a certain cash benefit, whose value will depend on the contributions made and the result of investments. In this system, the amount of future benefits is not pre-defined and depends largely on the effects of investment in the financial market. Needless to say, also in a defined contribution scheme the amount of future pension is essentially determined by the value of contributions made to a pension scheme and can be predicted with reasonable accuracy. It requires certain financial simulations, making assumptions regarding the profitability of investments, situation of the national and

global economy, and in the financial markets, etc. However, the amount of future benefits is not guaranteed by the occupational pension contract in proportion to the remuneration, as in defined benefit schemes.

In many DC pension plans, payment of the "purchased" in this way of an occupational pension is divided into two parts. The first part is a one-time payment of a sum of gathered pension capital (lump sum), usually exempt from tax, and the second part forms a constant stream of payments (annuities) coming from the resources left in a pension fund or from benefits purchased with these resources, provided by third parties (mainly the insurance companies).

DC occupational pension schemes are usually fully funded. In such systems, the real money is invested in the financial markets [World Bank 1994: p. 172]. Very often the DC systems are managed by external financial institutions (insurance companies, mutual funds, banks).

In general, it is assumed that these systems are safer for an employer, who is not required to pay in future a pension of a predetermined value. Investment risk is largely passed on the employee. In exchange for risk allocation from employer to employee, program participants can count on certain benefits. They often have the choice between several investment funds and at least a partial influence on investment strategies (for example, what percentage of the premiums paid to a pension scheme is to be invested in stocks, bonds or other financial instruments). DC schemes are more transparent to employees, who can systematically track the status of their individual savings account in a given program. Typically, the DC schemes are not subject to so many restrictions regarding defined benefit schemes. However, there is a number of risks regarding DC equity funded pension schemes, which their participants are not always aware of (cf. Table 3).

In an unprotected pension plans there are no guarantees from either the pension fund or from a financial services provider as to the rate of return on investment, or other obligations regarding the entire pension plan. The protected pension plans, on the other hand, offer such guarantees. For example, the return on investment not lower than the yield of safe debt securities, or higher than the rate of inflation or other benchmark indicator.

In the further part of this study, the subject of analyzes will be the investment risk occurring in additional pension systems in Poland (namely, business and individual ones) with a defined premium (DC).

4.4 Analysis of Risk Management in Some of the PPEs in Poland

From the free forms of occupational pension schemes operating on the market since 1999, only statistical data on the value of participation units (after deducting service costs) of PFEs were published in a systematic manner and can be used to assess their investment performance (cf. Fig. 3). PPE in the form of an agreement with

Table 3 Types of risks associated with DC occupational pension schemes. *Source* Daykin [24], Oxera [51], and own elaboration

Type of risk	Characteristic
Market risk	The value of investments reported on the individual account of a pension scheme participant may fluctuate and decline significantly due to adverse financial market conditions (e.g., slump on the stock exchange)
Investment risk	Pension fund investment risk comes from three main sources: risk that the fund will fall in a value, risk that the pension fund's returns will not keep pace with inflation (real returns are negative), and risk that the pension fund does not perform well enough to keep pace with the growth in the cost of providing pension benefits
Economic risk	The real rates of return on investments (rate of return above inflation) may prove to be unsatisfactory due to the difficult conditions in the economy or bad economic policy, for example, due to inflation or low economic growth
Default risk	Investments made on behalf of the participants of an occupational pension scheme can bring effects later than proposed time limits of return on investment or lose value due to the financial difficulties of the institutions managing the program, such as an insurance company
Management risk	Managers may prove incompetent, and sometimes even commit a criminal of-fence while managing the fund
Interest rate risk	The value of the benefit which can be purchased with the sum of the accumulated premiums paid into the program and the interest on the capital investments will mainly depend on the interest rates on the financial markets at the time of ending the saving phase and converting accumulated savings into lifetime benefits (annuity)
Longevity risk	The increase of average life expectancy and the associated extension of the benefits receiving period are taken into account in the calculations (based on actuarial calculations) of financial institutions, in which life-time benefits are purchased with the resources accumulated in DC (e.g., insurance companies) and they have impact life expectancy for the population (cohort) of retirees
Operational risk	Managers of pension fund may lose the capability of adequate control at the operational level (current investing of funds gathered in the individual accounts of program participants). This phenomenon may be caused by a lack of necessary information following natural disasters, failure of an IT system or other random events
Insolvency risk	The company managing the pension program or providing pensions from the funds accumulated in the program (e.g., insurance company) may become insolvent and file for bankruptcy. Consequences of bankruptcy may vary for program participants, depending on the legal and institution solutions adopted in the country (in many countries there are legal protections—such as compulsory insurance of funds accumulated in the program, insurance from bankruptcy, reinsurance of insurance companies involved in the payment of annuities for participants of pension schemes, etc.)
Expense risk	The cost of administering pension scheme or accepted level of remuneration (commission) for the managing institution may prove too high and unfairly passed on fees collected from program participants' savings accounts (e.g., assets management fee, distribution fees)

(continued)

Table 3 (continued)

Type of risk	Characteristic
Fiscal risk	The government may change the tax rules for occupational pension schemes, reducing the rate of return on investment in such programs (e.g., withdrawing the previously granted tax incentives or reducing their level)
Regulatory risk	Institutions supervising the occupational pension schemes may not see in time the risk in the manner of managing of a particular pension fund or on the contrary—revoke the license of the management institution, causing perturbations to its participants
Political risk	The government may directly interfere with the operations of pension funds, change the rules of paying contributions, investing the collected funds, impose in-vesting in government debts or in economic undertakings generating lower rate of return than could be achieved when freely investing in the financial market

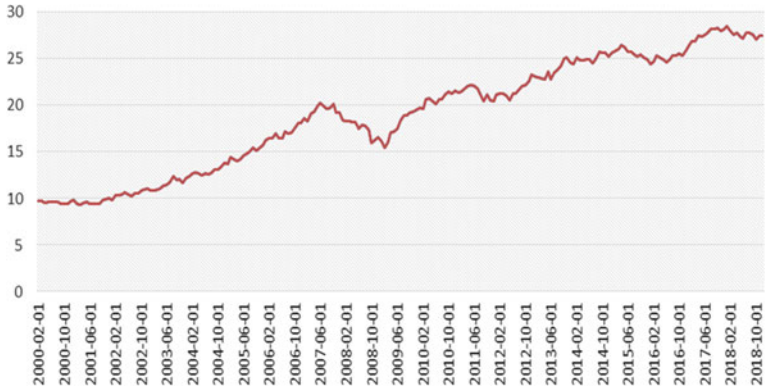


Fig. 3 The value of PFE "Nowy Świat" participation units from (2000–2018) in PLN. *Source* Polish Supervision Financial Authority

an investment fund or life insurance company were of individual nature, based on arrangements between financial services providers and with the companies' employers and the representation of employees (mostly trade unions). That and the service costs were not made public, and it is not possible to analyze precisely their rates of return.

The standard deviation of the rates of return on investments realized by PFE in the years 2000–2018 (cf. Fig. 4) was 8,16292. Only in 2008, at the peak of the global financial crisis, there was a negative double-digit rate of return (−13.5%). It is difficult to make it different, because during this period the value of financial assets dropped sharply in most countries of the world, while the value of assets of pension funds in many EU countries was higher than in Poland (e.g., in Ireland minus 30%). As early as in 2009, the value of participation units of PFE Nowy Świat has increased

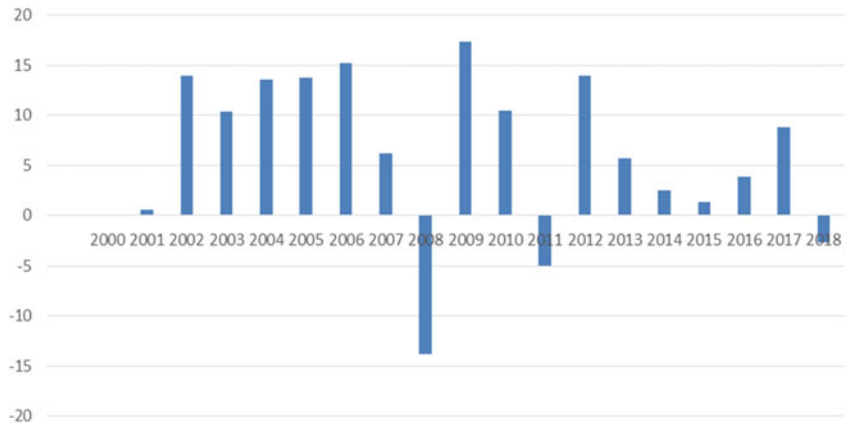


Fig. 4 Rates of return of investment in Employee Pension Fund (PFE) "Nowy Świat" 2000–2018. *Source* Polish Financial Supervision Authority 2018

enough to fully recuperate the 2008 losses (+15.2%). In the nearly twenty-year period of operation of the tested PFE, the rate of return on investments was negative also in 2011 and 2018 (−5% in 2011 and −2,2%), but it was much lower than in the timer of crisis in 2008. The recession from 2011 was connected with the Greek debt crisis, which strongly affected global equity markets, while 2018 correction was directly caused by US interest rates increase. A similar distribution of investment returns can be observed on the OFEs market in the public pension system 1999–2014. In this period, both PFEs and OFEs had similar portfolio structure, typical for mixed assets stable growth funds.

As a result of systemic legal changes and retreat from mandatory pension funds in public pension system,[5] the OFEs at the beginning of February 2014 became de facto funds of Polish shares and changed their status from obligatory to voluntary. At the end of March 2018, up to 12 OFEs had 16 million members, and their net assets amounted to PLN 166 billion. At that time, up to 2 employees of pension funds had 35 thousand members, and their assets amounted to PLN 1.8 billion. Up to 8 voluntary pension funds accounted for 99 thousand members and their assets amounted to PLN 316 million.

Due to the limited possibility of rebuilding OFE portfolios, in particular a prohibition on investing in OFE assets in government bonds, limited supply of instruments from other asset classes and a significant current involvement of OFE in shares listed on the WSE (Warsaw Stock Exchange), the investment profile of OFE is forced by the reform. At the end of March 2018, the share of domestic equity instruments in the OFE portfolio was 78%, the debt part constituted only 8%, bank deposits 7%, and

[5] The same process has been observed in other Eastern and Central European Union countries (Bielawska et al. [10]).

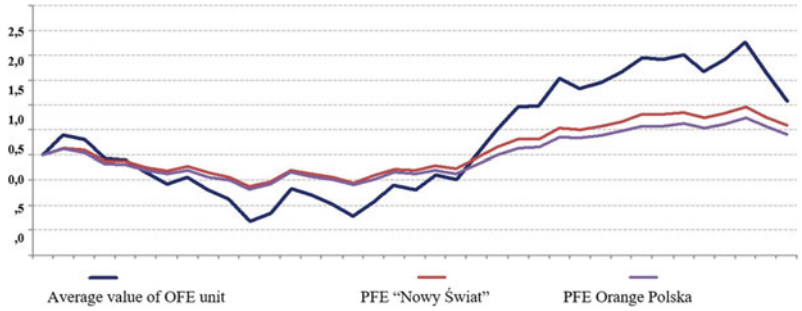


Fig. 5 Change of the value of settlement unit of the OFE and PFE (in points) from 31.12.2015 till 30.03.2018. Value from 31.03.2015 = 10. *Source* Polish Financial Supervision Authority 2018

investments in foreign securities 7%. This resulted in a greater variability in the return on investment realized by OFEs and by PFEs (cf. Fig. 5), while the non-diversified risk of investments in OFE is correspondingly higher. In the case of the PFE portfolio, the treasury bonds (66%) and shares listed on the WSE (26%) constituted the basis.

The PPEs operate in Poland nearly two decades. This is a relatively short time in the perspective of a professional career and saving for retirement, where a typical savings period (accumulation phase) is about 40 years. That is why the phenomenon of the so-called "a bad date"—the need to payout of occupational savings during the financial crisis, when the value of pension assets drops sharply—it only occurred to a very small group of employees paying their occupational pension in 2018 (less than 500 people according to data from KNF).

To avoid a similar situation ("bad-date" risk), the newly created employee capital plans will be defined-date funds (life-cycle funds). It should be adjusted to take into account the need to limit the level of investment risk as the participant approaches retirement age. In connection with the conclusion of a contract for the conduct of a PPK, funds collected by a participant may be placed in one of at least five funds of a defined date, applying different investment policy principles, appropriate for the date of birth of the participant (cf. Fig. 6).

At the participant's request, it is possible to change the fund of a defined date. This means that in the first phase of accumulating pension capital, most of the funds from contributions paid by the employee, employer and subsidies from the state budget will be invested in shares, while as the retirement age approaches, an increasingly larger share in the investment portfolio will have more secure debt securities, mostly treasury bonds and treasury bills.

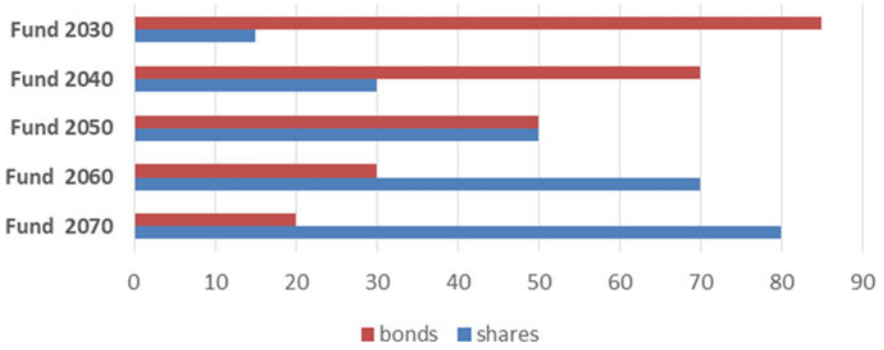


Fig. 6 The model of proportions of shares (equities) to bonds in investment portfolio of PPKs according to the law regulations (Article 40 of the Act on Employee Capital Plans). *Source* Instytut Emerytalny (Pensions Institute) Warsaw

4.5 Analysis of Investment Policy and Risk Management in Individual Additional Pension Schemes

In Poland, the first form of individual additional pension schemes were IKEs, which started to operate in 2004. IKEs were dedicated to people who cannot make saving for retirement using PPE.

IKE is a form of pension protection, which gives a wide spectrum of investment possibilities. Participants can choose from a variety of instruments, depending on their risk appetite, knowledge and available time for pension assets management. IKE can be conducted by five types of institutions: mutual funds, brokerage firms, insurance companies, banks and voluntary pension funds. Each member may have only one IKE and must be at least 16 years old to start saving for retirement. Participants can pay retirement contributions on a monthly, quarterly or annual basis. One of the important aspects of IKE is tax privilege. In Poland, income from financial instruments is taxable at 19%. In the case of IKE, the income is calculated as the difference between the sum of funds accumulated on the account and the sum of contributions made on them. Participants of IKE are exempted from this tax, if they meet the following criteria: pay contributions for at least five years and begin to pay out funds after reaching the age of 60. The maximum annual contribution to the IKE cannot be higher the sum of three average monthly salaries which was projected in the Polish national economy for a current year (cf. Fig. 7).

Another form of individual additional pension schemes are IKZEs, which were introduced in 2012. IKE and IKZE have quite similar principles of operation and may be conducted by the same type of institutions. However, IKZE and IKE differ from each other primarily by the annual contributions limit and the type of tax privilege for the account holder. The annual contributions to IKZE cannot exceed the amount equal to 1.2 times the average monthly salaries which was projected in the Polish national economy for a current year. Moreover, in the case of an IKZE, the

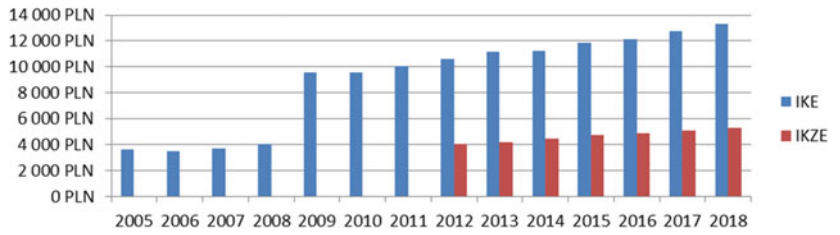


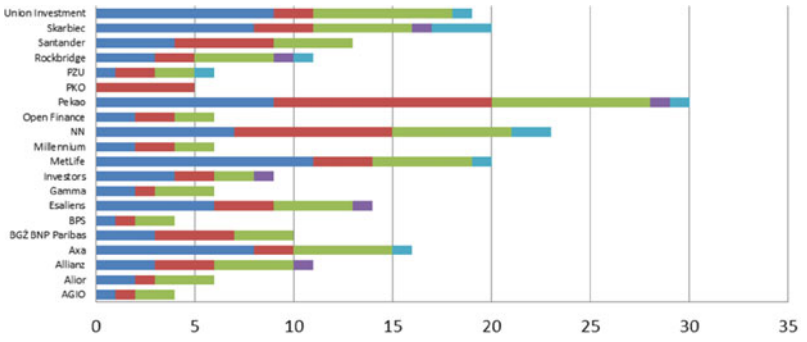
Fig. 7 Maximum allowed annual contributions in IKE and IKZE. *Source* Polish Ministry of Family, Labor and Social Policy 2018

tax privilege consists in deducting the sum of paid contributions from the personal income tax base.[6] The pension payment from IKZE takes place at the request of the account holder after reaching the age of 65. An additional condition is the payment of contributions for at least 5 calendar years. It is possible to withdraw the savings once or in installments.

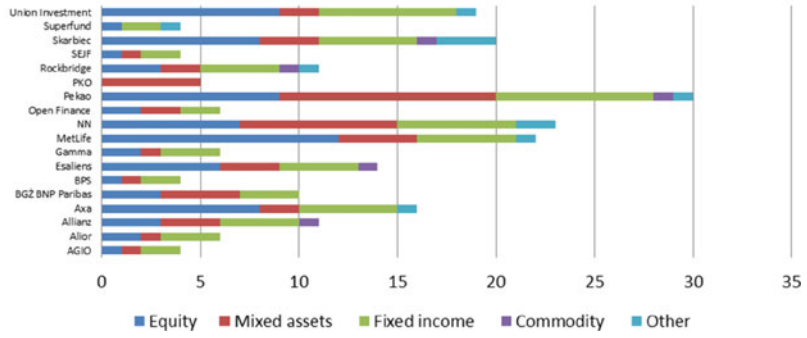
The selection of a financial institution, where IKE and IKZE are available, has a significant impact on the investment risk. Accounts maintained by banks take the form of simple and secure savings accounts, and their profitability depends only on the level of interest rates. Brokerage firms provide access to brokerage accounts; therefore, it is possible to invest directly in the capital market by individuals. In this case, each person has the opportunity to create an investment portfolio, which is tailored to individual needs and risk appetite. Saving in mutual funds, insurance companies and voluntary pension funds assumes the use of the idea of collective investment. The largest part of pension assets at IKE and IKZE is managed by mutual funds and insurance companies. In 2018, in the case of IKE, this was 31.0% and 30.9% of total assets, while in the case of IKZE, 39.5% and 35.0% respectively. In the majority of cases, insurance companies in Poland do not manage the retirement savings by themselves, but purchase mutual fund units available on the Polish market (Dopierala [27]). Moreover, the unit linked insurances are a pure DC pension plans. It follows that profitability and investment risks in IKE and IKZE depend mainly on the mutual fund portfolios. At the same time, savers take the entire investment risk in the case of all types of institutions.

Among the mutual funds operating within IKE and IKZE, the equity funds are the largest group (IKE 86 funds, IKZE 78 funds; cf. Fig. 8). In this group, there is a large diversity of applied investment strategies. Examples include funds investing in small and medium-sized companies, but also funds investing in large capitalization companies paying dividends. There is also a significant geographic diversity of asset placement. Although the Polish funds dominate, a group of abroad funds is also available. Among them are those that invest in developed markets as well as emerging markets. The funds investing in fixed income instruments are the second largest group (73 IKE funds, 67 IKZE funds), which invest mainly in the bond and money market. The bond funds invest both in government and corporate bonds in Poland and abroad.

[6] There are two PIT rates in Poland: 17% and 32%.



(a) Case of IKE



(b) Case of IKZE

Fig. 8 Number of mutual funds available under IKE and IKZE by company and by fund type. *Source* own elaboration based on: https://www.analizy.pl/ (access: 22.02.2018)

The mixed funds are also a large group (IKE 63 funds, IKZE 54 funds) investing in both shares and debt instruments. The mixed funds group includes maturity target funds (defined-date funds, life-cycle funds) in which the investment portfolio changes from aggressive to safe as the fund approaches the target date. For example, this type of funds are provided by the investment company Universal Savings Bank (in Polish: "Powszechna Kasa Oszczędności"; in short: PKO), which offer five funds with different maturity dates from 2020 to 2060. At the moment, the funds differ significantly in the volatility of investment results (cf. Fig. 9). Moreover, for the participants who save on IKE and IKZE it is also possible to choose the funds that invest in the commodity and alternative assets market.

In Poland, the forms of individual additional pension schemes operate under complex rules. It is possible to invest in various markets both individually and collectively using IKE and IKZE. In addition, financial institutions in which IKE and

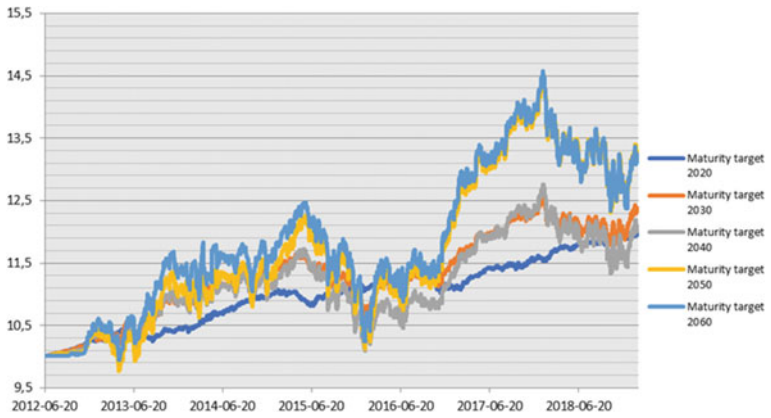


Fig. 9 Value of participation units of PKO maturity target funds available under IKE and IKZE (2012–2018) in PLN. *Source* own elaboration based on Refinitiv data available under an agreement between the University of Gdansk and Refinitiv

IKZE are available, offering the opportunity to choose from a wide range of investment strategies. However, the low level of financial knowledge of participants may lead to difficulties in choosing the best method of pension saving. Moreover, savers can use both IKE and IKZE in the same time. It follows that the construction of real investment portfolios depends on many factors. For this reason, the development of an optimal investment portfolio is an important issue.

4.6 Conclusions Concerning Managing Additional Pension Schemes in Poland

Neither PPEs nor PPKs guarantee a minimum rate of return or protection of capital from pension contributions. The same applies to individual forms of additional retirement savings—IKEs and IKZEs. These are typical funds with the DC formula, where the entire investment risk has been passed on to the program participant. Until now, there has been no discussion in the Polish literature on the subject of pension economics and finances regarding the optimal management of pension funds in conditions of risk and uncertainty.

5 Concluding Remarks

During the last decade, and especially after the financial (credit) crisis, the problem of providing supplementary pensions to the retirees has attracted a lot of attention (under the framework of the second and the third pillar, as outlined by the World Bank) from

official bodies, as well as private financial institutions (e.g., insurance companies and banks), worldwide. In this effort, there exist various possible solutions, one of which (in fact, the most popular) is provided by pension fund schemes. In plain words, a pension fund scheme represents accumulated wealth stemming from pooled contributions of its members. This accumulated wealth is collected in a portfolio of assets and is invested over a very long period of time in the financial market, in order to provide its members with retirement benefits. Hence, it is clear that the success of such a plan, heavily depends on the successful investment of the available accumulated funds. In fact, this remark places the problem of optimal pension fund management within a broader stochastic (due to the random character of the underlying financial variables) portfolio selection framework.

In the present chapter, we presented general ideas and preliminary results of our planned research project on defined contribution pension funds management, that is, on pension schemes according to which the contributions are fixed but benefits are unknown, as they depend on the performance of the fund portfolio. Our aim is to provide a detailed study for this problem, within an (as much as possible) realistic (stochastic) framework, by fully exploring its two different dimensions, to wit, risk and uncertainty. To be more precise, risk arises due to the exposure of the fund portfolio to the several macroeconomic and microeconomic factors that govern the evolution of the underlying financial markets (and the surrounding social/economic environment) and, in advance, of the financial variables that compose the fund portfolio. In this direction, we focused on the design of a pension fund scheme, from a risk management point of view, presenting a detailed study concerning the case of Poland. On the other hand, uncertainty (in the Knightian sense) arises when the fund manager distrusts the model according to which he/she makes the investment decisions. Model uncertainty is a very important concept, as it places the first stones towards realistic modeling and risk management (see, e.g., Hansen and Sargent [38]). In fact, there exists a limited amount of literature that considers defined contribution pension fund schemes within a model uncertainty framework (especially in discrete time), something that highlights the importance of our research.

From a mathematical point of view, in order to effectively study the problem of DC pension fund management under risk and uncertainty, we will resort to Stochastic Analysis and, in particular, to stochastic (and robust) optimal control theory. The first step of our research is to construct a general, robust investment scheme for the optimal management of DC pension funds. This represents a huge amount of work that lies in the interplay between Mathematics, Finance and risk management. In a second step, we will embaptize the aforementioned derived robust scheme to the new Polish pension system. This will be carried out in two major ways: (a) by taking into account (and modeling) all the legal restrictions entitled by the new Polish pension system, and (b) by calibrating the results obtained in the first step to real market data, with focus on the Polish economy (e.g., interest rates, inflation, salaries, etc.). It is our strong belief that the results will act as a very useful benchmark to pension fund managers when trying to decide the appropriate investment (or hedging) strategy.

References

1. Akume, D., Weber, G.-W.: Risk-constrained dynamic portfolio management. Dyn. Contin. Discrete Impulsive Syst. (Series B) **17**, 113–129 (2010)
2. Anderson, E., Ghysels, E., Juergens, J.: The impact of risk and uncertainty on expected returns. J. Finan. Econ. **94**, 233–263 (2009)
3. Anderson, E., Hansen, L., Sargent, T.: A quartet of semigroups for model specification, robustness, prices of risk, and model detection. J. Europ. Econ. Assoc. **1**, 68–123 (2003)
4. Authority, P.F.S.: Sektor funduszy emerytalnych w Polsce-ewolucja, kształt, perspektywy. Warszawa (2016)
5. Baltas, I., Xepapadeas, A., Yannacopoulos, A.: Robust portfolio decisions for financial institutions. J. Dyn. Games **5**, 61–94 (2018)
6. Baltas, I., Yannacopoulos, A.: Optimal investment and reinsurance policies in insurance markets under the effect of inside information. Appl. Stochastic Models Bus. Industry **28**, 506–528 (2012)
7. Baltas, I., Yannacopoulos, A.: Uncertainty and inside information. J. Dyn. Games **3**, 1–24 (2016)
8. Bates, D.: Post-87 crash fears in the s&p 500 futures option market. J. Econ. **94**, 181–238 (2000)
9. Battocchio, P., Menoncin, F.: Optimal portfolio strategies with stochastic wage income and inflation. The case of a defined contribution pension fund. CeRP Working Papers 19, 1–24 (2002)
10. Bielawska, K., Chlon-Dominczak, A., Stanko, D. Retreat from mandatory pension funds in countries of the eastern and central europe in result of financial and fiscal crisis: Causes, effects and recommendations for fiscal rules. Research financed from research grant number UMO-2012/05/B/HS4/04206 from National Science Centre in Poland, Warsaw (2015)
11. Bodie, Z., Detemple, J., Otruba, S., Walter, S.: Optimal consumption-portfolio choices and retirement planning. J. Econ. Dyn. Control **28**, 1115–1148 (2004)
12. Boulier, J., Huang, S., Taillard, G.: Optimal management under stochastic interest rates: the case of a protected defined contribution pension fund. Insur. Math. Econ. **28**, 173–189 (2001)
13. Branger, N., Larsen, L., Munk, C.: Robust portfolio choice with ambiguity and learning predictability. J. Banking Finan. **37**, 1397–1411 (2013)
14. Breen, R., van der Werfhorst, H., Jaeger, M.: Deciding under doubt: a theory of risk aversion, time discounting preferences, and educational decision-making. Europ. Sociol. Rev. **30**, 258–270 (2014)
15. Brennan, M.J., Schwartz, E.S.: Regulation and corporate investment policy. J. Finan. **37**, 289–300 (1982)
16. Brock, W., Xepapadeas, A., Yannacopoulos, A.: Robust control and hot spots in spatiotemporal economic systems. Dyn. Games Appl. **4**, 257–289 (2014)
17. Brown, R.H., Schaefer, S.M.: The term structure of real interest rates and the cox, ingersoll, and ross model. J. Finan. Econ. **35**, 3–42 (1994)
18. Browne, S.: Optimal investment policies for a firm with a random risk process: Exponential utility and minimizing the probability of ruin. Mathem. Oper. Res. **20**, 937–958 (1995)
19. Buchowiec, M.: Ustawowe mierniki efektywnosci inwestycyjnej otwartych funduszy emerytalnych w polsce w latach 1999–2012—metodyka obliczania, ocena oraz postulowane kierunki zmian. zeszyty naukowe uniwersytetu szczecinskiego. Finanse, Rynki Finansowe 59, pp. 409–422 (2013)
20. Chen, J., Xiong, X., Zhu, J., Zhu, X.: Asset prices and economic fluctuations: the implications of stochastic volatility. Econ. Model. **64**, 128–140 (2017)
21. Chybalski, F.: Miary oceny efektow dzialalnosci inwestycyjnej ofe. Wiadomosci Statystyczne **10**, 22–35 (2006)
22. Crandall, M., Ishii, H., Lions, P.: User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. **27**, 1–67 (1992)

23. Czerwińska, T.: Aktywność inwestycyjna otwartych funduszy emerytalnych na gieldowym rynku akcji. projekcja rozwoju. zeszyty naukowe uniwersytetu szczecińskiego. Finanse, Rynki Finansowe **33**, 66–67 (2010)
24. Daykin, C.: Risk management and regulation of defined contribution schemes. Int. Soc. Secur. Assoc. ISSA/ACT/SEM/02/IV(a), 1–22 (2002)
25. Deelstra, G.: Long-term returns in stochastic interest rate models: applications. Astin Bull. **30**, 123–140 (2000)
26. Deelstra, G.: Optimal design of the guarantee for defined contribution funds. J. Econ. Dyn. Control **28**, 2239–2260 (2004)
27. Dopierala, L.: Polityka inwestycyjna zakladów ubezpieczeń w ramach indywidualnych kont emerytalnych. Ubezpieczenia Spoleczne. Teoria i praktyka **2**, 127–147 (2016)
28. Drazenovic, B., Buterin, V., Nikolaj, S.: Institutional challenges for mandatory pension funds in Central and Eastern Europe. 45th International Scientific Conference on Economic and Social Development (2019)
29. Duarte, I., Pinheiro, D., Pinto, A., Pliska, S.: Optimal life insurance purchase, consumption and investment on a financial market with multi-dimensional diffusive terms. Optimization **63**, 1737–1760 (2014)
30. Engle, R.: Auroregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. Econometrica **50**, 987–1007 (1982)
31. Fama, E., French, K.: The equity premium. J. Finan. **57**, 987–1007 (2002)
32. Fleming, W., Souganidis, P.: On the existence of value functions of two players zero sum stochastic differential games. Indiana Univ. Math. J. **38**, 293–314 (1989)
33. Flor, C., Larsen, L.: Robust portfolio choice with stochastic interest rates. Ann. Finan. **10**, 243–265 (2014)
34. Giancinto, M.D., Federico, S., Gozzi, F.: Pension funds with a minimum guarantee: a stochastic control approach. Finan. Stochastics **15**, 297–342 (2011)
35. Góra, M.: System Emerytalny. Warszawa, PWE (2003)
36. Guan, G., Liang, Z.: Optimal management of dc pension plan in a stochastic interest rate and stochastic volatility framework. Insurance: Math. Econ. **57**, 58–66 (2004)
37. Guan, G., Liang, Z.: Optimal management of dc pension plan under loss aversion and value-at-risk constraints. Insur.: Math. Econ. **69**, 224–237 (2016)
38. Hansen, L., Sargent, T.: Robust control and model uncertainty. Amer. Econ. Assoc. **91**, 60–66 (2001)
39. Jakubowski, S.: New legal standards for investment policy of open pension funds. Econ. Environ. Studies **33**, 77–94 (2015)
40. Jakubowski, S.: Reversal of the pension reform in poland. Social Res. **39**, 40–46 (2016)
41. Jones, C.S.: The dynamics of stochastic volatility: evidence from underlying and option markets. J. Econ. **116**, 181–224 (2003)
42. Kara, G., Ozmen, A., Weber, G.-W.: Stability advances in robust portfolio optimization under parallelepiped uncertainty. CEJOR **27**, 241–261 (2019)
43. Korn, R. Worst case scenario investment for insurers. Insur.: Math. Econ. **36**, 1–11 (2005)
44. Lin, C., Zeng, L., Wu, H.: Multi-period portfolio optimization in a defined contribution pension plan during the decumulation phase. J. Industr. Manag. Optim. **15**, 401–427 (2019)
45. Maringer, D.: Risk preferences and loss aversion in portfolio optimization. In: Kontoghiorghes, E.J., Rustem, B., Winker, P. (Eds.) Computational Methods in Financial Engineering. Springer (2008)
46. Markellos, R., Psychoyios, D.: Interest rate volatility and risk management: evidence from cboe treasury options. Quart. Rev. Econ. Finan. **68**, 190–202 (2018)
47. Mataramvura, S., Øksendal, B.: Risk minimizing portfolios and hjbi equations for stochastic differential games. Stochastics: Int. J. Probab. Stochastic Process. **80**, 317–337 (2008)
48. Mazurek-Krasodomska, E.: Ryzyko otwartych funduszy emerytalnych. Polityka spoleczna **24**, 24–27 (2011)
49. Merton, R.: Lifetime portoflio selection under uncertainty: the continuous case. Rev. Econ. Stat. **51**, 247–257 (1969)

50. Ortiz, I., Durán-Valverde, F., Urban, S., Wodsak, V. (Eds).: Reversing Pension Privatisation. Rebuilding public pension systems in Eastern Europe and Africa. International Labour Organization (2018)
51. Oxera. Study on the position of savers in private pension products. Avalible on line: https://www.oxera.com/publications/study-on-the-position-of-savers-in-private-pension-products
52. Pinar, M.: On robust mean-variance portfolios. Optimization **65**, 1039–1048 (2016)
53. Rieder, U., Wopperer, C.: Robust consumption-investment problems with random market coefficients. Math. Finan. Econ. **6**, 295–311 (2012)
54. Rutecka, J.: Country case: Poland. In: Pension Savings: the Real Return. A Research Report by Better Finance for All (2014)
55. Savku, E., Azevedo, N., and Weber, G.-W.: Optimal control of stochastic hybrid models in the framework of regime switches. In: Pinto, A., Zilberman, D. (Eds.), Modeling, Dynamics, Optimization and Bioeconomics II. Springer, Proceedings in Mathematics & Statistics 195 (2014)
56. Sun, J., Li, Y., Zhang, L.: Robust portfolio choice for a defined contribution pension plan with stochastic income and interest rates. Commun. Stat. Theory Methods, pp. 4106–4130 (2018)
57. Szczepański, M.: Stymulatory i bariery rozwoju zakladowych systemów emerytalnych na przykladzie Polski. Wydawnictwo Politechniki Poznańskiej, Poznań (2010)
58. Zawisza, D.: Robust consumption-investment problem on infinite horizon. Appl. Math. Optim. **72**, 469–491 (2015)
59. Zhang, C., Rong, X.: Optimal investment strategies for dc pension with stochastic salary under the affine interest rate model. Discrete Dyn. Nat. Soc. Article ID **297875**, 1–11 (2013)
60. Zhang, C., Rong, X., Zhao, H., Hou, R.: Optimal investment for the defined-contribution pension with stochastic salary under a cev model. Appl. Math. J. Chinese Univ. **28**, 187–203 (2013)

Collaborative Innovation of Spanish SMEs in the European Context: A Compared Study

María Bujidos-Casado, Julio Navío-Marco, and Beatriz Rodrigo-Moya

Abstract This chapter aims to go in depth into the relationship between the SME and innovation, especially in collaboration with other organizations. After going through the principal findings in literature on the subject, the chapter compares data of the Spanish situation and its evolution with results of the European environment and obtains conclusions that allow us to make a solid diagnosis of the innovative and collaborative SME's situation in our country in the European context. For this case, we use data from the Community Innovation Survey (CIS) by Eurostat, for 2004 and 2012. In this sense it is noted that there has been a deterioration of innovation of the SMEs in Spain, when at the European level, companies have augmented their innovative activity, including the smallest ones. On the contrary, a general improvement in cooperation for innovation has been detected from 2004 to 2012 on the European level as well as the Spanish level. There are also interesting phenomena noted like the increasing collaboration with competitors with a decrease in perceived value of collaboration with providers.

Keywords Innovation · Collaboration · Small · Medium-sized enterprises (SME) · Co-creation · Innovation networks · European union (EU)

M. Bujidos-Casado (✉) · J. Navío-Marco · B. Rodrigo-Moya
Universidad Nacional de Educación a Distancia, Madrid, Spain
e-mail: mbujidos1@alumno.uned.es

J. Navío-Marco
e-mail: jnavio@cee.uned.es

B. Rodrigo-Moya
e-mail: brodrigo@cee.uned.es

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_4

1 Introduction

A long tradition of academic research [19] attests that innovation is clearly considered as one of the fundamental strategic elements to improve business competitiveness of the SME. Business performance has been related to innovation and innovative companies can reach up to double the profitability [21].

Reference [4] show that innovation not only drives growth but also improves a wide variety of company capabilities that allow it to improve the ability to enter new markets and attract customers.

The strategy used by the company to position itself in the market is a factor that plays a key role, not only in the performance but also in the innovative attitude that it adopts [19]. For this reason, analysing the relationship between collaboration strategy and innovative activity is a central theme of business management, being especially important in the case of the SME, because of their weight in our economy and employment.

For small and medium companies, where internal innovation capacity is limited, largely dedicating its efforts and resources to the product and its marketing, co-creation and collaboration through partners or innovation collaboration networks, are particularly relevant, therefore it is interesting to analyze who develops and what results are obtained.

In Spain, due to the preponderance of SMEs (especially micro-businesses with 1 to 9 employees), the number of companies is one of the highest in the EU. According to the Central Companies Directory (DIRCE), produced annually by the Spanish National Statistical Institute (INE), the total number of companies in Spain stood at 3,236,582 in January of 2016. This is 1.6% more than the previous year. Reference [9] points out that the SMEs provide 66% of gross added value and 75% of jobs in Spain.

In the rest of the chapter we will delve deeper into the starting situation of Spanish SMEs as they address this process of collaborative innovation and, following a review of the literature on the subject matter, we will compare the situation in the European context, using as a baseline the data from the Community Innovation Survey (CIS) by Eurostat for 2004 and 2012, to finally draw the conclusions that emerge from said analysis.

2 Innovation in the SMEs: Brief Literature Review

First, we present the results from academic studies that have addressed innovation in SMEs in regard to how they approach it, their comparison with large companies, their motivations and types of partners. This is a particularly important aspect when discussing collaborative innovation of these kinds of businesses.

There is abundant literature about the relation between the SMEs and innovation. We find systematic reviews that address the management of innovation in this

kind of company, especially focused on the SME in the classical sense [11] and also abundant literature related to the innovative entrepreneurial SME, to which is attributed advantages when innovating in relation to large businesses [15] in that it is often characterized by organizational lightweight and agile structures and which may be in the adequate position to radically innovate and compete successfully in niche markets [24].

The comparison with innovation in large businesses is a recurring theme in literature and, except for the aforementioned possible advantage of the entrepreneurial SME, there is consensus affirming the superiority of the large firm in the field of innovation. This is mainly because the SME's capacity of innovation is conditioned by the limited amount of resources [22], forcing them to focus on innovation initiatives on a small scale linked to specific products or services, instead of substantial strategic innovation. We find a lack of human and financial resources and even fewer resources for the acquisition of external technologies, while large businesses can rely on their formal processes, and internal capabilities and skills to develop innovation; capabilities that have been considered strategic assets closely controlled by the company [13].

To fill these gaps, SMEs often open up to cooperation with networks and external firms [14]. The companied form alliances for two reasons; first, the investigation on the basis of economics and strategic management supposes that the complementarity of resources and the potential synergistic value creation can lead the businesses to form alliances [25]. Second, the sociologic perspective argues that social structures play an important role in the formation of alliances and the experiences of direct and indirect relations of businesses help the formation of future bonds [1]. In either case, the alliances are engines of value creation. On the negative side, when participating in an alliance, the company can suffer opportunistic behaviours of the partners [8]. This concern is especially relevant for SMEs in their knowledge and technology-based products, as they have relatively less bargaining power in relation to large companies [16]. In addition to the specific factors of the company, a high level of environmental uncertainty may discourage SMEs from engaging in partnerships. To alleviate these concerns, trust plays a central role in mitigating the fears of possible opportunistic behaviours [1].

Seen the reasons for cooperation, it is pertinent to ask who this cooperation is established with.

The literature on innovation indicates that there has been a systematic and fundamental change in the way companies have taken on innovative activities in recent years. In particular, there has been enormous growth in the use of networks with external companies of all sizes [26]. On the other hand, in the era of open innovation, according to [5], companies increase external sources of innovation and the use of a wider range of externalities, knowledge networks and resources that are indispensable for the creation of successful innovations for SMEs.

Internal relations of the network can be vertical, horizontal or lateral, including the networks of relationships between customers, suppliers and communities. Consequently, cooperation networks based on innovation include a heterogeneous group of different people and organizations, including representatives of companies, univer-

sities, organizations, centers of technology and development [14]. Reference [7] has postulated a relatively superior performance of SMEs in this context, which could reflect the greater capacity of these to exploit their network relationships through exchanging information and sharing of resources.

According to [10], the lack of external partners is an important barrier to the realization of product innovation for small businesses and he considers it to be an important difference between small and large companies. It is necessary to analyze the specific role that innovation networks play as a possible determining factor in the development of the innovation capacity of SMEs [22], which do not have the benefits of scale and scope provided by the size of large companies. Reference [23] notes that SMEs rely on external knowledge networks as an input of innovation more than large companies. Since small business seem to have potentially more to gain from innovative partnerships than large firms, the success of SMEs in relation to larger competitors may be due to their ability to use external networks more efficiently [22]. Thus, we see that networks and alliances are a double-edged sword for SMEs [8]: in order to access resources, the alliance can support innovation but can also expose the company to opportunistic behaviours of other members of the alliance [8].

Let us analyze the studies on innovation partners (types) with SMEs:

With regard to relationships in the value chain or vertical relationships (suppliers and customers), according to [13], bilateral cooperation with larger companies or strategic alliances with other SMEs in supplier-customer relationships has generated substantial literature. Since SMEs are generally more specialized, their participation in networks can effectively enable them to enter the broader markets and acquire additional resources to improve their chances with larger competitors [17]. Therefore, the network, seen as a specific type of relationship that binds a group of people, objects or events, is a very suitable model for SMEs [17]. Well managed innovation networks can therefore offer clear benefits for SMEs [13].

Many researchers have focused their analysis on networks of vertical collaboration, these being the most intuitive and close way to establish a cooperation network [12]. The numerous empirical studies about the theme are noteworthy [22], showing positive results in this vertical cooperation. In this regard, it is worth mentioning the study of [7] of 435 SMEs from West Midlands, UK that found a positive relationship between the companies cooperating with customers and suppliers with product innovation, while process innovation was only significant with cooperation of suppliers.

With regard to horizontal relationships, the seemingly paradoxical collaboration with competitors or coopetition, promoting collaboration in some stages of the product's life-cycle or in certain technical or production areas, has become a strategic imperative for companies in the networked business world. This phenomenon also occurs in cooperation with SMEs [13] and started to become popular. More generally, it has been accepted that the horizontal cooperation among SMEs can accelerate product development, provide economies of scale and mitigate the risk associated with shortage of resources for R&D and technology, allowing them to compete with larger players [18]. However, competitive cooperation introduces a risk of leakage of technology to rivals and a loss of control over the innovation process.

Empirical evidence has been found of the positive correlation between coopetition and innovation in SMEs [20], but other authors have not found significant evidence [7].

Cooperation with universities and research centers provides a cheaper, less risky and faster access to knowledge as well as technical and infrastructure support and expertise for the development of innovation activities [2], to add to it, collaboration with investigation centers and universities can compensate for the lack of a well developed absorption capacity, which is an obstacle that can be found when participating in innovation networks ([3]). Therefore, this type of collaboration is particularly relevant to SMEs, since, as explained, they can meet possible shortages of resources needed to innovate. On the contrary, some authors find SMEs to be less proactive to interact with such partners [6] so that cooperation with them is less than would be expected.

Finally, in the field of Government and public institutions, innovation management with public administration can be a difficult and complex process [2], especially for SMEs. This difficulty to manage public bureaucracy (administrative procedures, competitive bidding, budget proposals and expenditure control) can discourage the small business, limited by resources and negatively affect the results of innovation and slow development.

3 Data, Methodology and Analysis

This chapter uses data from the Community Innovation Survey (CIS) by Eurostat, based on innovation statistics that are a part of the science and technology statistics data from the EU. The surveys are conducted biannually and for this chapter we have used comparable data for Spain and the EU from CIS 2012 and CIS 4, covering three year periods of 2010–2012 and 2002–2004 respectively. In the analysis we maintain the distribution by the different sizes of the companies that are collected in the survey, specifying the information about small (10–49 employees), medium (51–249 employees) and large businesses (250 or more). The poll does not collect, despite their interest, data for micro companies (less than 10 employees).

As shown in Table 1, we start off in a sample universe with a total of 71,801 companies in 2012 in Spain, (80,958 in 2004). From this table one can get the first results: a deterioration of innovative activity in small business is observed in Spain, the percentage of companies that innovate is reduced from 34.7% in 2004 to 33.6% in 2012. This reduction is particularly pressing in smaller companies, while in the larger ones innovative activity increases. On the other hand, the opposite effect is observed on the European level, companies have increased their innovative activity, including the smallest ones. Therefore we can see the effect is particularly serious in Spain because of the aforementioned deterioration in relation to the improvement in the European context.

In this line, [9] also notes a disturbing fact: Between 2008 and 2010, the number of SMEs that carried out innovative activities decreased by 33%, while the large

Table 1 Innovative firms by business size

Spain								
	2012				2004			
	10–49 employees	50–249 employees	250 or more employees	Total	10–49 employees	50–249 employees	250 or more employees	Total
Enterprises with innovative activities, number	17.650	5.164	1.345	24.159	21.893	4.996	1.228	28.117
Enterprises with innovative activities, % of total	29.0%	55.7%	78.2%	33.6%	32.3%	43.8%	66.0%	34.7%
Total firms	60.817	9.264	1.720	71.801	67.695	11.403	1.860	80.958
Size of firm percentage of total firms	84.7%	12.9%	2.4%	100.0%	83.6%	14.1%	2.3%	100.0%
EU 28								
	2012				2004			
	10–49 employees	50–249 employees	250 or more employees	Total	10–49 employees	50–249 employees	250 or more employees	Total
Enterprises with innovative activities, number	282.189	80.178	21.666	384.033	207.429	70.503	21.102	299.034
Enterprises with innovative activities, % of total	42.5%	60.5%	76.4%	48.9%	34.9%	52.8%	70,8%	39.5%
Total firms	624.377	132.510	28.357	785.243	593.722	113.454	29.809	756.985
Size of firm percentage of total firms	79.5%	16.9%	3.6%	100.0%	78.4%	17.6%	3.9%	100.0%

Source Compilation based on Community Innovation Survey (2004, 2012). Eurostat

ones fell by 7%. The drop in the number of the ones conducting internal R&D was of 34% and 12% respectively. The main difficulty alleged by the Spanish SMEs to address innovative activities is cost, followed by the perception that it is not necessary to innovate, the difficulties to access the market and lastly the lack of adequate knowledge. The percentage of SMEs that considered a major cost difficulty was 34% in 2007 which rose to 45% in 2010.

Table 2 shows data on cooperative innovation collected by type of partner. In total numbers, when analyzing companies engaged in collaborative innovation and comparing periods 2004, 2012, Spanish companies reduce the distance with European companies; from a distance of 7.3% points, the gap is reduced to 1.9% points highlighting the increased collaborative innovation in companies with 10 to 49 employees (form 14.5% to 23.4%), higher increase than any other kind of company, but still far from the cooperation of the large Spanish company (54%).

The behaviour improvement of Spanish businesses is mainly focused on the national level, especially in the SMEs, where cooperative innovation with national partners increases prominently (with percentage increases of around 10 points).

Next we will analyze in detailed fashion the evolution of innovation by the type of partner who they cooperate with:

Cooperation with other companies in the group, on the European level an increase of 3% points is perceived, from 9.5% in 2004 to 12.5% in 2012, in Spain there is also an even greater increase of 4.7% points (from 3.8% in 2004 to 8.5% in 2012); despite this positive increase higher than the European average it is noted that the situation in Spain remains worse than the European, 4% points lower.

Regarding the cooperation with competitors, in the evolution from 2004 to 2012, at a European level, we can see a small noticeable improvement from 8.3% to 8.7% compared to a marked improvement at a Spanish level from 3% to 6.7%. If we analyze evolution according to the size of companies, Spain is at the European level, with a very prominent evolution, except for SMEs that although improving (from 2.1% to 4.9%) remain far from reaching the European average of 7.8%.

With respect to the cooperation with public sector clients, despite the lack of data available for Europe and for the period of 2004, the existing information for 2012 shows a more significant collaboration of large companies with the public sector in Spain. In addition, absolute levels are still very low, compared to the collaboration with companies in the private sector.

Furthermore, in the cooperation with suppliers, small improvements can be observed in all types of businesses both at a European and Spanish level, although the Spanish results are below the European ones. At the Spanish level, it is worth noting the decrease in value given to this type of collaboration, especially in the case of the large companies.

In the cooperation with universities and higher education institutions is observed a higher growth in Spain, from 4.7% to 10.3% compared to the European level from 8.8% to 13%. By type of business, we can see relevant growth of around 5% points. The case of small businesses is particularly striking for their poor starting point, 2.8% in 2004, reaching 7.2% in 2012. However, the collaboration of small businesses with universities and institutions is far from reaching the levels of large businesses.

Table 2 Collaborative innovation by type of partner

	2012								2004							
	Total		10–49 employees		50–249 employees		250 or more employees		Total		10–49 employees		50–249 employees		250 or more employees	
Cooperation	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %
Enterprises co-operating with other enterprises within the enterprise group	12.5	8.5	8.7	3.8	17.2	14.2	37.3	33.0	9.5	3.8	6.2	1.7	12.8	8.5	30.4	23.5
Enterprises for which cooperation with other enterprises within the enterprise group is the most valuable method		5.4		2.4		9.5		18.9		2.6		1.4		5.1		12.5
Enterprises co-operating with competitors or other enterprises of the same sector	8.7	6.7	7.8	4.9	9.0	8.4	17.4	17.2	8.3	3.0	7.1	2.1		4.8		12.5
Enterprises for which cooperation with competitors or other enterprises of the same sector is the most valuable method		2.0		1.8		2.0		3.0		1.4		1.3		1.8		3.2
Enterprises co-operating with clients of costumers from the private sector		9.2		7.3		11.9		18.6								
Enterprises for which cooperation with clients or customers from the private sector is the most valuable method		3.3		2.9		4.0		4.0								
Enterprises co-operating with clients or customers from the public sector		3.0		2.0		4.0		8.5								
Enterprises for which cooperation with clients or customers from the public sector is the most valuable method		0.4		0.4		0.4		0.7								
Enterprises co-operating with suppliers of equipment, materials, components or software	18.3	13.2	15.2	10.0	22.2	16.4	38.5	31.7	16.5	9.5	13.8	7.5	19.3	14.3	34.1	26.4

(continued)

Table 2 (continued)

	2012								2004							
	Total		10–49 employees		50–249 employees		250 or more employees		Total		10–49 employees		50–249 employees		250 or more employees	
Cooperation	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %
Enterprises for which cooperation with suppliers of equipment, materials, components or software is the most valuable method		6.3		5.9		6.7		8.9		6.7		6.0		8.9		10.3
Enterprises co-operating with universities or other higher institutions	13.0	10.3	10.0	7.2	16.4	13.3	33.9	28.5	8.8	4.7	6.3	2.8	11.2	8.3		22.5
Enterprises for which cooperation with universities or other higher education institutions is the most valuable method		4.2		3.5		5.2		7.3		2.0		1.4		3.5		5.5
Enterprises co-operating with Government, public or private research institutes	8.9	11.5		8.2	11.2	15.6	23.7	28.0								
Enterprises for which cooperation with Government, public or private research institutes is the most valuable method		5.5		4.4		7.6		8.4								
Enterprises co-operating with consultants or commercial labs	11.0	7.9	8.9	5.6	13.1	10.1	26.9	22.7								
Enterprises for which cooperation with consultants or commercial labs is the most valuable method		2.3		2.1		2.7		3.5								
Enterprises engaged in any type of co-operation	31.2	29.3	26.8	23.4	37.9	38.2	56.9	54.5	25.5	18.2	21.5	14.5	30.0	27.0	50.0	49.8

(continued)

Table 2 (continued)

	2012								2004							
	Total		10–49 employees		50–249 employees		250 or more employees		Total		10–49 employees		50–249 employees		250 or more employees	
Cooperation	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 28 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %	EU 27 %	Spain %
Enterprises engaged in any type of innovation co-operation with a partner in China or India		1.1		0.4		1.5		5.5								
Enterprises engaged in any type of innovation co-operation with a partner in EU countries, EFTA or EU candidates countries (except a national partner)	13.0	8.0	9.6	4.3	17.0	12.3	36.0	28.1		4.3		2.2		8.8		23.3
Enterprises engaged in any type of innovation co-operation with a national partner	27.1	27.8	22.8	22.5	33.4	35.7	51.9	50.0		17.2		13.9		24.5		46.3
Enterprises engaged in any type of innovation co-operation with a partner in all other countries except in EU countries, EFTA or EU candidates countries, United States, China or India		29.0		23.2		37.9		53.6								
Enterprises engaged in any type of innovation co-operation with a partner in EU countries, EFTA or EU candidates countries (incl. national parter)	4.6	1.9		1.1	5.4	2.5	11.8	6.8								
Enterprises engaged in any type of innovation co-operation with a partner in United States		2.0		0.9		2.7		9.1								

Source Compilation based on Community Innovation Survey (2004, 2012). Eurostat

The available data on collaboration with government and research institutes is scarce but allows to conclude that Spain is above the European levels of cooperation in this field, especially in large companies.

Finally, the collaboration of Spanish companies with EU and EFTA partners shows some improvement particularly concentrated in large companies. On the other hand, collaboration with partners outside this area (e.g. USA or China) is much more limited for 2012 and is focused on large companies.

When specifying on products and processes development with innovation, as made clear from the data collected in Table 3, it is observed that the development of innovative products within the company itself is comparable in Spain and Europe, and the Spanish SME behaves similarly to the European one. However, in the development of innovative products in cooperation with other companies or institutions, the Spanish situation is much worse than in Europe (16.7% vs. 29.3%) and in particular the situation of SMEs, (11.8% vs. 26%).

As far as innovation in processes, a different situation is observed, innovation in processes within the company itself is lower in Spain than in Europe, (56.4% vs. 62.5%) and the situation is even worse in innovation of processes in cooperation with other companies and institutions (16.4% vs. 37.2%). The situation is especially alarming in the case of SMEs (10.5% vs. 32.1%).

4 Conclusions

As was already indicated, a deterioration of innovative activity in small businesses has been found in Spain, whereas on the European level, companies have increased their innovative activity, including the smallest ones.

However, a general improvement is seen in cooperative innovation from 2004 to 2012 both at European and Spanish levels, with a remarkable improvement of SMEs. While the gap between small businesses and large companies is reduced, large companies are still doubling small ones in innovation. The small company works primarily with national partners, and European associates to a smaller extent. Collaborative innovation with partners outside this environment (e.g. United States or China) is practically irrelevant.

The good behaviour and evolution of collaboration with competitors in the Spanish case is noteworthy with better results than the European average, which remains virtually unchanged. This result is attributable especially to businesses of a large and medium-size; small businesses have not yet reached the European average. However, In Spain, it is remarkable that despite the increase in collaboration with competitors there is a decrease in the perceived value of collaborating with suppliers. This fact could denote a shift in the shapes of innovation based on coopetition and in any case, could denote sophistication in establishing collaboration relationships for innovation seeking a real value contribution (e.g. knowledge) despite the difficulties that may be involved. This result deserves further analysis in future work and research in this area.

Table 3 Collaborative innovation in products and processes

EU 28 2012	Total		10–49 employees		50–249 employees		2 or more employees	
	In product (%)	In process (%)	In product (%)	In process (%)	In product (%)	In process (%)	In product (%)	In process (%)
Enterprises that innovate by themselves	58.2	62.5	56.5	63.9	62.1	58.7	62.3	61.2
Enterprises that innovate in cooperation with other enterprises or institutions	29.3	37.2	26.0	32.1	34.5	45.3	43.6	57.0
SPAIN 2012	Total		10–49 employees		50–249 employees		2 or more employees	
	In product (%)	In process (%)	In product (%)	In process (%)	In product (%)	In process (%)	In product (%)	In process (%)
Enterprises that innovate by themselves	56.2	56.4	56.7	55.7	58.6	60.2	46.6	50.9
Enterprises that innovate in cooperation with other enterprises or institutions	16.7	16.4	11.8	10.5	20.7	23.6	34.5	40.4

Source Compilation based on Community Innovation Survey (2004, 2012). Eurostat

In the field of collaboration with public sector clients in Spain, there is a long way to go in promoting and achieving results in collaborative innovation with the public sector and especially when it comes to small businesses. In Spain, collaboration with government and research institutes is better than in Europe but the small business is still well below the large enterprise in this type of collaboration.

In Spain, the collaboration with universities and higher education institutions is growing faster than in Europe and is approaching the European level, but there is still a wide gap between small and large businesses in this area of collaboration. Reversing this situation and further research on the cooperation relationship between companies and universities is an important line of future work.

In general, Spanish businesses have a long way to go to reach European standards in the development of innovative processes. Furthermore, this gap also exists in the ability to innovate in cooperation with companies and institutions both in products and processes. The case of SMEs is particularly alarming.

In the Spanish case, the lack of importance accorded by Spanish firms to sources of information other than the companies themselves is also noticeable. Percentages above 40% of the companies surveyed do not use relevant sources of information from their own competitors, fairs, conferences or other companies in the sector. Establishing mechanisms to achieve an increased impact of these sources in the innovation of the companies and their positive valuation as such, is a challenge that should be addressed.

References

1. Adobor, H.: Trust as sense making: the micro dynamics of trust in inter firm alliances. J. Bus. Res. **58**(3), 330–337 (2005)
2. Antolin-Lopez, R., Martinez-del-Rio, J., Cespedes-Lorente, J.J., Perez-Valls, M.: The choice of suitable cooperation partners for product innovation: differences between new ventures and established companies. Europ. Manag. J. **33**(6), 472–484 (2015)
3. Bruton, G.D., Rubanik, Y.: Resources of the firm, Russian high-technology startups, and firm growth. J. Bus. Venturing **17**(6), 553–566 (2002)
4. Charan, R., Lafley, A.G.: Why innovation matters. Fast Company. 30 May. Retrieved from https://www.fastcompany.com/874798/why-innovation-matters (2008)
5. Chesbrough, H.: Open Innovation: The New Imperative for Creating and Profiting from Technology. Harvard Business School Press, Boston (2003)
6. Cooke, P., Boekholt, P., Todtling, F.: The Governance of Innovation in Europe. Pinter, London (2000)
7. De Propris, L.: Types of innovation and inter-firm co-operation. Entrepreneurship Regional Dev. **14**, 337–353 (2002)
8. Dickson, P.H., Weaver, K.M., Hoy, F.: Opportunism in the R&D alliances of SMES: the roles of the institutional environment and SME size. J. Bus. Venturing **21**(4), 487–513 (2006)
9. Fundación Cotec para la Innovación Tecnológica: La Innovación en las Pymes Españolas. Cotec, Madrid (2013)
10. Hewitt-Dundas, N.: Resource and capability constraints to innovation in small and large plants. Small Bus. Econ. **26**, 257–277 (2006)

11. Hörte, S., Barth, H., Chibba, A., Florén, H., Frishammer, J., Halila, F., Rundquist, J., Tell, J.: Product development in SMEs: a literature review. Int. J. Technol. Intell. Plann. **4**, 299–325 (2008)
12. Huizingh, E.K.R.E.: Open innovation: state of the art and future perspectives. Technovation **31**, 2–9 (2011)
13. Iturrioz, C., Aragón, C., Narvaiza, L.: How to foster shared innovation within SMEs' networks: social capital and the role of intermediaries. Europ. Manag. J. **33**(2), 104–115 (2015)
14. Kamalian, A.R., Rashki, M., Hemmat, Z., Jolfaie, S.A.: Cooperation networks and innovation performance of small and medium-sized enterprises (SMEs). Int. J. Manag. Account. Econ. **2**(3), 233–242 (2015)
15. Klewitz, J., Hansen, E.G.: Sustainability-oriented innovation of SMEs: a systematic review. J. Cleaner Product. **65**, 57–75 (2014)
16. Lavie, D.: Alliance portfolios and firm performance: a study of value creation and appropriation in the U.S. software industry. Strategy Manag. J. **28**, 1187–1212 (2007)
17. Lee, S., Park, G., Yoon, B., Park, J.: Open innovation in SMEs-an intermediated network model. Res. Policy **39**, 290–300 (2010)
18. Morris, M.H., Kocak, A., Ozer, A.: Co-opetition as a small business strategy: implications for performance. J. Small Bus. Strategy **18**(1), 35–55 (2007)
19. Moya, M.M., Alemán, J.L.M., de Lema, D.G.P.: La innovación en las pymes españolas: un estudio exploratorio. Información Comercial Española, ICE: Revista de economía **860**, 99–114 (2011)
20. Najib, M., Kiminami, A.: Innovation, co-operation and business performance: some evidence from Indonesian small food processing cluster. J. Agribusiness Dev. Emerg. Econ. **1**(1), 75–96 (2011)
21. Navío, J.: Las pymes y la nueva industrialización. BIT **199**, 37–40 (2015)
22. Nieto, M.J., Santamaria, L.: Technological collaboration: bridging the innovation gap between small and large firms. J. Small Bus. Manag. **48**(1), 44–69 (2010)
23. Rogers, M.: Networks, firm size and innovation. Small Bus. Econ. **22**, 141–153 (2004)
24. Schaltegger, S., Wagner, M.: Sustainable entrepreneurship and sustainability innovation: categories and interactions. Bus. Strat. Environ. **20**, 222–237 (2011)
25. Wassmer, U.: Alliance portfolio: a review and research agenda. J. Manag. **36**(1), 141–171 (2010)
26. Zeng, S.X., Xie, X.M., Tam, C.M.: Relationship between cooperation networks and innovation performance of SMEs. Technovation **30**(3), 181–194 (2010)

Haar Systems, KMS States on von Neumann Algebras and C^*-Algebras on Dynamically Defined Groupoids and Noncommutative Integration

G. G. de Castro, Artur O. Lopes, and G. Mantovani

Abstract We analyse Haar systems associated to groupoids obtained by certain equivalence relations of dynamical nature on sets like $\{1, 2, ..., d\}^{\mathbb{Z}}$, $\{1, 2, ..., d\}^{\mathbb{N}}$, $S^1 \times S^1$, or $(S^1)^{\mathbb{N}}$, where S^1 is the unitary circle. We also describe properties of transverse functions, quasi-invariant probabilities and KMS states for some examples of von Neumann algebras (and also C^*-Algebras) associated to these groupoids. We relate some of these KMS states with Gibbs states of Thermodynamic Formalism. We will show new results but we will also describe in detail several examples and basic results on the above topics. Some known results on non-commutative integration are presented, more precisely, the relation of transverse measures, cocycles and quasi-invariant probabilities. We describe the results in a language which is more familiar to the people in Dynamical Systems. Our intention is to study Haar systems, quasi-invariant probabilities and von Neumann algebras as a topic on measure theory (intersected with ergodic theory) avoiding questions of algebraic nature (which, of course, are also extremely important)

Keywords Haar systems · KMS states · von Neumann algebras · Groupoids · Quasi-invariant probabilities

1 The Groupoid Associated to a Partition

We will analyze properties of Haar systems, quasi-invariant probabilities, transverse measures, C^*-algebras and KMS states related to Thermodynamic Formalism and Gibbs states. We will consider a specific particular setting where the groupoid will be defined by some natural equivalence relations on the sets of the form $\{1, 2, ..., d\}^{\mathbb{N}}$ or $\{1, 2, ..., d\}^{\mathbb{Z}}$, $S^1 \times S^1$, or $(S^1)^{\mathbb{N}}$. These equivalence relations will be of dynamic origin.

We will consider only the so called subgroupoids of pair groupoids (see [68]).

We will denote by X any one of the above sets.

G. G. de Castro · A. O. Lopes (✉) · G. Mantovani
Instituto de Matemática e Estatística - UFRGS, Porto Alegre, Brazil

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_5

The main point here is that we will use a notation which is more close to the one used on Ergodic Theory and Thermodynamic Formalism.

On Sect. 2 we introduce the concept of transverse functions associated to groupoids and Haar systems.

On Sect. 3 we consider modular functions and quasi-invariant probabilities on groupoids. In the end of this section we present a new result concerning a (non-)relation of the quasi-invariant probability with the SBR probability of the generalized Baker map.

On Sect. 4 we consider a certain von Neumann algebra and the associated KMS states. On Proposition 1 we present a new result concerning the relation between probabilities satisfying the KMS property (quasi-invariant) and Gibbs (DLR) probabilities of Thermodynamic Formalism on the symbolic space $\{1, 2, ..., d\}^{\mathbb{N}}$ for a certain groupoid. Proposition 3 shows that the KMS probability is not unique on this case.

Reference [23–25] are the classical references on measured groupoids and von Neumann algebras. KMS states and C^*-algebras are described on [46, 52].

On Sect. 5 we present a natural expression—based on quasi-invariant probabilities—for the integration of a transverse function by a transverse measure. Some basic results on non-commutative integration (see [12] for a detailed description of the topic) are briefly described.

On Sect. 6 we present briefly the setting of C^*-algebras associated to groupoids on symbolic spaces. We present the well known and important concept of approximately proper equivalence relation and its relation with the direct inductive limit topology (see [17–19, 56]).

On Sect. 7 we present several examples of quasi-invariant probabilities for different kinds of groupoids and Haar systems.

Results on C^*-algebras and KMS states from the point of view of Thermodynamic Formalism are presented in [1, 21, 22, 29, 39, 50, 56, 61, 62].

The paper [7] considers equivalence relations and DLR probabilities for certain interactions on the symbolic space $\{1, 2, ...d\}^{\mathbb{Z}}$ (not in $\{1, 2, ...d\}^{\mathbb{N}}$ like here).

Theorem 6.2.18 in Vol II of [4, 9] describe the relation between KMS states and Gibbs probabilities for interactions on certain spin lattices (on the one-dimensional case corresponds to the space $\{1, 2, .., d\}^{\mathbb{Z}}$).

We point out that Lecture 9 in [15] presents a brief introduction to C^*-Algebras and non-commutative integration.

In [34, 35] are presented generalizations of some of the results (of Sects. 4 and 5) described here.

We are indebted to Ruy Exel for many helpful discussions and useful comments during the procedure of writing this paper. We also thanks Ali Tahzibi for some fruitful remarks.

We denote $\{1, 2, ..., d\}^{\mathbb{N}} = \Omega$ and consider the compact metric space with metric d where for $x = (x_0, x_1, x_2, ..) \in \Omega$ and $y = (y_0, y_1, y_2, ..) \in \Omega$

$$d(x, y) = 2^{-N},$$

where N is the smallest natural number $j \geq 0$, such that, $x_j \neq y_j$.

We also consider $\{1, 2, ..., d\}^{\mathbb{Z}} = \hat{\Omega}$ and elements in $\hat{\Omega}$ are denoted by $x = (..., x_{-n}, ..., x_{-1} \,|\, x_0, x_1, .., x_n, ..)$.

We will use the notation $\overleftarrow{\Omega} = \{1, 2, ..., d\}^{\mathbb{N}}$ and $\overrightarrow{\Omega} = \{1, 2, ..., d\}^{\mathbb{N}}$.

Given $x = (..., x_{-n}, ..., x_{-1} \,|\, x_0, x_1, .., x_n, ..) \in \hat{\Omega}$, we call $(..., x_{-n}, ..., x_{-1}) \in \overleftarrow{\Omega}$ the past of x and $(x_0, x_1, .., x_n, ..) \in \overrightarrow{\Omega}$ the future of x.

In this way we express $\hat{\Omega} = \overleftarrow{\Omega} \times | \overrightarrow{\Omega}$.

Sometimes we denote

$$(..., a_{-n}, ..., a_{-1} \,|\, b_0, b_1, .., b_n, ..) =< a \,|\, b >,$$

where $a = (..., a_{-n}, ..., a_{-1}) \in \overleftarrow{\Omega}$ and $b = (b_0, b_1, .., b_n, ..) \in \overrightarrow{\Omega}$.

On $\hat{\Omega}$ we consider the usual metric d, in such way that for $x, y \in \hat{\Omega}$ we set

$$d(x, y) = 2^{-N},$$

$N \geq 0$, where for

$$x = (..., x_{-n}, ..., x_{-1} \,|\, x_0, x_1, .., x_n, ..) \ , y = (..., y_{-n}, ..., y_{-1} \,|\, y_0, y_1, .., y_n, ..),$$

we have $x_j = y_j$, for all j, such that, $-N + 1 \leq j \leq N - 1$ and, moreover $x_N \neq y_N$, or $x_{-N} \neq y_{-N}$.

The shift $\hat{\sigma}$ on $\hat{\Omega} = \{1, 2, ..., d\}^{\mathbb{Z}}$ is such that

$$\hat{\sigma}(..., y_{-n}, ..., y_{-2}, y_{-1} \,|\, y_0, y_1, ..., y_n, ...) = (..., y_{-n}, ..., y_{-2}, y_{-1}, y_0 \,|\, y_1, ..., y_n, ...).$$

On the other hand the shift σ on $\Omega = \{1, 2, ..., d\}^{\mathbb{N}}$ is such that

$$\sigma(y_0, y_1, ..., y_n, ...) = (y_1, ..., y_n, ...).$$

A general equivalence relation R on a space X define classes and we will denote by $x \sim y$ when two elements x and y are on the same class. We denote by $[y]$ the class of $y \in X$.

Definition 1 Given an equivalence relation $\sim$ on X, where X is any of the sets Ω, $\hat{\Omega}$, $(S^1)^{\mathbb{N}}$, or $S^1 \times S^1$, we denote by G the subset of $X \times X$, containing all pairs (x, y), where $x \sim y$. We call G the groupoid associated to the equivalence relation $\sim$.

We also denote by G^0 the set $\{(x, x) \,|\, x \in X\} \sim X$, where X denote any of the sets Ω, $\hat{\Omega}$, $(S^1)^{\mathbb{N}}$, or $S^1 \times S^1$.

Remark There is a general definition of groupoid (see [12]) which assumes more structure but we will not need this here. For all results we will consider there is no need for an additional algebraic structure (on the class of each point). In this way we can consider a simplified definition of groupoid as it is above. Our intention is to

study C^*-algebras and Haar systems as a topic on measure theory (intersected with ergodic theory) avoiding questions of algebraic nature.

There is a future issue about the topology we will consider induced on G. One possibility is the product topology, which we call the standard structure, or, a more complex one which will be defined later on Sect. 6 (specially appropriate for some C^*-algebras).

We will present several examples of dynamically defined groupoids. The equivalence relation of most of our examples is proper (see Definition 21).

Example 1 For example consider on $\{1, 2, ..., d\}^{\mathbb{N}}$ the equivalence relation R such that $x \sim y$, if $x_j = y_j$, for all $j \geq 2$, when $x = (x_1, x_2, x_3, ...)$ and $y = (y_1, y_2, y_3, ...)$. This defines a groupoid G. In this case $G^0 = \Omega = \{1, 2, ..., d\}^{\mathbb{N}}$.

For a fixed $x = (x_1, x_2, x_3, ...)$ the equivalence class associated to x is the set $\{ (j, x_2, x_3, ...), j = 1, 2.., d \}$. We call this relation the **bigger than two relation**.

Example 2 Consider an equivalence relation R which defines a partition η_0 of $\{1, 2, ..., d\}^{\mathbb{Z}-\{0\}} = \hat{\Omega}$ such its elements are of the form

$$a \times | \overrightarrow{\Omega} = a \times \{1, 2, ..., d\}^{\mathbb{N}} = (..., a_{-n}, ..., a_{-2}, a_{-1}) \times | \{1, 2, ..., d\}^{\mathbb{N}},$$

where $a \in \{1, 2, ..., d\}^{\mathbb{N}} = \overleftarrow{\Omega}$. This defines an equivalence relation $\sim$.

In this way two elements x and y are related if they have the same past.

There exists a bijection of classes of η_0 and points in $\overleftarrow{\Omega}$.

Denote $\pi = \pi_2 : \hat{\Omega} \to \overleftarrow{\Omega}$ the transformation such that takes a point and gives as the result its class.

In this sense

$$\pi^{-1}(x) = \pi^{-1}((..., x_{-n}, ..., x_{-1} \,|\, x_1, .., x_n, ..)) =$$

$$(..., x_{-n}, ..., x_{-2}, x_{-1}) \times | \overrightarrow{\Omega} \cong \Omega.$$

The groupoid obtained by this equivalence relation can be expressed as $G = \{(x, y), \pi(x) = \pi(y)\}$. In this way $x \sim y$ if they have the same past.

In this case the number of elements in each fiber is not finite.

Using the notation of page 46 of [12] we have $Y \subset \hat{\Omega} \times \hat{\Omega}$ and $X = \overleftarrow{\Omega}$.

In this case each class is associated to certain $a = (a_{-1}, a_{-2}, ..) \in \overleftarrow{\Omega} = \{1, 2, ..., d\}^{\mathbb{N}}$.

We use the notation $(a \,|\, x)$ for points on a class of the form

$$(a \,|\, x) = (..., a_{-n}, ..., a_{-2}, a_{-1} \,|\, x_1, ..., x_n, ...) \ .$$

Example 3 A particulary important equivalence relation R on $\hat{\Omega} = \{1, 2, ..., d\}^{\mathbb{Z}}$ is the following: we say $x \sim y$ if

$$x = (..., x_{-n}, ..., x_{-2}, x_{-1} \,|\, x_0, x_1, ..., x_n, ...),$$

and,

$$y = (..., y_{-n}, ..., y_{-2}, y_{-1} \,|\, y_0, y_1, ..., y_n, ...)$$

are such that there exists $k \in \mathbb{Z}$, such that, $x_j = y_j$, for all $j \leq k$.

The groupoid G_u is defined by this relation $x \sim y$.

By definition the unstable set of the point $x \in \hat{\Omega}$ is the set

$$W^u(x) = \{y \in \hat{\Omega}\,, \text{ such that } \lim_{n\to\infty} d(\hat{\sigma}^{-n}(x), \hat{\sigma}^{-n}(y)) = 0\}.$$

One can show that the unstable manifold of $x \in \hat{\Omega}$ is the set

$$W^u(x) = \{y = (..., y_{-n}, ..., y_{-2}, y_{-1} \,|\, y_0, y_1, ..., y_n, ...) \,|\, \text{there exists}$$

$$k \in \mathbb{Z}, \text{ such that } x_j = y_j, \text{for all } j \leq k\}.$$

If we denote by G_u the groupoid defined by the above relation, then, $x \sim y$, if and only if, $y \in W^u(x)$.

An equivalence relation of this sort—for hyperbolic diffeomorphism—was considered on [38, 58].

Example 4 An equivalence relation on $\overrightarrow{\Omega} = \{1, 2, ..., d\}^{\mathbb{N}}$ similar to the previous one is the following: we say $x \sim y$ if

$$x = (x_0, x_1, ..., x_n, ...),$$

and,

$$y = (y_0, y_1, ..., y_n, ...)$$

are such that there exists $k \in \mathbb{N}$, such that, $x_j = y_j$, for all $j \geq k$.

Example 5 Another equivalence relation on $\overrightarrow{\Omega}$ is the following: fix $k \in \mathbb{N}$, and we say $x \sim_k y$, if when

$$x = (x_0, x_1, ..., x_n, ...),$$

and,

$$y = (y_0, y_1, ..., y_n, ...)$$

we have $x_j = y_j$, for all $j \geq k$.

In this case each class has d^k elements.

Example 6 Given $x, y \in \hat{\Omega} = \{1, 2, ..., d\}^{\mathbb{Z}}$, we say that $x \sim y$ if

$$\lim_{k \to +\infty} d(\hat{\sigma}^k x, \hat{\sigma}^k y) = 0$$

and

$$\lim_{k \to -\infty} d(\hat{\sigma}^k x, \hat{\sigma}^k y) = 0. \tag{1}$$

This means there exists an $M \geq 0$, such that, $x_j = y_j$ for $j > M$, and, $j < -M$. In other words, there are only a finite number of i's such that $x_i \neq y_i$. This is the same to say that x and y are homoclinic.

For example in $\hat{\Omega} = \{1, 2\}^{\mathbb{Z}}$ take

$$x = (..., x_{-n}, ..., x_{-7}, 1, 2, 2, 1, 2, 2 \,|\, 1, 2, 1, 2, 1, 1, x_7, ...x_n, ..)$$

and

$$y = (..., y_{-n}, ..., y_{-7}, 1, 2, 2, 1, 2, 2 \,|\, 1, 2, 1, 1, 1, 2, y_7, ...y_n, ..)$$

where $x_j = y_j$ for $|j| \geq 7$.

In this case $x \sim y$.

This relation is called the **homoclinic relation** on $\hat{\Omega}$. It was considered for instance by Ruelle and Haydn in [26, 57] for hyperbolic diffeomorphisms and also on more general contexts (see also [7, 33, 43] for the symbolic case).

Example 7 Consider an expanding transformation $T : S^1 \to S^1$, of degree two, such that $\log T'$ is Holder and $\log T'(a) > \log \lambda > 0$, $a \in S^1$, for some $\lambda > 1$.

Suppose $T(x_0) = 1$, where $0 < x_0 < 1$. We say that $(0, x_0)$ and $(x_0, 1)$ are the domains of injectivity of T.

Denote $\psi_1 : [0, 1) \to [0, x_0)$ the first inverse branch of T and $\psi_2 : [0, 1) \to [x_0, 1]$ the second inverse branch of T.

In this case for all y we have $T \circ \psi_1(y) = y$ and $T \circ \psi_2(y) = y$.

The **associated T-Baker map** is the transformation $F : S^1 \times S^1$ such that satisfies for all a, b the following rule:

(1) if $0 \leq b < x_0$

$$F(a, b) = (\psi_1(a), T(b)),$$

and

(2) if $x_0 \leq b < 1$

$$F(a, b) = (\psi_2(a), T(b)).$$

In this case we take as partition the one associated to (local) unstable manifolds for F, that is, sets of the form $W_a = \{ (a, b) \,|\, b \in S^1 \}$, where $a \in S^1$.

Given two points $z_1, z_2 \in S^1 \times S^1$ we say that they are related if the the first coordinate is equal.

On $S^1 \times S^1$ we use the distance d which is the product of the usual arc length distance on S^1.

The bijection F expands vertical lines and contract horizontal lines.

As an example one can take $T(a) = 2a \pmod 1$ and we get (the inverse of) the classical Baker map (see [59]).

One can say that the dynamics of such F in some sense looks like the one of an Anosov diffeomorphism.

2 Kernels and Transverse Functions

A general reference for the material of this section is [12] (see also [27]).

We consider over $G \subset X \times X$ the Borel sigma-algebra $\mathscr{B}$ (on G) induced by the natural product topology on $X \times X$ and the metric d on X ([23, 24] also consider this sigma algebra). This will be fine for the setting of von Neumann algebras. Later, another sigma-algebra will be considered for the setting C^*-algebras.

We point out that the only sets X which we are interested are of the form $\hat{\Omega}$, Ω, $(S^1)^{\mathbb{N}}$, or $S^1 \times S^1$.

We denote $\mathscr{F}^+(G)$ the space of Borel measurable functions $f : G \to [0, \infty)$ (a function of two variables (a, b)).

$\mathscr{F}(G)$ is the space of Borel measurable functions $f : G \to \mathbb{R}$. Note that $f(x, y)$ just make sense if $x \sim y$.

There is a natural involution on $\mathscr{F}(G)$ which is $f \to \tilde{f}$, where $\tilde{f}(x, y) = f(y, x)$.

We also denote $\mathscr{F}^+(G^0)$ the space of Borel measurable functions $f : G^0 \to [0, \infty)$ (a function of one variable a).

There is a natural identification of functions $f : G^0 \to \mathbb{R}$, of the form $f(x)$, with functions $g : G \to \mathbb{R}$ which depend only on the first coordinate, that is $g(x, y) = f(x)$. This will be used without mention, but if necessary we write $(f \circ P_1)(x, y) = f(x)$ and $(f \circ P_2)(x, y) = f(y)$.

Definition 2 A **measurable groupoid** G is a groupoid with the topology induced by the product topology over $X \times X$, such that, the following functions are measurable for the Borel sigma-algebra:

$P_1(x, y) = x$, $P_2(x, y) = y$, $h(x, y) = (y, x)$ and $Z((x, s), (s, y)) = (x, y)$, where $Z : \{ ((x, s), (r, y)) \mid r = s \} \subset G \times G \to G$.

Now, we will present the definition of kernel (see beginning of Sect. 2 in [12]).

Definition 3 A **G-kernel** ν on the measurable groupoid G is an application of G^0 in the space of measures over the sigma-algebra $\mathscr{B}$, such that,

(1) for any $y \in G^0$, we have that ν^y has support on $[y]$,
and
(2) for any $A \in \mathscr{B}$, the function $y \to \nu^y(A)$ is measurable.

The set of all G- kernels is denoted by $\mathscr{K}^+$.

Example 8 As an example consider for the case of the groupoid G associated to the bigger than two relation, the measure ν^y, for each $y = (y_1, y_2, y_3, ...)$, such that $\nu(j, y_2, y_3, ...) = 1$, $j = 1, 2, ..., d$. In other words we are using the counting measure on each class.We call this the standard G-kernel for the the bigger than two relation.

More precisely, the **counting measure** is such that $\nu^y(A) = \#(A \cap [y])$, for any $A \in \mathscr{B}$.

Example 9 Another possibility is to consider the G-kernel such that ν^y, for each $y = (y_1, y_2, y_3, ...)$, is such that $\nu(j, y_2, y_3, ...) = \frac{1}{d}$, We call this the normalized standard G-kernel for the the bigger than two relation.

Example 10 Given any groupoid G another example of kernel is the **delta kernel** ν which is the one such that for any $y \in G^0$ we have that $\nu^y(dx) = \delta_y(d\,x)$, where δ_y is the delta Dirac on y. We denote by $\mathfrak{d}$ such kernel.

We denote by $\mathscr{F}_\nu(G)$ the set of ν-integrable functions.

Definition 4 Given a G-kernel ν and an integrable function $f \in \mathscr{F}_\nu(G)$ we denote by $\nu(f)$ the function in $\mathscr{F}(G^0)$ defined by

$$\nu(f)\ (y) = \int f\ (s, y)\ \nu^y(ds), \quad y \in G^0.$$

A kernel ν is characterized by the law

$$f \in \mathscr{F}_\nu(G) \ \ \rightarrow \ \ \nu(f) \in \mathscr{F}(G^0).$$

In other words, for a kernel ν we get

$$\nu : \mathscr{F}_\nu(G) \rightarrow \mathscr{F}(G^0).$$

By notation given a kernel ν and a positive $f \in \mathscr{F}_\nu(G)$ then the kernel $f\,\nu$ is the one defined by $f(x, y)\nu^y(dx)$. In other words the action of the kernel $f\,\nu$ get rid of the first coordinate:

$$h(x, y) \ \rightarrow \ \int h(s, y) f\ (s, y)\ \nu^y(ds).$$

In this way if $f \in \mathscr{F}_\nu(G^0)$ we get $f(x)\nu^y(dx)$.
Note that $\nu(f)$ is a function and $f\,\nu$ is a kernel.

Definition 5 A **transverse function** is a G-kernel ν, such that, if $x \sim y$, then, the finite measures ν^y and ν^x are the same. The set of transverse functions for G is denoted by $\mathscr{E}^+$. We call probabilistic transverse function any one such that for each $y \in G^0$ we get that ν^y is a probability on the class of y.

The above means that

$$\int f(a)\nu^x(d\,a) = \int f(a)\nu^y(d\,a),$$

if x and y are related. In the above we have $x \sim y \sim a$.

Remark 1 The above equality implies that a transverse function is left (and right) invariant.

The concept of transverse function considered here is a particular version of the general definition presented in [12].

The standard G-kernel for the bigger than two relation (see Example 8) is a transverse function.

The normalized standard G-kernel for the bigger than two relation (see Example 9) is a probabilistic transverse function.

If we consider the equivalence relation such that each point is related just to itself, then the transverse functions can be identified with the positive functions defined on X.

The difference between a function and a transverse function is that the former takes values on the set of real numbers and the later on the set of measures.

If ν is transverse, then $\nu^x = \nu^y$ when $x \sim y$, and we have from Definition 18:

$$(\nu * f)(x, y) = \int f\,(x, s)\,\nu^x(ds) = \nu(\tilde{f})(x), \quad \forall y \sim x \tag{2}$$

and,

$$(f * \nu)(x, y) = \int f\,(s, y)\,\nu^y(ds) = \nu(f)(y), \quad \forall x \sim y. \tag{3}$$

Definition 6 The pair $(G,\ \nu)$, where $\nu \in \mathscr{E}^+$, is called the **measured groupoid for the transverse function** ν or a **Haar System**. We assume any ν we consider is such that ν^y is not the zero measure for any y.

In the case ν is such that, $\int \nu^y(ds) = 1$, for any $y \in G^0$, the Haar system will be called a probabilistic Haar system.

Note that the delta kernel $\mathfrak{d}$ is not a transverse function.

Given a measured groupoid (G, ν) and two measurable functions $f, g \in \mathscr{F}_\nu(G)$, we define $(f \underset{\nu}{*} g) = h$ in such way that for any $(x, y) \in G$

$$(f \underset{\nu}{*} g)(x, y) = \int g(x, s)\, f(s, y)\ \nu^y(ds) = h(x, y).$$

$(f \underset{\nu}{*} g)$ is called the **convolution** of the functions f, g for the measured groupoid (G, ν).

Example 11 Consider the groupoid G of Example 1 and the family ν^y, $y \in \{1, 2, ..., d\}^{\mathbb{N}}$, of measures (where each measure ν^y has support on the equivalence class of y), such that, ν^y is the counting measure. This defines a transverse function (Haar system) called the **standard Haar system**.

Example 12 Consider the groupoid G of Example 1 and the normalized standard family ν^y, $y \in \{1, 2, ..., d\}^{\mathbb{N}}$. This defines a transverse function called the **normalized standard Haar system**.

More precisely the family ν^y, $y \in \{1, 2, ..., d\}^{\mathbb{N}}$, $y = (y_1, y_2, y_3, ...)$, of probabilities on the set

$$\{ (a, y_2, y_3, ...), a \in \{1, 2, .., d\} \},$$

is such that, $\nu^y(\{(a, y_2, y_3, ...)\}) = \frac{1}{d}$, $a \in \{1, 2, .., d\}$

Example 13 In Example 5 in which k is fixed consider the transverse function ν such that for each $y \in G^0$, we get that ν^y is the counting measure on the set of points $x \sim_k y$.

Example 14 Suppose $J : \{1, 2, ..., d\}^{\mathbb{N}} \to \mathbb{R}$ is continuous positive function such that for any $x \in \Omega$ we have that $\sum_{a=1}^{d} J(ax) = 1$. For the groupoid G of Example 1, the family ν^y, $y \in \{1, 2, ..., d\}^{\mathbb{N}}$, of probabilities on $\{ (a, y_2, y_3, ...), a \in \{1, 2, .., d\} \}$, such that, $\nu^y(a, y_2, y_3, ...) = J(a, y_2, y_3, ...)$, $a \in \{1, 2, .., d\}$ defines a Haar system. We call it the **probability Haar system associated to** J.

Example 12 is a particular case of the present example.

Example 15 On the groupoid over $\{1, 2, ..., d\}^{\mathbb{Z}}$ described on Example 2, where we consider the notation: for each class specified by $a \in \overleftarrow{\Omega}$ the general element in the class is given by

$$(a \,|\, x) = (..., a_{-n}, ..., a_{-2}, a_{-1} \,|\, x_1, ..., x_n, ...),$$

where $x \in \overrightarrow{\Omega}$.

Consider a fixed probability μ on $\overrightarrow{\Omega}$. We define the transverse function $\nu^a(dx) = \mu(dx)$ independent of a.

Example 16 We will show that the above defined concept of convolution generalize (in some sense) the product of matrices. Consider over the set $G^0 = \{1, 2.., d\}$ the equivalence relation where all points are related. In this case $G = \{1, 2.., d\} \times \{1, 2.., d\}$. Take ν as the counting measure. A function $f : G \to \mathbb{R}$ is denoted by $f(i, j)$, where $i \in \{1, 2.., d\}$, $j \in \{1, 2.., d\}$.

In this case the set of functions $f : G \to \mathbb{R}$ can be identified with the set of d by d matrices with real entries. A matrix A can be identified with $A = (f_{i,j})_{i \in \{1,2..,d\}, j \in \{1,2..,d\}}$.

The convolution product is

$$(f \underset{\nu}{*} g)(i, j) = \sum_k g(i, k)\, f(k, j).$$

The convolution is just the product of matrices.

Example 17 The so called generalized XY model consider space $(S^1)^{\mathbb{N}}$, where S^1 is the unitary circle and the shift acting on it (see).

We can consider the equivalence relation R such that $x \sim y$, if $x_j = y_j$, for all $j \geq 2$, when $x = (x_1, x_2, x_3, ...)$ and $y = (y_1, y_2, y_3, ...)$. This defines a groupoid G. In this case $G^0 = (S^1)^{\mathbb{N}}$.

For a fixed $x = (x_1, x_2, x_3, ...)$ the equivalence class associated to x is the set $\{ (a, x_2, x_3, ...), a \in S^1 \}$. We call this relation the **bigger than two relation for the XY model** and G the **standard XY groupoid** over $(S^1)^{\mathbb{N}}$.

Given the class $\{(x, x_2, x_3, ..), x \in S^1\}$, where $(x_2, x_3, ..)$ is fixed, the transverse function could be dx for instance.

We refer the reader to [6, 36, 40, 44] for general results on the Thermodynamic Formalism for the XY model. We point out that as in this example the cardinality of each set is not countable (and the transverse is dx in each class) it will be natural to consider a Ruelle operator with an a priori probability equal to dx. The dynamics is given by the shift. These papers are useful for the understanding of Haar systems in such type of groupoids. This claim is related to the future Theorem 1 (see [35])

Example 18 In the Example 7 we consider the partition of $S^1 \times S^1$ given by the sets $W_a = \{ (a, b) \,|\, b \in S^1\}$, where $a \in S^1$. For each $a \in S^1$, consider a probability $\nu^a(d\, b)$ over $\{ (a, b) \,|\, b \in S^1\}$ such that for any Borel set $K \subset S^1 \times S^1$ we have that $a \to \nu^a(K)$ is measurable. This defines a probabilistic transverse function and a Haar system.

Consider a continuous function $A : S^1 \times S^1 \to \mathbb{R}$. For each a consider the kernel ν^a such that $\int f(b)\nu^a(db) = \int f(b)\, e^{A(a,b)} db$, where db is the Lebesgue measure. This defines a transverse function.

We call the **standard Haar system on $\mathbf{S^1} \times \mathbf{S^1}$** the case where for each a we consider as the probability $\nu^a(d\, b)$ over $\{ (a, b) \,|\, b \in S^1\}$ the Lebesgue probability on S^1.

We will present several properties of kernels and transverse functions on Sect. 5.

A question of notation: for a fixed groupoid G we will describe now for the reader the common terminology on the field (see [12, 27, 53, 56]). It is usual to denote a general pair $(x, y) \in G$ by γ (of related elements x, y). The γ is called the directed arrow from x to y. In this case we call $s(\gamma) = x$ and $r(\gamma) = y$ (see [45] for a more detailed description of the arrow's setting).

Here, for each pair of related elements (x, y) there exist an unique directed arrow γ satisfying $s(\gamma) = x$ and $r(\gamma) = y$. Note that, since se are dealing with equivalence relations, (y, x) denotes another arrow. In category language: there is a unique morphism γ that takes $\{x\}$ to $\{y\}$, whenever x and y are related, and this morphism is associated in a unique way to the pair (x, y).

In this notation $r^{-1}(y)$ is the set of all arrows that end in y. This is in a bijection with all elements on the same class of equivalence of y. We call $r^{-1}(y)$ the fiber over y. If $x \sim y$, then $r^{-1}(y) = r^{-1}(x)$.

We adapt the notation in [12, 27] to our notation. We use here the expression (s, y) instead of $\gamma\,\gamma'$. This makes sense considering that $\gamma = (x, y)$ and $\gamma' = (s, x)$. We use the expression (y, s) for $(\gamma')^{-1}\,\gamma$, where in this case, $\gamma = (x, y)$ and $\gamma' = (s, y)$, and, finally, $\nu^y(\gamma')$ means $\nu^y(ds)$ for $\gamma' = (s, y)$.

In the case of the groupoid G associated to the bigger than two relation we have for each $x = (x_1, x_2, x_3, ...)$ the property $r^{-1}(x) = \{\,(j, x_2, x_3, ...),\ j = 1, 2.., d\,\}$.

The terminology of arrows will not be essentially used here. It was introduced just for the reader to make a parallel (a dictionary) with the one commonly used on papers on the topic.

Using the terminology of arrows Definition 5 is equivalent to say that: if, $\gamma = (x, y) = (s(\gamma), r(\gamma))$, then,

$$\nu^y = \gamma\,\nu^x.$$

Related results on Haar systems and transverse functions appear in [34, 35, 37].

3 Quasi-invariant Probabilities

Definition 7 A function $\delta : G \to \mathbb{R}$ such that

$$\delta(x, z) = \delta(x, y)\,\delta(y, z),$$

for any $(x, y), (y, z) \in G$ is called a **modular function** (also called a multiplicative **cocycle**).

In the arrow notation this is equivalent to say that

$$\delta(\gamma_1\gamma_2) = \delta(\gamma_1)\delta(\gamma_2).$$

Note that $\delta(x, y)\,\delta(y, y) = \delta(x, y)$ and it follows that for any y we have $\delta(y, y) = 1$. Moreover, $\delta(x, y)\,\delta(y, x) = \delta(x, x) = 1$ is true. Therefore, we get $\tilde{\delta} = \delta^{-1}$.

Example 19 Given $W : G^0 \to \mathbb{R}$, $W(x) > 0, \forall x$, a natural way to get a modular function is to consider $\delta(x, y) = \frac{W(x)}{W(y)}$. In this case we say that the modular function is derived from W.

Example 20 In the case of Example 7 the equivalence relation is: given two points $z_1, z_2 \in S^1 \times S^1$ they are related if the first coordinate is equal.

Consider a expanding transformation T and the associated Baker map F. Note que $F^n(a, b) = (*, T^n(b))$ for some point $*$.

Given two points $z_1 \sim z_2$, for each n there exist z_1^n and z_2^n, such that, respectively, $F^n(z_1^n) = z_1$ and $F^n(z_2^n) = z_2$, and $z_1^n \sim z_2^n$.

For each pair $z_1 = (a, b_1)$ and $z_2 = (a,\ b_2)$, and $n \geq 0$, the elements z_1^n, z_2^n are of the form $z_1^n = (a^n, b_1^n)$, $z_2^n = (a^n, b_2^n)$.

In this case $T^n(b_1^n) = b_1$ and $T^n(b_2^n) = b_2$.

Note also that $T^n(a) = a^n$.

The distances between b_1^n and b_2^n are exponentially decreasing with n.

We denote

$$\delta(z_1, z_2) = \Pi_{j=1}^{\infty} \frac{T'(b_1^n)}{T'(b_2^n)} < \infty.$$

This product is well defined because

$$\sum_n \log \frac{T'(b_1^n)}{T'(b_2^n)} = \sum_n [\log T'(b_1^n) - \log T'(b_2^n)]$$

converges. This is so because $\log T'$ is Holder and for all n we have $|b_1^n - b_2^n| < \lambda^{-n}$, where $T'(x) > \lambda > 1$ for all x.

This δ is a cocycle.

In the case of Example 18 considered a Holder continuous function $A(a, b)$, where $A : S^1 \times S^1 \to \mathbb{R}$.

Define for $z_1 = (a, b_1)$ and $z_2 = (a, b_2)$

$$\delta(z_1, z_2) = \Pi_{j=1}^{\infty} \frac{e^{A(z_1^n)}}{e^{A(z_2^n)}}.$$

The modular function $\delta(z_1, z_2)$ is well defined because A is Holder.

We will show that δ can be expressed in the form of Example 19. Indeed, fix a certain $b_0 \in S^1$, then, taking $z_1 = (a, b_1)$ consider $z_0 = (a, b_0)$. We denote in an analogous way z_1^n and z_0^n the ones such that $F^n(z_1^n) = z_1$ and $F^n(z_0^n) = z_0$.

Define $V : G^0 \to \mathbb{R}$ by

$$V(z_1) = \Pi_{j=1}^{\infty} \frac{e^{A(z_1^n)}}{e^{A(z_0^n)}}. \tag{4}$$

V is well defined and if $z_1 \sim z_2$ we get that

$$\delta(z_1, z_2) = \frac{V(z_1)}{V(z_2)}.$$

We will show later (see Proposition 11) that $V(a, b)$ does not depend on a, and then we can write $V(b)$, and finally

$$\delta(z_1, z_2) = \frac{V(b_1)}{V(b_2)}.$$

Example 21 Consider a fixed Holder function $\hat{A} : \{1, 2, ..., d\}^{\mathbb{Z}} \to \mathbb{R}$ and the groupoid given by the equivalence relation of Example 3. Denote for any (x, y)

$$\delta(x, y) = \Pi_{j=1}^{\infty} \frac{\hat{A}(\hat{\sigma}^{-j}(s(\gamma)))}{\hat{A}(\hat{\sigma}^{-j}(r(\gamma)))} = \Pi_{j=1}^{\infty} \frac{\hat{A}(\hat{\sigma}^{-j}(x))}{\hat{A}(\hat{\sigma}^{-j}(y))}.$$

The modular function δ is well defined because $\hat{A}$ is Holder. Indeed, this follows from the bounded distortion property.

In a similar way as in the last example one can show that such δ can be expressed on the form of Example 19.

Definition 8 Given a measured groupoid G for the transverse function ν we say that a probability M on G^0 is **quasi-invariant** for ν if there exist a modular function $\delta : G \to \mathbb{R}$, such that, for any integrable function $f : G \to \mathbb{R}$ we have

$$\int \int f(s, x) \nu^x(ds) dM(x) = \int \int f(x, s)\, \delta^{-1}(x, s)\, \nu^x(ds) dM(x). \quad (5)$$

In a more accurate way we say that M is quasi-invariant for the transverse function ν and the modular function δ.

For the existence of quasi-invariant probabilities see [41, 42].

Note that if $\delta(x, s) = \frac{B(x)}{B(s)}$ we get that the above condition (5) can be written as

$$\int \int f(s, x) B(s) \nu^x(ds) dM(x) = \int \int f(x, s)\, B(s)\, \nu^x(ds) dM(x). \quad (6)$$

Indeed, in (5) replace $f(s, x)$ by $B(s) f(s, x)$.

Quasi-invariant probabilities will be also described as the ones which satisfies the so called the KMS condition (on the setting of von Neumann algebras, or C^*-algebras) as we will see later on Sect. 4.

As an extreme example consider the equivalence relation such that each point is related to just itself. In this case a modular function δ takes only the value 1. Given any transverse function ν the condition

$$\int \int f(s, x) \nu^x(ds) dM(x) = \int \int f(x, s)\, \delta^{-1}(x, s)\, \nu^x(ds) dM(x) \quad (7)$$

is satisfied by any probability M on X. In this case the set of probabilities is the set of quasi-invariant probabilities.

Example 22 Quasi invariant probability and the SBR probability for the Baker map

We will present a particular example where we will compare the probability M satisfying the quasi invariant condition with the so called SBR probability. We will consider a different setting of the case described on [58] (considering Anosov systems) which, as far as we know, was never published.

We will show that the **quasi invariant probability is not the SBR probability**.

We will address later on the end of this example the kind of questions discussed on [38, 58].

We will consider the groupoid of Example 7, that is, we consider the equivalence relation: given two points $z_1, z_2 \in S^1 \times S^1$ they are related if the first coordinate is equal.

In Example 7 we consider an expanding transformation $T : S^1 \to S^1$ and F denotes the associated T-Baker map. The associated SBR probability is the only absolutely continuous F-invariant probability over $S^1 \times S^1$.

The dynamical action of F in some sense looks like the one of an Anosov diffeomorphism.

Consider the measured groupoid $(G\,,\,\nu)$ where in each vertical fiber over the point a we set ν^a as the Lebesgue probability db over the class (a, b), $0 \leq b \leq 1$.

This groupoid corresponds to the local unstable foliation for the transformation F.

We fix a certain point $b_0 \in (0, 1)$. For each pair $x = (a, b)$ and $y = (a, b_0)$, where $a, b \in S^1$, and $n \geq 0$, the elements z_1^n, z_2^n, $n \in \mathbb{N}$, are such that $F^n(z_1^n) = x = (a, b)$ and $F^n(z_2^n) = y = (a, b_0)$. Note that they are of the form $z_1^n = (a^n, b^n)$, $z_2^n = (a^n, s^n)$. We use the notation $z_1^n(x)$, $b^n(x)$, $n \in \mathbb{N}$, to express the dependence on x.

We denote for $x \in S^1 \times S^1$

$$V(x) = V(a, b) = \Pi_{n=1}^{\infty} \frac{T'(b^n(x))}{T'(s^n)} = \Pi_{n=1}^{\infty} \frac{T'(b^n(a, b))}{T'(s^n)} < \infty.$$

This is finite because s^n and $b^n(x)$ are on the same domain of injectivity of T for all n and T' is of Holder class.

In a similar fashion as in [58] we define δ by the expression

$$\delta(\,(a, y_1)\,,\,(a, y_2)\,) = \frac{V(a, y_1)}{V(a, y_2)} = \frac{V(y_1)}{V(y_2)} = \Pi_{n=1}^{\infty} \frac{T'(b^n(a, y_1))}{T'(b^n(a, y_2))},$$

where $(a, y_1) \sim (a, y_2)$.

Consider the probability M on $S^1 \times S^1$ given by

$$dM(a, b) = \frac{V(a, b)}{\int V(a, c)dc}\, db\, da.$$

The density $\psi(a, b) = \frac{V(a,b)}{\int V(a,c)dc}$ satisfies the equation

$$\psi(a, b)\frac{1}{T'(b)} = \psi(F(a, b)). \tag{8}$$

Denote $F(a, b) = (\tilde{a}, \tilde{b})$, then, it is known that the density $\varphi(a, b)$ of the SBR probability for F satisfies

$$\varphi(a,b)\frac{T'(\tilde{a})}{T'(b)} = \varphi(F(a,b)). \tag{9}$$

This follows from the F-invariance of the SBR
Therefore, M is not the SBR probability—by uniqueness of the SBR.
We will show that M satisfies the quasi invariant condition.
Note that

$$\int\int f((a,b),(a,s))\ \nu^a(d\,s)\ dM(a,b) =$$

$$\int\int\int f((a,b),(a,s))\ \ \frac{V(a,b)}{\int V(a,c)dc}\,ds\,db\,da.$$

On the other hand

$$\int\int f((a,s),(a,b))\ \frac{V(a,s)}{V(a,b)}\ \nu^a(d\,s)\ dM(a,b) =$$

$$\int\int\int f((a,s),(a,b))\ \frac{V(a,s)}{V(a,b)}\ \ \frac{V(a,b)}{\int V(a,c)dc}\,ds\,db\,da =$$

$$\int\int\int f((a,s),(a,b))\ \ \ \frac{V(a,s)}{\int V(a,c)dc}\,ds\,db\,da.$$

If we exchange the variables b and s, and using Fubini's theorem, we get that M satisfies the quasi invariant condition.

The relation of quasi-invariant probabilities and transverse measures is described on Sect. 5.

The result considered on Theorem 6.18 in [58] for an Anosov diffeomorphism concerns transverse measures and cocycles. [58] did not mention quasi-invariant probabilities.

Note that from Eqs. (8) and (9) one can get that the conditional disintegration along unstable leaves of both the SRB and the quasi-invariant probability M are equal (see page 533 in [32]).

Using the relation of quasi-invariant probabilities, cocycles and transverse measures one can say that one of the main claims in [58] (see Theorem 6.18) and [38] (both considering the case of Anosov Systems) can be expressed in some sense via the above mentioned property about conditional disintegration along unstable leaves (using the analogy with the case of the above Baker map F).

In Sect. 7 we will present more examples of quasi-stationary probabilities.

4 Von Neumann Algebras Derived from Measured Groupoid

We refer the reader to [2, 12, 27] as general references for von Neumann algebras related to groupoids.

Here $X \sim G^0$ will be either $\hat{\Omega}$, Ω or $S^1 \times S^1$. We will denote by G a general groupoid obtained by an equivalence relation R.

Definition 9 Given a measured groupoid G for the transverse function ν and two measurable functions $f, g \in \mathscr{F}_\nu(G)$, we define the convolution $(f \underset{\nu}{*} g) = h$, in such way that, for any $(x, y) \in G$

$$(f \underset{\nu}{*} g)(x, y) = \int g(x, s)\, f(s, y)\, \nu^y(ds) = h(x, y).$$

In the case there exists a multiplicative neutral element for the operation $*$ we denote it by $\mathbb{1}$.

The above expression in some sense resembles the way we get a matrix from the product of two matrices.

For a fixed Haar system ν the product $\underset{\nu}{*}$ defines an algebra on the vector space of ν-integrable functions $\mathscr{F}_\nu(G)$.

As usual function of the form $f(x, x)$ are identified with functions $f : G^0 \to \mathbb{R}$ of the from $f(x)$.

Example 23 In the particular case where ν^y is the counting measure on the fiber over y then

$$(f \underset{\nu}{*} g)(x, y) = \sum_s g(x, s)\, f(s, y).$$

Denote by I_Δ the indicator function of the diagonal on $G^0 \times G^0$. In this case, I_Δ is the neutral element for the product $\underset{\nu}{*}$ operation.

In this case $\mathbb{1} = I_\Delta$.

Note that I_Δ is measurable but generally not continuous. This is fine for the von Neumann algebra setting. However, we will need a different topology (and σ-algebra) on $G^0 \times G^0$—other than the product topology—when considering the unit $\mathbb{1} = I_\Delta$ for the C^*-algebra setting (see [14, 55, 56]).

Remark The indicator function of the diagonal on $G^0 \times G^0$ is not always the multiplicative neutral element on the von Neumann algebra obtained from a general Haar system (G, ν).

Example 24 Another example: consider the standard Haar system of Example 11.

In this case

$$(f \underset{\nu}{*} g)(x,y) = \int g(x,s)\, f(s,y) \nu^y(ds) =$$

$$\frac{1}{d} \sum_{a=1}^{d} g(x,\ (a,x_2,x_3,...))\ f(\,(a,x_2,x_3,...)\,,y)\ = h(x,y).$$

The neutral element is $d\, I_\Delta = \mathbb{1}$.

Example 25 Suppose $J : \{1,2,...,d\}^{\mathbb{N}} \to \mathbb{R}$ is a continuous positive function such that for any $x \in \Omega$ we have that $\sum_{a=1}^{d} J(ax) = 1$. The measured groupoid (G, ν) of Example 14, where ν^y, $y \in \{1,2,...,d\}^{\mathbb{N}}$, is such that given $f, g : G \to \mathbb{R}$, we have for any $(x,y) \in G$, $x = (x_1,x_2,x_3,...)$, $y = (y_1,x_2,x_3,..)$ that

$$(f \underset{\nu}{*} g)(x,y) = \int g(x,s)\, f(s,y) \nu^y(ds) =$$

$$\sum_{a=1}^{d} g(x,\ (a,x_2,x_3,...))\ f(\,(a,x_2,x_3,...)\,,y)\ J(a,x_2,x_3,...) = h(x,y).$$

Note that $x_j = y_j$ for $j \geq 2$.

Suppose that f is such that for any string $(x_2,x_3,...)$ and $a \in \{1,2,.,d\}$ we get

$$f(\,(a,x_2,x_3,...)\,,(a,x_2,x_3,...)\,)\ =\ \frac{1}{J(a,x_2,x_3,...)},$$

and, $a, b \in \{1,2,.,d\}$, $a \neq b$

$$f(\,(a,x_2,x_3,...)\,,(b,x_2,x_3,...)\,)\ =\ 0.$$

In this case the neutral multiplicative element is $\mathbb{1}(x,y) = \frac{1}{J(x)} I_\Delta(x,y)$.

Consider a measured groupoid (G,ν), $\nu \in \mathscr{E}$, then, given two functions ν-integrable $f, g : G \to \mathbb{R}$, we had defined before an algebra structure on $\mathscr{F}_\nu(G)$ in such way that $(f \underset{\nu}{*} g) = h$, if

$$(f \underset{\nu}{*} g)(x,y) = \int g(x,s)\, f(s,y) \nu^y(ds) = h(x,y),$$

where $(x,y) \in G$ and $(s,y) \in G$.

To define the von Neumann algebra associated to (G,ν), we work with complex valued functions $f : G \to \mathbb{C}$. The product is again given by the formula

$$(f \underset{\nu}{*} g)(x,y) = \int g(x,s)\, f(s,y) \nu^y(ds).$$

The involution operation $*$ is the rule $f \to \tilde{f} = f^*$, where $\tilde{f}(x, y) = \overline{f(y, x)}$. The functions $f \in \mathscr{F}(G^0)$ are of the form $f(x) = f(x, x)$ are such that $\tilde{f} = f$.

Following Hanh [25], we define the I-norm

$$\|f\|_I = \max\left\{\left\| y \mapsto \int |f(x, y)|\, \nu^y(dx)\right\|_\infty, \left\| y \mapsto \int |f(y, x)|\, \nu^y(dx)\right\|_\infty\right\},$$

and the algebra $I(G, \nu) = \{f \in L^1(G, \nu) : \|f\|_I < \infty\}$ with the product and involution as above. An element $f \in I(G, \nu)$ defines a bounded operator L_f of left convolution multiplication by a fixed f on $L^2(G, \nu)$. This gives the left regular representation of $I(G, \nu)$.

Definition 10 Given a measured groupoid (G, ν), we define the **von Neumann Algebra associated to** (G, ν), denoted by $W^*(G, \nu)$, as the the von Neumann generated by the left regular representation of $I(G, \nu)$, that is, $W^*(G, \nu)$ is the closure of $\{L_f : f \in I(G, \nu)\}$ in the weak operator topology.

The multiplicative unity is denoted by $\mathfrak{1}$.

In the case ν is such that such that, $\int \nu^y(ds) = 1$, for any $y \in G^0$, we say that the von Neumann algebra is normalized.

In the setting of von Neumann Algebras we do not require that $\mathfrak{1}$ is continuous.

Definition 11 We say an element $h \in W^*(G\,, \nu)$ is positive if there exists a g such that $h = g * \tilde{g}$.

This means

$$h(x, y) = (g \underset{\nu}{*} \tilde{g})(x, y) = \int g(x, s)\, \overline{g(y, s)} \nu^y(ds) = h(x, y).$$

Note que $h(x, x) = (g \underset{\nu}{*} \tilde{g})(x, x) \geq 0$.

The next example is related to Example 16.

Example 26 Consider the equivalence relation where all points are related on the set $G^0 = \{1, 2.., d\}$. In this case $G = \{1, 2.., d\} \times \{1, 2.., d\}$. Take ν as the counting measure. A function $f : G \to \mathbb{C}$ is denoted by $f(i, j)$, where $i \in \{1, 2.., d\}$, $j \in \{1, 2.., d\}$.

The convolution product is

$$(f \underset{\nu}{*} g)(i, j) = \sum_k g(i, k)\, f(k, j).$$

In this case the associated von Neumann algebra (the set of functions $f : G \to \mathbb{C}$) is identified with the set of matrices with complex entries. The convolution is the product of matrices and the identity matrix is the unit $\mathfrak{1}$. The involution operation is to take the Hermitian A^* of a matrix A.

Note that the diagonal elements of a positive matrix (A is of the form $A = B\, B^*$) are non negative real numbers.

Example 27 For the groupoid G of Example 1 and the counting measure, given $f, g : G \to \mathbb{C}$, we have that

$$(f \underset{\nu}{*} g)(x, y) =$$

$$\sum_{a \in \{1,2,..,d\}} g((x_1, x_2, ...), (a, x_2, x_3, ..)) \; f((a, x_2, x_3, ..), (y_1, x_2, ...)).$$

We call standard von Neumann algebra on the groupoid G (of Example 1) the associated von Neumann algebra. For this $W^*(G, \nu)$ the neutral element $\mathtt{1}$ (or, more formally $L_{\mathtt{1}}$) is the indicator function of the diagonal (a subset of G). In this case $\mathtt{1}$ is measurable but not continuous.

Example 28 For the probabilistic Haar system (G, ν) of Example 14, given $f, g : G \to \mathbb{C}$, we get

$$(f \underset{\nu}{*} g)(x, y) =$$

$$\sum_{a \in \{1,2,...,d\}} \varphi(a, x_2, x_3, ...)\, g((x_1, x_2, ...), (a, x_2, x_3..)) \; f((a, x_2, x_3, ...), (y_1, x_2, ...)),$$

where φ is Holder and such that $\sum_{a \in \{1,2,...,d\}} \varphi(a, x_1, x_2, ...) = 1$, for all $x = (x_1, x_2, ...)$.

This φ is a Jacobian.

The neutral element is described in Example 25.

Example 29 In the case $\nu^y = \delta_{x_0}$ for a fixed x_0 independent of y, then

$$(f \underset{\nu}{*} g)(x, y) = g(x, x_0)\, f(x_0, y).$$

Proposition 1 *If (G, ν) is a measured groupoid, then for $f, g \in I(G, \lambda)$.*

$$(f \underset{\nu}{*} g)^{\sim} = \tilde{g} \underset{\nu}{*} \tilde{f}.$$

Proof Remember that for (x, y) in G

$$(f \underset{\nu}{*} g)(x, y) = \int g(x, s)\, f(s, y) \nu^y(ds) = h(x, y).$$

Then,

$$(f \underset{\nu}{*} g)^{\sim}(x, y) = \int \overline{g(y, s)\, f(s, x)} \nu^x(ds).$$

On the other hand

$$(\tilde{g} \underset{\nu}{*} \tilde{f})(y, x) = \int \overline{f(s, x)}\, \overline{g(y, s)} \nu^y(ds).$$

As $\nu^y = \nu^x$ we get that the two expressions are equal. □

Then by Proposition 1 we have for the involution $*$ it is valid the property

$$(f \underset{\lambda}{*} g)^* = g^* \underset{\lambda}{*} f^*.$$

For more details about properties related to this definition we refer the reader to chapter II in [53] and Sect. 5 in [27].

We say that $c : G \to \mathbb{R}$ is a linear cocycle function if $c(x, y) + c(y, z) = c(x, z)$, for all x, y, z which are related. If c is a linear cocycle then e^{δ} is a modular function (or, a multiplicative cocycle).

Definition 12 Consider the von Neumann algebra $W^*(G, \nu)$ associated to (G, ν).

Given a continuous cocycle function $c : G \to \mathbb{R}$ we define the **group homomorphism** $\alpha : \mathbb{R} \to \text{Aut}(W^*(G, \nu))$, where for each $t \in \mathbb{R}$ we have that $\alpha_t \in \text{Aut}(W^*(G, \nu))$ is defined by: for each fixed $t \in \mathbb{R}$ and $f : G \to \mathbb{R}$ we set $\alpha_t(f) = e^{t\,i\,c}\, f$.

Remark Observe that in the above definition that for each fixed $t \in \mathbb{R}$ and any $f : G^0 \to \mathbb{R}$, we have $\alpha_t(f) = f$, since $c(x, x) = 0$ for all $x \in G^0$.

We are particularly interested here in the case where $G^0 = \Omega$ or $G^0 = \hat{\Omega}$.

The value t above is related to temperature and not time. We are later going to consider complex numbers z in place of t. Of particular interest is $z = \beta i$ where β is related to the inverse of temperature in Thermodynamic Formalism (or, Statistical Mechanics).

Definition 13 Consider the von Neumann Algebra $W^*(G, \nu)$ with unity $\mathbb{1}$ associated to (G, ν). A von Neumann **dynamical state** is a linear functional w (acting on the linear space $W^*(G, \nu)$) of the form $w : W^*(G, \nu) \to \mathbb{C}$, such that, $w(a) \geq 0$, if a is a positive element of $W^*(G, \nu)$, and $w(\mathbb{1}) = 1$.

Example 30 Consider over $\Omega = \{1, 2, .., d\}^{\mathbb{N}}$ the equivalence relation R of Example 1 and the Haar system (G, ν) associated to the counting measure in each fiber $r^{-1}(x) = \{(a, x_2, x_3, ...)\, | a \in \{1, 2, ..., d\}\}$, where $x = (x_1, x_2, ...)$.

Given a probability μ over Ω we can define a von Neumann dynamical state φ_μ in the following way: for $f : G \to \mathbb{C}$ define

$$\varphi_\mu(f) = \int f(x, x)\, d\mu(x) = \int f((x_1, x_2, x_3, ...), (x_1, x_2, x_3, ...))\, d\mu(x). \tag{10}$$

If h is positive, that is, of the form $h(x, y) = \int g(x, s)\, \overline{g(y, s)} \nu^y(ds)$, then

$$\varphi_\mu(h) = \int (\int \|g(x, s)\|^2 \nu^x(ds))\, d\mu(x) \geq 0.$$

Note that $\varphi_\mu \mathbb{1} = 1$.

Then, φ_μ is indeed a von Neumann dynamical state.
In this case given $f, g : G \to \mathbb{C}$

$$\varphi_\mu(f \underset{\nu}{*} g) =$$

$$\int \sum_{a\in\{1,2,..,d\}} f((x_1, x_2, ...), (a, x_2, x_3..))\ g((a, x_2, x_3..), (x_1, x_2, ...))\, d\mu(x).$$

It seems natural to try to obtain dynamical states from probabilities M on G^0 (adapting the reasoning of the above example). Then, given a cocycle c it is also natural to ask: what we should assume on M in order to get a KMS state for c?

Example 31 For the von Neumann algebra of complex matrices of Example 26 taking $p_1, p_2, .., p_d \geq 0$, such that $p_1 + p_2 + .. + p_d = 1$, and $\mu = \sum_{j=1}^{d} \delta_j$, we consider φ_μ such that

$$\varphi_\mu(A) = A_{11}p_1 + A_{22}p_2 + ... + A_{dd}p_d,$$

where A_{ij} are the entries of A.

Note first that $\varphi_\mu(I) = 1$.

If $B = A\,A^*$, then the entries $B_{jj} \geq 0$, for $j = 1, 2, ..., d$.

Therefore, φ_μ is a dynamical state on this von Neumann algebra.

Example 32 Consider over $\Omega = \{1, 2, .., d\}^{\mathbb{N}}$ the equivalence relation R of Example 14 and the associated probability Haar system ν.

Given a probability μ over Ω we can define a von Neumann dynamical state φ_μ in the following way: given $f : G \to \mathbb{C}$ we get $\varphi_\mu(f) = \int f(x, x)\, J(x)\, d\mu(x)$. In this way given f, g we have

$$\varphi_\mu(f \underset{\nu}{*} g) =$$

$$\int \sum_{a\in\{1,2,..,d\}} J(a, x_2, ..)\, g((x_1, x_2, ...), (a, x_2, ..))\ f((a, x_2, ..), (x_1, x_2...))\, J(x) d\mu(x).$$

For the neutral multiplicative element $\mathfrak{1}(x, y) = \frac{1}{J(x)} I_\Delta(x, y)$ we get

$$\varphi_\mu(\mathfrak{1}) = \int \frac{1}{J(x)} I_\Delta(x, x)\, J(x) d\mu(x) = 1.$$

Consider G a groupoid and a von Neumann Algebra $W^*(G, \nu)$, where ν is a transverse function, with the algebra product $f \underset{\nu}{*} g$ and involution $f \to \tilde{f}$.

Given a continuous cocycle $c : G \to \mathbb{R}$ we consider $\alpha : \mathbb{R} \to \mathrm{Aut}(W^*(G, \nu))$, $t \mapsto \alpha_t$, the associated homomorphism according to Definition 12: for each fixed $t \in \mathbf{R}$ and $f : G \to \mathbb{R}$ we set $\alpha_t(f) = e^{t\,i\,c}\, f$.

Definition 14 An element $a \in W^*(G, \nu)$ is said to be **analytical** with respect to α if the map $t \in \mathbb{R} \mapsto \alpha_t(a) \in W^*(G, \nu)$ has an analytic continuation to the complex numbers.

More precisely, there is a map $\varphi : \mathbb{C} \to W^*(G, \nu)$, such that, $\varphi(t) = \alpha_t(a)$, for all $t \in \mathbb{R}$, and moreover, for every $z_0 \in \mathbb{C}$, there is a sequence $(a_n)_{n\in\mathbb{N}}$ in $W^*(G, \nu)$, such that, $\varphi(z) = \sum_{n=0}^{\infty}(z - z_0)^n a_n$ in a neighborhood of z_0.

The analytical elements are dense on the von Neumann algebra (see [49]).

Definition 15 We say that a von Neumann dynamical state w is a **KMS state for** β **and** c if

$$w\,(\,b \underset{\nu}{*} (\alpha_{i\,\beta}(a))\,) = w\,(a \underset{\nu}{*} b\,),$$

for any b and any analytical element a.

It follows from general results (see [49]) that it is enough to verify: for any $f, g \in I(G, \nu)$ and $\beta \in \mathbb{R}$ we get

$$w\,(\,g \underset{\nu}{*} \alpha_{\beta\,i}(\,f\,)\,) = w\,(\,g \underset{\nu}{*} (\,e^{-\beta\,c}\,f\,)\,) = w\,(f \underset{\nu}{*} g\,). \tag{11}$$

Consider the functions

$$u(x, y) = (f * g)(x, y) = \int g(x, s)\, f(s, y)\, \nu^y(ds),$$

and

$$v(x, y) = (\,g * (e^{-\beta\,c}\,f)\,)\,(x, y) = \int e^{-\beta c(x,s)} f(x, s)\, g(s, y)\, \nu^y(ds).$$

Equation (11) means

$$w(\,u(x, y)\,) = w(\,v(x, y)\,). \tag{12}$$

Note that Eq. (11) implies that a KMS von Neumann (or, C^*)-dynamical state w satisfies:

(a) for any $f : G^0 \to \mathbb{C}$ and $g : G \to \mathbb{C}$:

$$w\,(\,g \underset{\nu}{*} f\,) = w\,(f \underset{\nu}{*} g\,). \tag{13}$$

This follows from the fact that for any $t \in \mathbb{R}$ and any $f : G^0 \to \mathbb{R}$, we have that $\alpha_t(f) = f$.

(b) if the function $\mathtt{1}$ depends just on $x \in G^0$, then, for any β

$$\alpha_{i\,\beta}(\mathtt{1}) = \mathtt{1}.$$

(c) w is invariant for the group α_t, $t \in \mathbb{R}$. Indeed,

$$w(\alpha_t(f)) = w(1 \underset{\nu}{*} \alpha^t(f)) = w(f \underset{\nu}{*} 1) = w(f).$$

Example 33 For the von Neumann algebra (C^*-algebra) of complex matrices of Examples 26 and 31 consider the dynamical evolution $\sigma_t = e^{i\,t\,H}$, $t \in \mathbb{R}$, where H is a diagonal matrix with entries the real numbers $H_{11} = U_1$, $H_{22} = U_2$, ..., $H_{dd} = U_d$. The KMS state ρ for β is

$$\rho(A) = A_{1\,1}\rho_1 + A_{2\,2}\rho_2 + ... + A_{d\,d}\rho_d,$$

where $\rho_i = \frac{e^{-\beta U_i}}{\sum_{j=1}^{d} e^{-\beta U_j}}$, $i = 1, 2, ..., d$, and $A_{i,j}$, $i, j = 1, 2, ..., d$, are the entries of the matrix A (see [56]).

The probability μ of Example 30 corresponds in some sense to the probability $\mu = (\rho_1, \rho_2, .., \rho_d)$ on $\{1, 2, ..., d\}$. That is, $\rho = \varphi_\mu$.

This is a clear indication that the μ associated to the KMS state has in some sense a relation with Gibbs probabilities. This property will appear more explicitly on Theorem 1 for the case of the bigger than two equivalence relation.

Remember that if c is a cocycle, then $c(x, z) = c(x, y) + c(y, z)$, $\forall x \sim y \sim z$, and, therefore,

$$\delta(x, y) = e^{\beta c(x,y)} = e^{-\beta c(y,x)}$$

is a modular function.

Definition 16 Given a cocycle $c : G \to \mathbb{R}$ we say that a probability M over G^0 satisfies the (c, β)-KMS condition for the groupoid (G, ν), if for any $h \in I(G, \nu)$, we have

$$\int \int h(s, x)\nu^x(ds)dM(x) = \int \int h(x, s)e^{-\beta\, c(x,s)}\, \nu^x(ds)dM(x), \qquad (14)$$

where $\beta \in \mathbb{R}$.

In this case we will say that M is a **KMS probability**.

The above means that M **is quasi-invariant for** ν **and** $\delta(x, s) = e^{-\beta c(s,x)}$.

When $\beta = 1$ and c is of the form $c(s, x) = V(x) - V(s)$ the above condition means

$$\int \int h(s, x)\nu^x(ds)e^{V(x)}dM(x) = \int \int h(x, s)e^{V(x)}\, \nu^x(ds)dM(x). \qquad (15)$$

Proposition 2 *(J. Renault—Proposition II.5.4 in [53]) Suppose that the state w is such that for a certain probability μ on G^0 we have that for any $h \in I(G, \nu)$ we get $w(h) = \int h(x, x)d\mu(x)$. Then, to say that μ satisfies the (c, β)-KMS condition for (G, ν) according to Definition 16 is equivalent to say that w is KMS for (G, ν), c and β, according to Eq. (11).*

Proof Note that for any f, g

$$(f \underset{\nu}{*} g)(x, y) = \int g(x, s)\, f(s, y) dv^x(s)$$

and

$$(g \underset{\nu}{*} (e^{-\beta c} f))(x, y) = \int f(x, s)\, g(s, y)\, e^{-\beta c(x,s)}\, dv^x(s).$$

We have to show that $\int u(x, x) d\mu(x) = \int v(x, x) d\mu(x)$ (see Eq. (12)).

Then, if the (c, β)-KMS condition for M is true, we take $h(s, x) = g(x, s)\, f(s, x)$ and we got Eq. (12) for such w.

By the other hand if (12) is true for such w and any f, g, then take $f(s, x) = h(s, x)$ and $g(s, x) = 1$. □

Example 34 In the case for each y we have that v^y is the counting measure we get that to say that a probability M over $\hat{\Omega}$ satisfies the (c, β)-KMS condition means: for any $h : G \to \mathbb{C}$

$$\sum_{y \sim x} \int h(x, y) e^{-\beta c(x,y)} dM(x) = \sum_{x \sim y} \int h(x, y)\, dM(y). \tag{16}$$

In the notation of [54] we can write the above in an equivalent way as

$$\int h\, e^{-\beta c} d(s^*(M)) = \int h\, d(r^*(M)).$$

Note that in [54] it is considered $r(x, y) = x$ and $s(x, y) = y$.

Suppose $c(x, y) = \varphi(x) - \varphi(y)$. Then, taking $h(x, y) = k(x, y)\, e^{\beta \varphi(x)}$ we get an equivalent expression for (16): for any $k(x, y)$

$$\sum_{y \sim x} \int k(x, y) e^{\beta \varphi(y)} dM(x) = \sum_{x \sim y} \int k(x, y)\, e^{\beta \varphi(x)} dM(y). \tag{17}$$

For a Holder continuous potential $A : \{1, 2.., d\}^{\mathbb{N}} \to \mathbb{R}$ the Ruelle operator $\mathscr{L}_A$ acts on continuous functions $v : \{1, 2.., d\}^{\mathbb{N}} \to \mathbb{R}$ by means of $\mathscr{L}_A(v) = w$, if

$$\mathscr{L}_A(v)(x_1, x_2, x_3, ...) = \sum_{a=1}^{d} e^{A(a, x_1, x_2, x_3, ...)} v(a, x_1, x_2, x_3, ...) = w(x).$$

For a Holder continuous potential $A : \{1, 2.., d\}^{\mathbb{N}} \to \mathbb{R}$ there exist a continuous positive eigenfunction f, such that, $\mathscr{L}_A(f) = \lambda\, f$, where λ is positive and also the spectral radius of $\mathscr{L}_A$ (see [47]).

The dual $\mathscr{L}_A^*$ of $\mathscr{L}_A$ acts on probabilities by Riesz Theorem (see [47]). We say that the probability m on $\{1, 2.., d\}^{\mathbb{N}}$ is **Gibbs for the potential** A, if $\mathscr{L}_A^*(m) = \lambda\, m$ (same λ as above). In this case we say that m is an eigenprobability for A.

Gibbs probabilities for Holder potentials A are also **DLR probabilities** on $\{1, 2, ...d\}^{\mathbb{N}}$ (see [11]).

Gibbs probabilities for Holder potentials A can be also obtained via Thermodynamic Limit from boundary conditions (see [11]).

We say that the **potential** A **is normalized** if $\mathscr{L}_A(1) = 1$. In this case a probability μ is Gibbs (equilibrium) for the normalized potential A if it is a fixed point for the dual of the Ruelle operator, that is, $\mathscr{L}_A^*(\mu) = \mu$.

Suppose $\Omega = \{-1, 1\}^{\mathbb{N}}$ and $A : \Omega \to \mathbb{R}$ is of the form

$$A(x_0, x_1, x_2, ...) = x_0a_0 + x_1a_1 + x_2a_2 + x_3a_3 + ... + x_na_n + ...$$

where $\sum a_n$ is absolutely convergent.

In [10] the explicit expression of the eigenfunction for $\mathscr{L}_A$ and the eigenprobability for the dual $\mathscr{L}_A^*$ of the Ruelle operator $\mathscr{L}_A$ is presented. The eigenprobability is not invariant for the shift.

In Example 27 consider $\Omega = \{1, 2\}^{\mathbb{N}}$ and take ν^y the counting measure on the class of y. Consider the von Neumann algebra associated to this measured groupoid (G, ν) where G is given by the bigger than two relation.

In this case $\mathtt{1}(x, y) = I_\Delta(x, y)$.

Consider $c(x, y) = \varphi(x) - \varphi(y)$, where φ is Holder. We do not assume that φ is normalized.

A natural question is: the eigenprobability μ for such potential φ is such that $f \to \varphi_\mu(f) = \int f(x, x)\, d\mu(x)$ defines the associated KMS state? For each modular function c?

The purpose of the next results is to analyze this question when $c(x, y) = \varphi(x) - \varphi(y)$.

Consider the equivalence relation on $\Omega = \{1, 2..., d\}^{\mathbb{N}}$ which is

$$x = (x_1, x_2, x_3, ..) \sim y = (y_1, y_2, y_3, ...)\ ,\ \text{if an only if}\ ,\ x_j = y_j \quad \text{for all} \quad j \geq 2.$$

In this case the class $[x]$ of $x = (x_1, x_2, x_3, ..)$ is

$$[x] = \{\, (1, x_2, x_3, ..),\ (2, x_2, x_3, ..), ..., (d, x_2, x_3, ..)\,\}.$$

The associated groupoid by $G \subset \Omega \times \Omega$, is

$$G = \{(x, y) \,|\, x \sim y\}.$$

G is a closed set on the compact set $\Omega \times \Omega$. We fix the measured groupoid (G, ν) where ν^x is the counting measure. The results we will get are the same if we take the Haar system as the one where each point y on the class of x has mass $1/d$.

In this case Eq. (14) means

$$\sum_j \int f(\,(j, x_2, x_3, .., x_n, ..)\,,\ (x_1, x_2, x_3, .., x_n, ..))dM(x) =$$

$$\sum_j \int f(\,(x_1, x_2, x_3, ..)\,,\ (j, x_2, x_3, ..))e^{-c(j,x_2,x_3,..)\,,\,(x_1,x_2,x_3,..))}dM(x). \qquad (18)$$

The first question: given a cocycle c does there exist M as above?
Suppose $c(x, y) = \varphi(y) - \varphi(x)$.
In this case Eq. (18) means

$$\sum_j \int f(\,(j, x_2, x_3, .., x_n, ..)\,,\ (x_1, x_2, x_3, .., x_n, ..))dM(x) =$$

$$\sum_j \int f(\,(x_1, x_2, x_3, ..)\,,\ (j, x_2, x_3, ..))e^{-\varphi(j,x_2,x_3,..)+\varphi(x_1,x_2,x_3,..)}dM(x). \qquad (19)$$

Among other things we will show later that if we assume that φ depends just on the first coordinate then we can take M as the independent probability (that is, such independent M satisfies the KMS condition (19)).

In Sect. 3.4 in [56] and in [29] the authors present a result concerning quasi-invariant probabilities and Gibbs probabilities on $\{1, 2, ..d\}^{\mathbb{N}}$ which has a different nature when compared to the next one. The groupoid is different from the one we will consider (there elements are of the form (x, n, y), $n \in \mathbb{Z}$). In [29, 56] for just one value of β you get the existence of the quasi invariant probability. Moreover, the KMS state is unique (here this will not happen as we will show on Theorem 3).

In [26, 33, 43, 57] the authors present results which have some similarities with the next theorem. They consider Gibbs (quasi-invariant) probabilities in the case of the symbolic space $\{1, 2, ..d\}^{\mathbb{Z}}$ and not $\{1, 2, ..d\}^{\mathbb{N}}$ like here. In all these papers the quasi-invariant probability is unique and invariant for the shift. In [7] the authors consider DLR probabilities for interactions in $\{1, 2, ..d\}^{\mathbb{Z}}$. The equivalence relation (the homoclinic relation of Example 6) in all these cases is quite different from the one we will consider.

The next result were generalized in [34, 35].

Theorem 1 *Consider the Haar system with the counting measure ν for the bigger than two relation on $\{1, 2, ..d\}^{\mathbb{N}}$. Suppose that φ depends just on the first k coordinates, that is, and*

$$\varphi(x_1, x_2, ..., x_k, x_{k+1}, x_{k+2}, ..) = \varphi(x_1, x_2, ..., x_k).$$

Then, the eigenprobability μ (a DLR probability) for the potential $-\varphi$ (that is, $\mathcal{L}^*_{-\varphi}(\mu) = \lambda\mu$, for some positive λ) satisfies the KMS condition (is quasi-invariant) for the associated modular function $c(x, y) = \varphi(y) - \varphi(x)$.

The same result is true, of course, for βc, where $\beta > 0$.

Proof We are going to show that the Gibbs probability μ for the potential $-\varphi$ satisfies the KMS condition.

We point out that in general the eigenvalue $\lambda \neq 1$.

We have to show that (19) is true when $M = \mu$. That is, μ is a KMS probability for the Haar system and the modular function.

Denote for any finite string $a_1, a_2, ...a_n$ and any n

$$p_{a_1,a_2,...a_n} = \frac{e^{-[\varphi(a_1,a_2,...a_n,\,1^\infty)+\varphi(a_2,...a_n,\,1^\infty)+...+\varphi(a_n,\,1^\infty)]}}{\sum_{b_1,b_2,...b_n} e^{-[\varphi(b_1,b_2,...b_n,\,1^\infty)+\varphi(b_2,...b_n,\,1^\infty)+...+\varphi(b_n,\,1^\infty)]}}.$$

Note that for $n > k$ we have that

$$e^{-[\,\varphi(a_1,a_2,...a_n,\,1^\infty)+\varphi(a_2,...a_n,\,1^\infty)+...+\varphi(a_n,\,1^\infty)\,]} =$$

$$e^{-[\,\varphi(a_1,a_2,...a_k)+\varphi(a_2,...a_{k+1})+...+\varphi(a_n,\underbrace{1,\,1,\,..,\,1}_{k-1})+(n-k)\varphi(\underbrace{1,\,1,\,..,\,1}_{k-1})]}.$$

Therefore,

$$\sum_{b_1,b_2,...b_n} e^{-[\,\varphi(b_1,b_2,...b_n,\,1^\infty)+\varphi(b_2,...b_n,\,1^\infty)+...+\varphi(b_n,\,1^\infty)]} =$$

$$e^{-(n-k)\varphi(\underbrace{1,\,1,\,..,\,1}_{k-1})} \sum_{b_1,b_2,...b_n} e^{-[\,\varphi(b_1,b_2,...b_k)+\varphi(b_2,...b_{k+1})+...+\varphi(b_n,\underbrace{1,\,1,\,..,\,1}_{k-1})\,]}.$$

Consider the probability μ_n, such that,

$$\mu_n = \sum_{a_1,a_2,...a_n} \delta_{(a_1,a_2,...a_n,\,1^\infty)}\; p_{a_1,a_2,...a_n} =$$

$$\sum_{a_1,...a_n} \delta_{(a_1,...a_n,\,1^\infty)} \frac{e^{-[\,\varphi(a_1,...a_k)+\varphi(a_2,...a_{k+1})+...+\varphi(a_n,\underbrace{1,\,..,\,1}_{k-1})]}}{\sum_{b_1,...b_n} e^{-[\,\varphi(b_1,...b_k)+\varphi(b_2,...b_{k+1})+...+\varphi(b_n,\underbrace{1,\,..,\,1}_{k-1})]}}.$$

and μ such that $\mu = \lim_{n\to\infty} \mu_n$.

Note that

$$p_{a_1,a_2,..,a_n} = \frac{e^{-[\,\varphi(a_1,...a_k)+\varphi(a_2,...a_{k+1})+...+\varphi(a_n,\underbrace{1,..,1}_{k-1})\,]}}{\sum_{b_1,...b_n} e^{-[\,\varphi(b_1,...b_k)+\varphi(b_2,...b_{k+1})+...+\varphi(b_n,\underbrace{1,..,1}_{k-1})\,]}}.$$

If φ is Holder it is known that the above probability μ is the eigenprobability for the dual of the Ruelle operator $\mathscr{L}_{-\varphi}$ (a DLR probability). That is, there exists $\lambda > 0$ such that $\mathscr{L}^*_{-\varphi}(\mu) = \lambda\mu$. This follows from the Thermodynamic Limit with boundary condition property as presented in [11].

We claim that the above probability μ satisfies the KMS condition.

Indeed, note that

$$\sum_j \int f(\,(j, x_2, x_3, .., x_n, ..)\,,\ (x_1, x_2, x_3, .., x_n, ..))d\mu(x) =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1,a_2,...a_n} f(\,(j, a_2, a_3, .., a_n,\ 1^\infty)\,,\ (a_1, a_2, a_3, .., a_n,\ 1^\infty)\,)\, p_{a_1,a_2,...,a_n} =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1} \sum_{a_2,...a_n} f((j, a_2, a_3, .., a_n,\ 1^\infty)(a_1, a_2, a_3, .., a_n, 1^\infty) p_{a_1,a_2,...,a_n}. \quad (20)$$

On the other hand

$$\sum_j \int f(\,(x_1, x_2, x_3, ..)\,,\ (j, x_2, x_3, ..))e^{-\varphi(j,x_2,x_3,...)+\varphi(x_1,x_2,x_3,..)} d\mu(x) =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1,a_2,...,a_n} f(\,(a_1, .., a_n,\ 1^\infty)\,,\ (j, a_2, .., a_n,\ 1^\infty)\,)e^{-\varphi(j,a_2,...,a_k)+\varphi(a_1,a_2,...a_k)} p_{a_1,a_2,...,a_n} =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1} \sum_{a_2,...a_n} f(\,(a_1, .., a_n,\ 1^\infty)\,,\ (j, a_2, .., a_n,\ 1^\infty)\,)e^{-\varphi(j,a_2,...,a_k)+\varphi(a_1,a_2,...a_k)}$$

$$\frac{e^{-[\,\varphi(a_1,a_2,...a_k)+\varphi(a_2,...a_{k+1})+...+\varphi(a_n,\underbrace{1,..,1}_{k-1})]}}{\sum_{b_1,...b_n} e^{-[\,\varphi(b_1,...b_k)+\varphi(b_2,...b_{k+1})+...+\varphi(b_n,\underbrace{1,..,1}_{k-1})]}} =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1} \sum_{a_2,...a_n} f(\,(a_1, .., a_n,\ 1^\infty)\,,\ (j, a_2, .., a_n,\ 1^\infty)\,)$$

$$\frac{e^{-[\varphi(j,a_2,,...a_k)+\varphi(a_2,...a_{k+1})+...+\varphi(a_n,\underbrace{1,..,1}_{k-1})]}}{\sum_{b_1,...b_n} e^{-[\varphi(b_1,...b_k)+\varphi(b_2,...b_{k+1})+...+\varphi(b_n,\underbrace{1,..,1}_{k-1})]}}. =$$

$$\lim_{n\to\infty} \sum_j \sum_{a_1} \sum_{a_2,...a_n} f((a_1,..,a_n,\ 1^\infty),\ (j,a_2,..,a_n,\ 1^\infty))\, p_{j,a_2,...,a_n}.$$

On this last equation if we exchange coordinates j and a_1 we get expression (20). Then, such μ satisfies the KMS condition. □

The above theorem can be extended to the case the potential φ is Holder. We refer the reader to [34] for more general results. This paper consider a more general relation defining the so called continuous groupoids. In the case the transverse function is not the counting measure (for instance when each class is not a countable set) a similar kind of results as above can be shown using Thermodynamic Formalism for the so called generalized XY model (see [35]).

We will show now that under the above setting the KMS probability **is not unique**.

Proposition 3 *Suppose μ satisfies the KMS condition for the measured groupoid (G, ν) where $c(x, y) = \varphi(y) - \varphi(x)$. Suppose φ is normalized for the Ruelle operator, where $\varphi : G^0 = \Omega \to \mathbb{R}$. Consider $v(x_1, x_2, x_3, ..)$ which does not depend of the first coordinate. Then, $v(x)d\mu(x)$ also satisfies the KMS condition for the measured groupoid (G, ν).*

Proof Suppose μ satisfies the (c, β)-KMS condition for the measured groupoid (G, ν). This means: for any $g \in I(G, \nu)$

$$\int \sum_{a\in\{1,2,..,d\}} g((a, y_2, y_3..),\ (y_1, y_2, ...))\, e^{\beta\varphi(a,y_2,y_3..)}\, d\mu(y) =$$

$$\int \sum_{a\in\{1,2,..,d\}} g((x_1, x_2, ...), (a, x_2, x_3..))\, e^{\beta\varphi(a,x_2,x_3..)}\, d\mu(x). \tag{21}$$

Take

$$h(x_1, x_2, x_3, ..), (y_1, y_2, y_3, ..)) =$$

$$k((x_1, x_2, x_3, ..), (y_1, y_2, y_3, ..))\, v(x_1, x_2, x_3, ..)) =$$

$$k((x_1, x_2, x_3, ..), (y_1, y_2, y_3, ..))\, v(x_2, x_3, ..)).$$

From the hypothesis about μ we get that

$$\int \sum_{a\in\{1,2,..,d\}} h((a, y_2, y_3..),\ (y_1, y_2, ...))\, e^{\beta\varphi(a,y_2,y_3..)}\, d\mu(y) =$$

$$\int \sum_{a\in\{1,2,..,d\}} h((x_1,x_2,x_3,...),(a,x_2,x_3..))\, e^{\beta\varphi(a,x_2,x_3..)}\, d\mu(x).$$

This means, for any continuous k the equality

$$\int \sum_{a\in\{1,2,..,d\}} k((a,y_2,y_3..),(y_1,y_2,...))\, e^{\beta\varphi(a,y_2,y_3..)}\, v(y_2,y_3,...)d\mu(y) =$$

$$\int \sum_{a\in\{1,2,..,d\}} k((x_1,x_2,x_3,...),(a,x_2,x_3..))\, e^{\beta\varphi(a,x_2,x_3..)}\, v(x_2,x_3,..)d\mu(x).$$

Therefore, $v(x)d\mu(x)$ also satisfies the (c,β)-KMS condition for the measured groupoid (G,ν).

□

It follows from the above result that the probability that satisfies the KMS condition for c and the measured groupoid (G,ν) is not always unique.

A probability ρ satisfies the Bowen condition for the potential $-\varphi$ if there exists constants $c_1, c_2 > 0$, and P, such that, for every

$$x = (x_1,...,x_m,...) \in \Omega = \{1,2,...d\}^{\mathbb{N}},$$

and all $m \geq 0$,

$$c_1 \leq \frac{\rho\{y : y_i = x_i, \quad \forall i = 1,...,m\}}{\exp\left(-Pm - \sum_{k=1}^{m} \varphi(\sigma^k(x))\right)} \leq c_2. \tag{22}$$

Suppose φ is Holder, then, if ρ is the equilibrium probability (or, if ρ is the eigenprobability for the dual of Ruelle operator $\mathscr{L}_{-\varphi}$) one can show that it satisfies the Bowen condition for $-\varphi$.

In the case v is continuous and does not depend on the first coordinate then $v(x)d\mu(x)$ also satisfies the Bowen condition for φ. The same is true for the probability $\hat{\rho}$ of Example 35 on the case $-\varphi = \log J$.

There is an analogous definition of the Bowen condition on the space $\{1,2,...d\}^{\mathbb{Z}}$ but it is a much more strong hypothesis on this case (see Sect. 5 in [33]).

Example 35 We will show an example where the probability μ of Theorem 1 (the eigenprobability for the potential $-\varphi$) is such that if f is a function that depends just on the first coordinate, then, $f\,\mu$ does not necessarily satisfies the KMS condition.

Suppose $\varphi = -\log J$, where $J(x_1,x_2,x_3,..) = J(x_1,x_2) > 0$, and $\sum_i P_{i,j} = 1$, for all i. In other words the matrix P, with entries $P_{i,j}$, $i,\ j \in \{1,2..,d\}$, is a column stochastic matrix. The Ruelle operator for $-\varphi$ is the Ruelle operator for $\log J$. The potential $\log J$ is normalized for the Ruelle operator.

We point out that in Stochastic Process it is usual to consider line stochastic matrices which is different from our setting.

There exists a unique right eigenvalue probability vector π for P (acting on vectors on the right). The Markov chain determined by the matrix P and the initial vector of probability $\pi = (\pi_1, \pi_2, .., \pi_d)$ determines an stationary process, that is, a probability ρ on the Bernoulli space $\{1, 2, ..., d\}^{\mathbb{N}}$, which is invariant for the shift acting on $\{1, 2, ..., d\}^{\mathbb{N}}$.

For example, we have that $\rho(\overline{21}) = P_{21}\pi_1$.

We point out that such ρ is the eigenprobability for the $\mathscr{L}^*_{\log J}$ (associated to the eigenvalue 1). Therefore, ρ satisfies the KMS condition from the above results.

The Markov Process determined by the matrix P and the initial vector of probability $\pi = (1/d, 1/d, ..., 1/d)$ defines a probability $\hat{\rho}$ on the Bernoulli space $\{1, 2, ..., d\}^{\mathbb{N}}$, which is not invariant for the shift acting on $\{1, 2, ..., d\}^{\mathbb{N}}$.

In this case, for example, $\hat{\rho}(\overline{21}) = P_{21}\, 1/d$.

Note that the probability ρ satisfies $\rho = u\,\hat{\rho}$ where u depends just on the first coordinate.

Note that unless P is double stochastic is not true that for any j_0 we have that $\sum_k P_{j_0,k} = 1$.

Assume that there exists j_0 such that $\sum_k P_{j_0,k} \neq 1$.

We will check that, in this case $\hat{\rho}$ does not satisfies the KMS condition for the function $f(x, y) = I_{X_1=i_0}(x)\, I_{X_1=j_0}(y)$.

Indeed, Eq. (19) means

$$\sum_j \int f(\,(j, x_2, x_3, .., x_n, ..)\,,\ (x_1, x_2, x_3, .., x_n, ..))d\hat{\rho}(x) =$$

$$\sum_j \int I_{X_1=i_0}(j, x_2, x_3, ..)\, I_{X_1=j_0}(x_1, x_2, x_3...)d\hat{\rho}(x) =$$

$$\int I_{X_1=i_0}(i_0, x_2, x_3, ..)\, I_{X_1=j_0}(x_1, x_2, x_3...)d\hat{\rho}(x) =$$

$$\int I_{X_1=j_0}(x_1, x_2, x_3...)d\hat{\rho}(x) = \hat{\rho}(\overline{j_0}) = 1/d =$$

$$\sum_j \int f(\,(x_1, x_2, x_3, ..)\,,\ (j, x_2, x_3, ..))e^{-\varphi(j,x_2,x_3,..)+\varphi(x_1,x_2,x_3,..)}d\hat{\rho}(x) =$$

$$\sum_j \int I_{X_1=i_0}(x_1, x_2, x_3, ..)\, I_{X_1=j_0}(j, x_2, x_3...)e^{-\varphi(j,x_2,x_3,..)+\varphi(x_1,x_2,x_3,..)}d\hat{\rho}(x) =$$

$$\int I_{X_1=i_0}(x_1, x_2, x_3, ..)\, I_{X_1=j_0}(j_0, x_2, x_3...)e^{-\varphi(j_0,x_2,x_3,..)+\varphi(x_1,x_2,x_3,..)}d\hat{\rho}(x) =$$

$$\int I_{X_1=i_0}(x_1, x_2, x_3, ..)\, e^{-\varphi(j_0,x_2,x_3,..)+\varphi(x_1,x_2,x_3,..)} d\hat{\rho}(x) =$$

$$\int_{X_1=i_0} e^{-\varphi(j_0,x_2,x_3,..)+\varphi(i_0,x_2,x_3,..)} d\hat{\rho}(x) =$$

$$\sum_k \int_{X_1=i_0\,,\, X_2=k} e^{-\varphi(j_0,x_2,x_3,..)+\varphi(i_0,x_2,x_3,..)} d\hat{\rho}(x) =$$

$$\sum_k \int_{X_1=i_0\,,\, X_2=k} P_{j_0,k}\ P_{i_0,k}^{-1} d\hat{\rho}(x) =$$

$$\sum_k P_{j_0,k}\ P_{i_0,k}^{-1}\, P_{i_0,k} 1/d =$$

$$\sum_k P_{j_0,k}\ 1/d\ \neq 1/d = \hat{\rho}(\overline{j_0}).$$

Therefore, $\hat{\rho}$ does not satisfies the KMS condition.

Example 36 Consider $\Omega = \{1, 2\}^{\mathbb{N}}$, a Jacobian J and take ν^y the probability on each class y given by $\sum_a J(a, y_2, y_3, ..)\delta_{(a,y_2,y_3,..)}$.

Note first that $\varphi = \log J$ is a normalized potential. Does the equilibrium probability for $\log J$ satisfies the KMS condition? We will show that this in not always true.

The question means: is it true that for any function k is valid

$$\int \sum_{a\in\{1,2\}} k(\,(a, y_2, ..),\ (y_1, y_2, ..)\,)\, e^{\varphi(a,y_2,...)}\, d\mu(y) =$$

$$\int \sum_{a\in\{1,2\}} k(\,(a, x_2, ..),\ (x_1, x_2, ..)\,)\, e^{\varphi(a,x_2,...)}\, d\mu(x) =$$

$$\int \sum_{a\in\{1,2\}} k(\,(x_1, x_2, ..),\ (a, x_2, ..)\,)\, e^{\varphi(a,x_2,...)}\, d\mu(x)? \tag{23}$$

Consider the example: take $c(x, y) = \varphi(x) - \varphi(y)$, for $\varphi : \{1, 2\}^{\mathbb{N}} \to \mathbb{R}$, such that,

$$\varphi(a,\ .\,,\ .\,, ...) = \log p$$

where $p = p_a$, for $a \in \{1, 2\}$, and $p_1 + p_2 = 1,\ p_1, p_2 > 0$.

The Gibbs probability μ for such φ is the independent probability associated to p_1, p_2.

Given such probability μ over Ω we can define a dynamical state φ_μ in the following way: given $f : G \to \mathbb{R}$ we get $\varphi_\mu(f) = \int f(x, x)\ d\mu(x)$.

Take $\beta = 1$. We will show that φ_μ is not KMS for c.

The Eq. (23) for such μ means for any $k(x, y)$

$$\int \sum_{a\in\{1,2\}} p_a\, k(\,(a, y_2, ..),\ (y_1, y_2, ..)\,)\ d\mu(y) =$$

$$\int \sum_{a\in\{1,2\}} k(\,(a, y_2, ..),\ (y_1, y_2, ..)\,)\, p(a, y_2, ...)\, d\mu(y) =$$

$$\int \sum_{b\in\{1,2\}} k(\,(x_1, x_2, ..),\ (b, x_2, ..)\,)\, p(b, x_2, ...)\, d\mu(x) =$$

$$\int \sum_{b\in\{1,2\}} p_b\, k(\,(x_1, x_2, ..),\ (b, x_2, ..)\,)\ d\mu(x).$$

It is not true that μ is Gibbs for the potential $\log p$.

Indeed, given k consider the function

$$g(y_1, y_2, y_3, y_4, ...) = k(\,(y_1, y_3, ...)\,,\ (y_2, y_3, ...)\,).$$

Note that

$$\mathscr{L}_{\log p}(g)(y_1, y_2, y_3, ..) = \sum_{a\in\{1,2\}} p(a, y_1, y_2, y_3, ...)\, g(a, y_1, y_2, ...)$$

$$\sum_{a\in\{1,2\}} p_a\, k(\,(a, y_2, y_3, ..),\ (y_1, y_2, ..)\,).$$

Then,

$$\int \sum_{a\in\{1,2\}} p_a\, k(\,(a, y_2, ..),\ (y_1, y_2, ..)\,)\ d\mu(y) =$$

$$\int \mathscr{L}_{\log p}(g)(y_1, y_2, y_3, ..)\, d\mu(y) = \int k(\,(y_1, y_3, ...)\,,\ (y_2, y_3, ...)\,)\, d\mu(y).$$

Now, given k consider the function

$$h(x_1, x_2, x_3, x_4, ...) = k(\,(x_2, x_3, ...)\,,\ (x_1, x_3, x_4, ...)\,).$$

Then,

$$\int \mathscr{L}_{\log p}(h)(x_1, x_2, x_3, ..)\, d\mu(y) = \int k(\,(x_2, x_3, ...)\,,\ (x_1, x_3, ...)\,)\, d\mu(x).$$

For the Gibbs probability μ for $\log p$ is not true that for all k

$$\int k(\,(x_2, x_3, ...)\,,\ (x_1, x_3, ...)\,)\, d\mu(x) = \int k(\,(x_1, x_3, ...)\,,\ (x_2, x_3, ...)\,) d\mu(x).$$

5 Noncommutative Integration and Quasi-invariant Probabilities

In non-commutative integration the transverse measures are designed to integrate transverse functions (see [12] or [27]).

In the same way we can say that a function can be integrated by a measure resulting in a real number we can say that the role of a transverse measure is to integrate transverse functions (producing a real number).

The main result here is Theorem 2 which describes a natural way to define a transverse measure from a modular function δ and a Haar system $(G\,, \hat{\nu})$.

We refer the reader to [35] for new results on the topic (for instance related to the entropy of transverse measures, etc.)

As a motivation for the topic of this section consider a foliation of the two dimensional torus where we denote each leave by l. This partition defines a grupoid with a quite complex structure. Each leave is a class on the associated equivalence relation. This motivation is explained with much more details in [13].

We consider in each leave l the intrinsic Lebesgue measure on the leave which will be denoted by ρ_l.

A random operator q is the association of a bounded operator $q(l)$ on $\mathscr{L}^2(\rho_l)$ for each leave l. We will avoid to describe several technical assumptions which are necessary on the theory (see page 51 in [13]).

The set of all random operators defines a von Neumann algebra under some natural definitions of the product, etc. (see Proposition 2 in page 52 in [13]) (*).

This setting is the formalism which is natural on **noncommutative geometry** (see [13]).

Important results on the topic are for instance the characterization of when such von Neumann algebra is of type I, etc. (see page 53 in [13]). There is a natural trace defined on this von Neumann algebra.

A more abstract formalism is the following: consider a fixed grupoid G. Given a transverse function λ one can consider a natural operator $L_\lambda : \mathscr{F}^+(G) \to \mathscr{F}^+(G)$, which satisfies

$$f \to \lambda * f = L_\lambda(f)\,.$$

L_λ acts on $\mathscr{F}^+(G)$ and can be extended to a linear action on the von Neumann algebra $\mathscr{F}(G)$. This defines a Hilbert module structure (see Sect. 3.2 in [31] or [27]).

Given λ we can also define the operator $R_\lambda : \mathscr{F}^+(G) \to \mathscr{F}^+(G)$ by

$$h(x, y) = R_\lambda(f)(\gamma) = \int f(s, y) d\lambda^x(d\, s),$$

for any (x, y).

Definition 17 Given two G-kernels λ_1 and λ_2 we get a new G-kernel $\lambda_1 * \lambda_2$, called the convolution of λ_1 and λ_2, where given the function $f(x, y)$, we get the rule

$$(\lambda_1 * \lambda_2)\,(f) = g \in \mathscr{F}(G^0),$$

given by

$$g(y) = \int \Big(\int f(s, y)\ \lambda_2^x(d\, s) \Big) \quad \lambda_1^y(dx).$$

In the above $x \sim y \sim s$.

In other words $(\lambda_1 * \lambda_2)$ is such that for any y we have

$$(\lambda_1 * \lambda_2)^y(dx) = \int \lambda_2^x(d\, s)\ \ \lambda_1^y(dx). \tag{24}$$

Note that

$$R_{\lambda_1 * \lambda_2} = R_{\lambda_1} \circ R_{\lambda_2}.$$

For a given fixed transverse function λ, for each class $[y]$ on the grupoid G, we get that R_λ defines an operator acting on functions $f(r, s)$, where $f : [y] \times [y] \to \mathbb{C}$, and where $R_\lambda(f) = h$.

In this way, each transverse function λ defines a random operator q, where $q([y])$ acts on $\mathscr{L}^2(\lambda^y)$ via R_λ.

A transverse measure can be seen as an integrator of transverse functions or as an integrator of random operators (which are elements on the von Neumann algebra (*) we mention before).

First we will present the basic definitions and results that we will need later on this section.

Remember that $\mathscr{E}^+$ is the set of transverse functions for the grupoid $G \subset X \times X$ associated to a certain equivalence relation $\sim$.

$\mathscr{F}^+(G)$ denotes the space of Borel measurable functions $f : G \to [0, \infty)$ (a real function of two variables (a, b)).

Definition 18 Given a G kernel ν and an integrable function $f \in \mathscr{F}_\nu(G)$ we can define two functions on G:

$$(x, y) \to (\nu * f)(x, y) = \int f\,(x, s)\, \nu^y(ds)$$

and

$$(x, y) \to (f * \nu)(x, y) = \int f(s, y)\, \nu^x(ds).$$

Note that $\nu * 1 = 1$ if ν^y is a probability for all y. Also note that $(f * \nu)(y, y) = \nu(f)\ (y)$ (see Definition 4).

About (24) we observe that

$$(\lambda_1 * \lambda_2)\,(f) = \lambda_1(f * \lambda_2).$$

A kind of analogy of the above concept of convolution (of kernels) with integral kernels is the following: given the kernels $K_1(s, x)$ and $K_2(x, y)$ we define the kernel

$$\hat{K}(s, y) = \int K_2(s, x)\, K_1(x, y)\, dx.$$

This is a kind of convolution of integral kernels.

This defines the operator

$$f(x, y) \to g(y) = \int f(s, y)\hat{K}(s, y)ds = \int (\int f(s, y)\, K_2(s, x)\, d\, s\,)\ K_1(x, y)\, dx.$$

Example 37 Given any kernel ν we have that $\mathfrak{d} * \nu = \nu$, where $\mathfrak{d}$ is the delta kernel of Example 10.

Indeed, for any $f \in \mathscr{F}(G)$

$$\int f\,(\mathfrak{d} * \nu)^y = \int \int f(s, y)\, \nu^x(d\, s)\, \mathfrak{d}^y(d\, x) = \int f(s, y)\, \nu^y(d\, s) = \int f\, \nu^y$$

In the same way for any ν we have that $\nu * \mathfrak{d} = \nu$.

Example 38 Given a fixed positive function $h(x, y)$ and a fixed kernel ν, we get that the kernel $\nu * (h\,\mathfrak{d}\,)$, where $\mathfrak{d}$ is the Dirac kernel, is such that given any $f(x, y)$,

$$(\nu * (h\,\mathfrak{d}\,))(f)(y) = \int (\int f(s, y)\, h(s, x)\, \mathfrak{d}^x(d\, s)\,)\quad \nu^y(dx) =$$

$$\int f(x, y)\, h(x, x)\quad \nu^y(dx).$$

Particularly, taking $h = 1$, we get $\nu * \mathfrak{d} = \nu$.

Example 39 For the bigger than two equivalence relation of Example 17 on $(S^1)^{\mathbb{N}}$, where S^1 is the unitary circle, the equivalence classes are of the form $\{\,(a, x_2, x_3, ...),\ a \in S^1\,\}$, where $x_j \in S^1$, $j \geq 2$, is fixed.

Given $x = (x_1, x_2, x_3, ...)$ we define $\nu^x(da)$ the Lebesgue probability on S^1, which can be identified with $S^1 \times (x_2, x_3, .., x_n, ...)$. This defines a transverse function where $G^0 = (S^1)^{\mathbb{N}}$. We call it the **standard XY Haar system**.

In this case given a function $f(x, y) = f(\,(x_1, x_2, x_3, ..), (y_1, y_2, y_3, ..)\,)$

$$(\nu * f)(x, y) = \int f(x, s)\, \nu^x(d\, s) = \int f\,(\,(x_1, x_2, x_3, ..), (s, x_2, x_3, ..)\,)\, ds,$$

where $s \in S^1$. Note that in the present example the information on y was lost after convolution.

Such ν is called in [6] the a priori probability for the Ruelle operator. Results about Ruelle operators and Gibbs probabilities for such kind of XY models appear in [6, 36].

After Proposition 8 we will present several properties of convolution of transverse function (we will need soon some of them).

Note that if ν is transverse and λ is a kernel, then $\nu * \lambda$ is transverse.

Remember that given a kernel λ and a fixed y the property $\lambda^y(1) = 1$ means $\int \lambda^y(dx) = 1$.

Definition 19 A transverse measure Λ over the modular function $\delta(x, y)$, $\delta : G \to \mathbb{R}$, is a linear function $\Lambda : \mathscr{E}^+ \to \mathbb{R}^+$, such that, for each kernel λ which satisfies the property $\lambda^y(1) = 1$, for any y, if ν_1 and ν_2 are transverse functions such that $\nu_1 * (\delta\lambda) = \nu_2$, then,

$$\Lambda(\nu_1) = \Lambda(\nu_2). \tag{25}$$

A measure produces a real number from the integration of a classical function (which takes values on the real numbers), and, on the other hand, the transverse measure produces a real number from a transverse function ν (which takes values on measures).

The assumptions on the above definition are necessary (for technical reasons) when considering the abstract concept of integral of a transverse function by Λ (as is developed in [12]). We will show later that there is a more simple expression providing the real values of such process of integration by Λ which is related to quasi-invariant probabilities.

If one consider the equivalence relation such that each point is related just to itself, any cocycle is constant equal 1 and the only kernel satisfying $\lambda^y(1) = 1$, for any y, is the delta Dirac kernel $\mathfrak{d}$. In this case if ν_1 and ν_2 are such that $\nu_1 * (\delta\lambda) = \nu_2$, then, $\nu_1 = \nu_2$ (see Example 37). Moreover, $\mathscr{E}^+$ is just the set of positive functions on X. Finally, we get that the associated transverse measure Λ is just a linear function $\Lambda : \mathscr{E}^+ \to \mathbb{R}^+$.

Example 40 Given a probability μ over G^0 we can define

$$\Lambda(\nu) = \int \int \nu^y(dz) d\mu(y).$$

Suppose that λ satisfies $\lambda^x(1) = 1$, for any x, and

$$\nu_1 * \lambda = \nu_2.$$

Then, $\Lambda(\nu_1) = \Lambda(\nu_2)$. This means that Λ is invariant by translation on the right side.

Indeed, note that,

$$\Lambda(\nu_1) = \int \int \nu_1^y(dz) d\mu(y),$$

and, moreover

$$\Lambda(\nu_2) = \int \int \nu_2^y(dz) d\mu(y) =$$

$$\int [\int (\int \lambda^x(d\,s))) \ \nu_1^y(dx))] d\mu(y) =$$

$$\int (\int \ \nu_1^y(dx)) \ d\mu(y).$$

Therefore, Λ is a transverse measure of modulus $\delta = 1$.

In this way for each measure μ on G^0 we can associate a transverse measure of modulus 1 by the rule $\nu \to \Lambda(\nu) = \int \int \nu^y(dz) d\mu(y) \in \mathbb{R}$.

The condition

$$\nu_1 * (\delta\lambda) = \nu_2$$

means for any f we get

$$\int f(x, y) \, (\nu_1 * (\delta\lambda))^y (d\,x) = \int (\int f(s, y) \, [\, \delta(s, x) \lambda^x(ds) \,]) \quad \nu_1^y(dx) =$$

$$\int f(x, y) \ \nu_2^y(dx). \tag{26}$$

We define before (see (5)) the concept of quasi-invariant probability for a given modular function δ, a grupoid G and a fixed transverse function ν.

For reasons of notation we use a little bit variation of that definition. In this section we say that M is quasi invariant probability for δ and ν if for any $f(x, y)$

$$\int \int f(y, x) \, \nu^y(x) dM(y) = \int \int f(x, y) \delta(x, y)^{-1} \nu^y(x) dM(y). \tag{27}$$

Proposition 4 *Given a modular function δ, a grupoid G and a fixed transverse function $\hat{\nu}$ denote by M the quasi invariant probability for δ.*

Assume that $\int \hat{\nu}^y(dr) \neq 0$ *for all* y.
If $\hat{\nu} * \lambda_1 = \hat{\nu} * \lambda_2$, *where* λ_1, λ_2 *are kernels, then,*

$$\int \delta^{-1}\lambda_1(1)\ d\,M = \int \delta^{-1}\lambda_2(1)\ d\,M.$$

This is equivalent to say that

$$\int\int \delta^{-1}(s,y)\lambda_1^y(ds)dM(y) = \int\int \delta^{-1}(s,y)\lambda_2^y(ds)dM(y).$$

Proof By hypothesis $g(y) = (\hat{\nu} * \lambda_1)(\delta^{-1})\,(y) = (\hat{\nu} * \lambda_2)(\delta^{-1})(y)$.
Then, we assume that (see (24))

$$\int g(y)\frac{1}{\int \hat{\nu}^y(dr)}dM(y) = \int\int\int \frac{1}{\int \hat{\nu}^y(dr)}\delta^{-1}(s,y)\lambda_1^x(ds)\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^y(dr)}\delta^{-1}(s,y)\lambda_2^x(ds)\hat{\nu}^y(dx)\,dM(y). \tag{28}$$

Therefore,

$$\int\int \delta^{-1}(s,y)\lambda_1^y(ds)dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^x(dr)}\delta(y,s)\lambda_1^y(ds)\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^y(dr)}[\,\delta(y,x)\,\delta(x,s)\,\lambda_1^y(ds)\,]\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^y(dr)}[\,\delta(x,y)^{-1}\,\delta(x,s)\,\lambda_1^y(ds)\,]\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^x(dr)}[\,\delta(y,x)^{-1}\,\delta(y,s)\,\lambda_1^x(ds)\,]\,\delta^{-1}(x,y)\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^x(dr)}\delta(y,s)\lambda_1^x(ds)\hat{\nu}^y(dx)\,dM(y) \tag{29}$$

On the above from the fourth to the fifth line we use the quasi-invariant expression (27) for M taking

$$f(y,x) = \int \frac{1}{\int \hat{\nu}^y(dr)}\,\delta(x,y)^{-1}\,\delta(x,s)\,\lambda_1^y(ds).$$

Note that if $\hat{\nu}$ is transverse $\int \hat{\nu}^x(dr)$ does not depends on x on the class $[y]$.
Finally, from the above equality (29) (and replacing λ_1^x by λ_2^x) it follows that

$$\int\int \delta^{-1}(s,y)\lambda_1^y(ds)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^y(dr)}\delta(y,s)\lambda_1^y(ds)\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int\int \frac{1}{\int \hat{\nu}^y(dr)}\delta(y,s)\lambda_2^y(ds)\hat{\nu}^y(dx)\,dM(y) =$$

$$\int\int \delta^{-1}(s,y)\lambda_2^y(ds)\,dM(y).$$

□

From now on we assume that $\int \hat{\nu}^y(dr) \neq 0$ for all y.

Theorem 2 *Given a modular function δ and a Haar system $(G, \hat{\nu})$, suppose M is quasi invariant for δ.*

We define Λ on the following way: given a transverse function ν there exists a kernel ρ such that $\nu = \hat{\nu} * \rho$ by Proposition 9. We set

$$\Lambda(\nu) = \int\int \delta(x,y)^{-1}\rho^y(dx)dM(y). \tag{30}$$

Then, Λ is well defined and it is a transverse measure.

Proof Λ is well defined by Proposition 4.

We have to show that if $\lambda^x(1) = 1$, for any x, and ν_1 and ν_2 are such that $\nu_1 * (\delta\lambda) = \nu_2$, then, $\Lambda(\nu_1) = \Lambda(\nu_2)$.

Suppose $\nu_1 = \hat{\nu} * \lambda_1$, then, $\nu_2 = \hat{\nu} * (\lambda_1 * (\delta\,\lambda))$.

Note that

$$\Lambda(\nu_1) = \int\int \delta(x,y)^{-1}\,\lambda_1^y(dx)dM(y).$$

On the other hand from (24)

$$\Lambda(\nu_2) = \int\int \delta(s,y)^{-1}\,(\lambda_1 * (\delta\,\lambda))^y(ds)dM(y) =$$

$$\int\int\int \delta(s,y)^{-1}\,\delta(s,x)\,\lambda^x(ds)\,\lambda_1^y(dx)\,dM(y) =$$

$$\int\int\int \delta(x,y)^{-1}\,\lambda^x(ds)\,\lambda_1^y(dx)\,dM(y) =$$

$$\int\int \delta(x,y)^{-1}\ \left(\int \lambda^x(ds)\right)\ \lambda_1^y(dx)\,dM(y) =$$

$$\int\int \delta(x,y)^{-1}\ \lambda_1^y(dx)\,dM(y) = \Lambda(\nu_1).$$

□

Remark The last proposition shows that given a quasi invariant probability M—for a transverse function $\hat{\nu}$ and a cocycle δ—there is a natural way to define a transverse measure Λ (associated to a grupoid G and a modular function δ).

One can ask the question: given transverse measure Λ (associated to a grupoid G and a modular function δ) is it possible to associate a probability on G_0? In the affirmative case, is this probability quasi invariant? We will elaborate on that.

Definition 20 Given a transverse measure Λ for δ we can associate by Riesz Theorem to a transverse function $\hat{\nu}$ a measure M on G^0 by the rule: given a non-negative continuous function $h : G^0 \to \mathbb{R}$ we will consider the transverse function $h(x)\,\hat{\nu}^y(dx)$ and set

$$h \to \Lambda(h\,\hat{\nu}) = \int h(x)dM(x).$$

Such M is a well defined measure (a bounded linear functional acting on continuous functions) and we denote such M by $\Lambda_{\hat{\nu}}$.

$\Lambda_{\hat{\nu}}$ means the rule $h \to \Lambda(h\,\hat{\nu}) = \Lambda_{\hat{\nu}}(h)$.

Proposition 5 *Given any transverse measure Λ associated to the modular function δ and any transverse functions ν and ν' we have for any continuous f that*

$$\Lambda_{\nu'}(\nu(\tilde{\delta}f)) = \Lambda(\nu(\tilde{\delta}f)\nu') = \Lambda(\nu'(\tilde{f})\nu) = \Lambda_{\nu}(\nu'(\tilde{f})).$$

Proof If $\lambda^y(1) = 1\ \forall y$, that is, $\int 1\lambda^y(ds) = 1\ \ \forall y$, then $\Lambda(\nu * \delta\lambda) = \Lambda(\nu)$. If $g(x) = \lambda^x(1) = \int 1\lambda^x(ds) \neq 1$, then we can write $\lambda'^x(ds) = \frac{1}{g(x)}\lambda^x(ds)$, where λ and λ' are just kernels. In this way $(\nu * \delta\lambda) = (g\nu) * \delta\lambda'$. Indeed, for $h(x,y)$,

$$\int h(x,y)\,(\nu * \delta\lambda)^y(dx) = \int h(s,y)\delta(s,x)\lambda^x(ds)\nu^y(dx)$$

$$= \int h(s,y)\delta(s,x)\lambda'^x(ds)g(x)\nu^y(dx) = \int h(x,y)((g\nu) * \delta\lambda')^y(dx).$$

Denoting $\lambda(1)(x) = g(x) = \lambda^x(1) = \int 1\lambda^x(ds)$, it follows that

$$\Lambda(\nu * \delta\lambda) = \Lambda(g\nu * \delta\lambda') = \Lambda(g\nu) = \Lambda_\nu(g) = \Lambda_\nu(\lambda(1)) = \Lambda_\nu(\int 1\lambda^x(ds)). \tag{31}$$

From, (2) if ν is a kernel and $f = f(x, y)$

$$(\nu * f)(x, y) = \nu(\tilde{f})(x),$$

and, from Lemma 2, if λ is a kernel and ν is a transverse function, then, for any $f = f(x, y)$,

$$\lambda * (f\nu) = (\lambda * f)\nu.$$

It follows that, for transverse functions ν and ν', we get

$$\nu * [(\delta \tilde{f})\nu'] = [\nu * (\delta \tilde{f})]\nu' = [\nu(\tilde{\delta} f)]\nu'.$$

As a consequence

$$\Lambda_{\nu'}(\nu(\tilde{\delta} f)) = \Lambda([\nu(\tilde{\delta} f)]\nu') = \Lambda(\nu * [(\delta \tilde{f})\nu']) = \Lambda(\nu * \delta(\tilde{f}\nu')) =$$

$$\Lambda_{\nu}((\tilde{f}\nu')(1)) = \Lambda_{\nu}(\int 1 \cdot \tilde{f}(s, y)\nu'^{y}(ds)) = \Lambda_{\nu}(\nu'(\tilde{f})).$$

Above we use Eq. (31) with $\lambda = \tilde{f}\nu'$. □

Corollary 1 *If $\nu \in \mathscr{E}^{+}$, then for any f*

$$\Lambda(\nu(\tilde{f})\,\nu) \;=\; \Lambda(\nu(\tilde{\delta}\, f)\;\nu) \tag{32}$$

Proof Just take $\nu = \nu'$ on last Proposition. □

Among other things we are interested on a modular function δ, a transverse function $\hat{\nu}$ and a transverse measure Λ (of modulo δ) such that $M_{\Lambda,\hat{\nu}} = M$ is Gibbs for a Jacobian J. What conditions are required from M?.

The main condition of the next theorem is related to the KMS condition of Definition 16.

Proposition 6 *Given a transverse measure Λ associated to the modular function δ, and a* **transverse function** *$\hat{\nu}$, consider the associated $M = \Lambda_{\hat{\nu}}$. Then, M is quasi invariant for δ. That is, M satisfies for all g*

$$\int \int g(s, x)\hat{\nu}^{x}(ds)dM(x) = \int \int g(x, s)\delta(x, s)\,\hat{\nu}^{x}(ds)dM(x). \tag{33}$$

Proof First we point out that (37) is consistent with (27) (we are just using different variables).

A transverse function $\hat{\nu}$ defines a function of $f \in \mathscr{F}(G) \to \mathscr{F}(G^0)$.

The probability M associated to $\hat{\nu}$ satisfies for any continuous function $h(x)$, where $h : G^0 \to \mathbb{R}$ the rule

$$h \to \Lambda(h\,\hat{\nu}) = \int h(x) dM(x),$$

where $h(x)\,\hat{\nu}^y(dx) \in \mathscr{E}^+$.

From Proposition 1 we have that for the continuous function $f(s,x) = \tilde{g}(s,x)$, where $f : G \to \mathbb{R}$, the expression.

$$\Lambda(\hat{\nu}\ (g)\,\hat{\nu}\,) = \Lambda(\hat{\nu}\ (\tilde{f})\,\hat{\nu}\,) = \Lambda(\hat{\nu}(\delta^{-1} f)\ \hat{\nu}\ ) = \Lambda(\hat{\nu}(\delta^{-1}\,\tilde{g})\ \hat{\nu}\ )$$

For a given function $g(s,x)$ it follows from the above that

$$\Lambda(\hat{\nu}\ (g)\,\hat{\nu}\,) = \int \hat{\nu}\ (g)(x)\, dM(x) = \int \int g(s,x)\hat{\nu}^x(ds)\, dM(x).$$

On the other hand

$$\Lambda(\hat{\nu}(\delta^{-1}\,\tilde{g})\ \hat{\nu}\ ) = \int \hat{\nu}(\delta^{-1}\,\tilde{g})(x) d\,M(x) = \int g(x,s)\delta^{-1}(s,x)\,\hat{\nu}^x(ds)\, dM(x).$$

□

Proposition 7 *Given a modular function δ, a grupoid G, a transverse measure Λ and a transverse function $\hat{\nu}$, suppose for any ν, such that $\nu = \hat{\nu} * \rho$, we have that*

$$\Lambda(\nu) = \Lambda(\hat{\nu} * \rho) = \int \int \delta(s,x)^{-1}\rho^x(ds) d\mu(x) = \int \delta^{-1}\rho(1)\, d\,\mu.$$

Then, $\mu = \Lambda_{\hat{\nu}}$.

Proof Given $f \in \mathscr{F}(G_0)$ consider λ the kernel such that $\lambda^x(ds) = f(x)\delta_x(ds)$, where δ_x is the Delta Dirac on x.

Then, using the fact that $\delta(x,x) = 0$ we get that the kernel $f(x)\hat{\nu}^y(dx)$ is equal to $\hat{\nu} * \delta\,\lambda$.

Then, taking $\rho = \delta\,\lambda$ on the above expression we get

$$\Lambda(\,f\hat{\nu}) = \Lambda(\hat{\nu} * (\delta\,\lambda)\,) =$$

$$\int \delta^{-1}\rho(1) d\mu = \int \delta^{-1}\,(\delta\lambda)(1) d\mu = \int \lambda(1) d\mu = \int f(x) d\mu(x).$$

Therefore, $\Lambda_{\hat{\nu}} = \mu$. □

Now we present a general procedure to get transverse measures.

Proposition 8 *For a fixed modular function δ we can associate to any given probability μ over G^0 a transverse measure Λ by the rule*

$$\nu \to \Lambda(\nu) = \int \int \delta(s,x)^{-1}\, \nu^x(ds)\ \ d\mu(x). \tag{34}$$

Proof Consider $\nu' \in \mathscr{E}^+$ and λ, such that, $\int \lambda^r(ds) = 1$, for all r, and moreover that $\nu' = \nu * (\delta\lambda)$.

We will write

$$(\nu * \delta\lambda)\,(\delta^{-1}) = \int \int \delta^{-1}(s,x)\,\delta(s,r)\,\lambda^r(ds)\nu^x(dr)$$

which is a function of x

Then

$$\Lambda(\nu') = \int \int \delta(s,x)^{-1}\, \nu'^x(ds)\ \ d\mu(x) = \int \nu'(\delta^{-1})(x)d\mu(x) =$$

$$= \int \Big(\nu * (\delta\lambda)\Big)(\delta^{-1})(x)d\mu(x) = \int\int\int \delta(s,x)^{-1}\delta(s,r)\lambda^r(ds)\nu^x(dr)d\mu(x) =$$

$$= \int\int\int \delta(r,x)^{-1}\lambda^r(ds)\nu^x(dr)d\mu(x) = \int\int \delta(r,x)^{-1}\nu^x(dr)d\mu(x) = \Lambda(\nu).$$

□

This last transverse measure is defined in a quite different way that the one described on Theorem 2.

Now we will present some general properties of convolution of transverse functions.

Lemma 1 *Suppose $\nu \in \mathscr{E}^+$ is a transverse function, ν_0 a kernel, and $g \in \mathscr{F}^+(G)$ is such that $\int g(s,x)\nu_0^y(dx) = 1$, for all s, y. Then, $\nu_0 * (g\,\nu) = \nu$, where $g\,\nu$ is a kernel.*

Remark The condition $\int g(s,x)\nu_0^y(dx) = 1$, for all s, y means $(\nu_0 * g)(s,y) = 1$ for all s, y, that is $\nu_0 * g \equiv 1$ (see Lemma 3).

Proof

$$z(y) = \int f(s,y)\ \nu^y(ds) =$$

$$\int f(s,y)\ [\int g(s,x)\nu_0^y(dx)]\,\nu^x(ds) \quad =$$

$$\int\ \int f(s,y)\ [\,g(s,x)\nu^x(ds)\,]\quad \nu_0^y(dx) =$$

$$\int f(s,y)\,(g\,\nu)^x(ds)\nu_0^y(dx) = \int f(s,y)\,(\,\nu_0 * (g\,\nu))^y(ds).$$

□

We say that the kernel ν is fidel if $\int \nu_0^y(ds) \neq 0$ for all y.

Proposition 9 *For a fixed transverse function ν_0 we have that for each given transverse function ν there exists a kernel λ, such that, $\nu_0 * \lambda = \nu$.*

Proof Given the kernel ν_0 take $g_0(s) = \frac{1}{\int 1\nu_0^s(dr)} \geq 0$. Note that $g_0(v)$ is constant for $v \in [s]$. Then $\nu_0(g) = 1$, that is, for each s we get that $\int g_0(s)\nu_0^s(dx) = 1$.

We can take $\lambda = g_0\, \nu$ as a solution. Indeed, in a similar way as last lemma we get

$$z(y) = \int f(s, y)\ \nu^y(ds) =$$

$$\int f(s, y)\ [\int g_0(s)\nu_0^s(dx)]\, \nu^x(ds) \quad =$$

$$\int \int f(s, y)\ [\, g_0(s)\nu^x(ds)\,] \quad \nu_0^y(dx) =$$

$$\int f(s, y)\, (g_0\, \nu)^x(ds)\nu_0^y(dx) = \int f(s, y)\, (\, \nu_0 * (g_0\, \nu))^y(ds) =$$

$$\int f(s, y)\, (\, \nu_0 * \lambda)^y(ds).$$

□

The next Lemma is just a more general form of Lemma 1.

Lemma 2 *Suppose $\nu \in \mathscr{E}^+$, $g \in \mathscr{F}^+(G)$ and λ a kernel, then $\lambda * (g\, \nu) = (\lambda * g)\ \nu$, where $g\, \nu$ is a kernel and $\lambda * g$ is a function.*

Proof Given $f \in \mathscr{F}(G)$ we get

$$(\lambda * (g\, \nu))(f)(y) = \int f(x, y)\, (\lambda * (g\, \nu))^y(dx)$$

$$= \int \int f(s, y)[(g\nu)^x(ds)]\, \lambda^y(dx) = \int \int f(s, y)[g(s, x)\nu^x(ds)]\, \lambda^y(dx).$$

On the other hand

$$[(\lambda * g)\ \ \nu](f)(y) = \int f(s, y)[(\lambda * g)\, \nu]^y(ds) = \int f(s, y)[(\lambda * g)(s, y)]\nu^y(ds)$$

$$= \int f(s, y)[\int g(s, x)\lambda^y(dx)]\nu^y(ds) = \int \int f(s, y)g(s, x)\nu^x(ds)\lambda^y(dx).$$

□

Proposition 10 *Suppose ν and λ are transverse. Given $f \in \mathscr{F}^+(G)$, we have that*

$$\lambda\,(\nu * f) \;=\; \nu\,(\lambda * \tilde{f}).$$

Proof Indeed, by Definition 18, $(\nu * f)(x, y) = g(x, y) = \int f(x, s)\nu^y(ds)$, and by Definition 4

$$\lambda\,(\nu * f)(y) = \lambda(g)(y) = \int g(x, y)\lambda^y(dx) = \int\int f(x, s)\nu^y(ds)\lambda^y(dx).$$

By the same arguments $(\lambda * \tilde{f})(x, y) = h(x, y) = \int \tilde{f}(x, s)\lambda^y(ds)$, and

$$\nu\,(\lambda * \tilde{f})(y) = \nu(h)(y) = \int h(x, y)\nu^y(dx) = \int\int \tilde{f}(x, s)\lambda^y(ds)\nu^y(dx) =$$

$$= \int\int f(s, x)\lambda^y(ds)\nu^y(dx) = \lambda\,(\nu * f)(y),$$

if we exchange the coordinates x and s.

Note that in the case $f \in \mathscr{F}(G^0)$ we denote $f(x, s) = f(x)$. In the same way

$$\lambda\,(\nu * f) \;=\; \nu\,(\lambda * \tilde{f})$$

in the following sense:

$$\int\int f(x)\nu^y(ds)\lambda^y(dx) = \int\int f(s)\lambda^y(ds)\nu^y(dx).$$

□

6 C^*-Algebras Derived from Haar Systems

In this section the functions $f : G \to \mathbb{R}$ will be required to be continuous (not just measurable).

An important issue here is that we need suitable hypotheses in such way that the indicator of the diagonal $\imath$ belongs to the underlying space we consider. On von Neumann algebras setting the unit is just measurable and not continuous (this is good enough). We want to consider another setting (certain C^*-algebras associated to Haar Systems) where the unit will be required to be a continuous function. In general terms, given a groupoid $G \subset \Omega \times \Omega$, as we will see, we will need another topology (not the product topology) on the set G for the C^*-Algebra formalism and for defining KMS states.

We will begin with some more examples. The issue here is to set a certain appropriate topology.

Example 41 For $n \in \mathbb{N}$ we define the partition η_n over $\overrightarrow{\Omega} = \{1, 2, ..., d\}^{\mathbb{N}}$, $d \geq 2$, such that two elements $x \in \overrightarrow{\Omega}$ and $y \in \overrightarrow{\Omega}$ are on the same element of the partition, if and only if, $x_j = y_j$, for all $j > n$. This defines an equivalence relation denoted by R_n.

Example 42 We define a partition η over $\overrightarrow{\Omega}$, such that two elements $x \in \overrightarrow{\Omega}$ and $y \in \overrightarrow{\Omega}$ are on the same element of the partition, if and only if, there exists an n such that $x_j = y_j$, for all $j > n$. This defines an equivalence relation denoted by R_∞.

Example 43 For each fixed $n \in \mathbb{Z}$ consider the equivalence relation on $\hat{\Omega}$: $x \sim y$ if

$$y = (..., y_{-n}, ..., y_{-2}, y_{-1} \mid y_0, y_1, ..., y_n, ...)$$

is such that $x_j = y_j$ for all $j \leq n$, where $\hat{\Omega} = \overleftarrow{\Omega} \times \overrightarrow{\Omega}$.

This defines a groupoid.

Example 44 Recall that by definition the unstable set of the point $x \in \hat{\Omega}$ is the set

$$W^u(x) = \{y \in \hat{\Omega}\,, \text{ such that } \lim_{n\to\infty} d(\hat{\sigma}^{-n}(x), d(\hat{\sigma}^{-n}(y))\,) = 0\,\}$$

One can show that the unstable manifold of $x \in \hat{\Omega}$ is the set

$$W^u(x) = \{y = (..., y_{-n}, ..., y_{-2}, y_{-1} \mid y_0, y_1, ..., y_n, ...) \mid \text{there exists}$$

$$k \in \mathbb{Z}, \text{ such that } x_j = y_j, \text{for all } j \leq k\}.$$

If we denote by G_u the groupoid defined by the above relation, then, $x \sim y$, if and only if $y \in W^u(x)$.

Definition 21 Given the equivalence relation R, when the quotient $\hat{\Omega}/R$ (or, $\overrightarrow{\Omega}/R$) is Hausdorff and locally compact we say that R is a proper equivalence.

For more details about proper equivalence see Sect. 2.6 in [56].

On the set $X = \overrightarrow{\Omega}$, if we denote $x = (x_1, x_2, ..x_n, ..)$, the family $U_x(m) = \{y \in \overrightarrow{\Omega}$, such that, $y_1 = x_1, y_2 = x_2, ..., y_m = x_m\}$, $m = 1, 2, ...$, is a fundamental set of open neighbourhoods on Ω.

Considering the relations R_m and R_∞ we get the corresponding groupoids

$$G_1 \subset G_2 \subset ... \subset G_m \subset ... \subset G_\infty \subset \overrightarrow{\Omega} \times \overrightarrow{\Omega} = X \times X.$$

The equivalence relation described in Example 42 (and also 3) is not proper if we consider the product topology on $\overrightarrow{\Omega}$ (respectively on $\hat{\Omega}$). The equivalence relation

described in Example 41 (and also 43) is proper if we consider the product topology on $\overrightarrow{\Omega}$ (respectively on $\hat{\Omega}$) (see [21]).

We consider over G_n the quotient topology.

Lemma 3 *Given $X = \overrightarrow{\Omega}$, for each n the map defined by the canonical projection $X \to G_n$ is open.*

Proof Given an open set $U \subset X$ take $V = \{y \in X \mid$ there exists $x \in X$, satisfying $y \sim x$ for the relation $R_n\}$. We will show that V is open.

Consider $y \in V$, $y \in U$, such that, $y \sim x$ for the relation R_n. There exists $m > n$, such that, $U_x(m) \subset U$. Then, $U_y(m) \subset V$. Indeed, if $z \in U_y(m)$, take $z' \in X$, such that $z'_j = x_j$, when $1 \leq j \leq m$, and $z'_j = z_j$, when $j > m$.

Then, $z' \sim z$ for the relation R_n. But, as $z \in U_y(m)$, this implies that $z_j = y_j$, when $1 \leq j \leq m$, and $y \sim x$, for R_n, implies that $y_j = x_j$, when $j > n$. Then, $z'_j = x_j$, if $1 \leq j \leq m$. Therefore, $z' \in U_x(m) \subset U$. □

Lemma 4 *Given $X = \overrightarrow{\Omega}$, for each n the map defined by the canonical projection $X \to G_\infty$ is open.*

Proof Given an open set $U \subset X$ take $V = \{y \in X \mid$ there exists $x \in X$, satisfying $y \sim x$ for the relation $R_n\}$ and $V_\infty = \{y \in X \mid$ there exists $x \in X$, satisfying $y \sim x$ for the relation $R_\infty\}$. Then, $V_\infty = \cup_{n=1}^{\infty} V_n$ is open. □

Lemma 5 *Given $X = \overrightarrow{\Omega}$, for each $n = 1, 2..., n, ...$, the set G_n is Hausdorff.*

Proof Given a fixed n, and $x, y \in X$, such that x and y are not related by R_n, then, there exists $m > n$ such that $x_m \neq y_m$. From this follows that no element of $U_x(m)$ is equivalent by R_n to an element of $U_y(m)$. By Lemma 3 it follows that G_n is Hausdorff. □

Lemma 6 *Given $X = \overrightarrow{\Omega}$ the set G_∞ is not Hausdorff.*

Proof If $x_m = (\underbrace{1, 1, ..., 1}_{m}, d, d, d...)$, then $\lim_{n\to\infty} x_n = (1, 1, 1..., 1, ...)$ and $(\underbrace{1, 1, ..., 1}_{m}, d, d, d...) \sim (d, d, d, ..., d, ...)$, for the relation R_∞. Note, however, that $(1, 1, 1..., 1, ...)$ is not in the class $(d, d, d, ..., d, ...)$ for the relation R_∞. □

Lemma 7 *Given $X = \overrightarrow{\Omega}$ denote by D the diagonal set on $X \times X$. Then, D is open on G_n for any n, where we consider on D the topology induced by $X \times X$.*

Proof Given $x \in X$, we have that $U_x(n) \times U_x(n)$ is an open set of $X \times X$ which contains (x, x).

Consider $y, z \in U_x(n)$ such that y and z are related by R_n. Then, $y_j = x_j = z_j$, when $1 \leq j \leq n$, and $y_j = z_j$, when $j > n$. Therefore, $y = z$.

From this we get that

$$U_x(n) \times U_x(n) \cap G_n \subset D$$

□

Definition 22 An equivalence relation R on a compact Hausdorff space X is said to be approximately proper if there exists an increasing sequence of proper equivalence relations R_n, $n \in \mathbb{R}$, such that $R = \cup_n R_n$, $n \in \mathbb{N}$. This in the sense that if $x \sim_R y$, then there exists an n such that $x \sim_{R_n} y$.

Example 45 Consider the equivalence relation R_∞ of Example 42 and R_n the one of Example 41. For each n the equivalence relation R^n is proper.

Then, $R_\infty = \cup_n R_n$, $n \in \mathbb{N}$ is approximately proper (see [21]).

Definition 23 Consider a fixed set K, a sequence of subsets $W_0 \subset W_1 \subset W_2 \subset ... \subset W_n \subset ... \subset K$ and a topology $\mathscr{W}_n$ for each set $W_n \subset K$.

By the direct inductive limit

$$t - \lim_{n\to\infty} \mathscr{W}_n = \mathscr{K}$$

we understand the set K endowed with the largest topology $\mathscr{K}$ turning the identity inclusions $W_n \to K$ into continuous maps.

The topology of $t - \lim_{n\to\infty} \mathscr{W}_n = \mathscr{K}$ can be easily described: it consists of all subsets $U \subset K$ whose intersection $U \cap W_n$ is in $\mathscr{W}_n$ for all n.

For more details about the inductive limit (see Sect. 2.6 in [56]).

In the case $W_n = G_n$ we consider as $\mathscr{W}_n$ the product topology.

Lemma 8 *Given $X = \overrightarrow{\Omega}$ if we consider over $K = G_\infty$ the inductive limit topology defined by the sequence of the $G_n \subset X \times X$, then, the indicator function $\mathbb{1}$ on the diagonal is continuous.*

Proof By Lemma 7 the diagonal D is an open set.

Moreover, $(G_\infty - D) \cap G_n = ((X \times X) - D) \cap G_\infty) \cap G_n = ((X \times X) - D) \cap G_n$ is open on G_n for all n. Then, $(G_\infty - D)$ is open on G_∞.

□

Remark Note that on G_∞ we have that D is not open on the induced topology by $X \times X$. Indeed, consider $a = (1, 1, 1, .., 1, ..)$ and $b_m = (\underbrace{1, 1, ..., 1}_{m-1}, d, 1, 1, 1, ..1, ...)$. Then, $\lim_{m\to\infty}(a, b_m) = (a, a) \in D$, and $(a, b_m) \in G_\infty$ but (a, b_m) is not on D, for all m.

Example 46 In the above Definition 23 consider $W_n = G_n \subset \overleftarrow{\Omega} \times \overrightarrow{\Omega}$, $n \in \mathbb{N}$, which is the groupoid associated to the equivalence relation R_n (see Example 41).

Then, $\cup_n W_n = K = G \subset \overleftarrow{\Omega} \times \overrightarrow{\Omega}$, where G is the groupoid associated to the equivalence relation R^∞. Consider on W_n the topology $\mathscr{W}_n$ induced by the product topology on $\overrightarrow{\Omega} \times \overrightarrow{\Omega}$.

For a fixed x the set $U = \{y \,|\, x_j = y_j \text{ for all } j \leq n\} \cap G_n$ is open on G_n, that is, an element on $\mathscr{W}_n$.

Note that $G_n \cap (U \times U)$ is a subset of the diagonal.

Points of the form

$$((x_1, x_2, ..., x_n, z_{n+1}, z_{n+2}, ...), (x_1, x_2, ..., x_n, z_{n+1}, z_{n+2}, ...))$$

are on this intersection.

Then, the diagonal $\{(y, y)\,,\ y \in \overrightarrow{\Omega}\}$ is an open set in the inductive limit topology $\mathscr{K}$ over G.

From this follows that the indicator function of the diagonal, that is, I_Δ, where $\Delta = \{(x, x) | \, x \in \overrightarrow{\Omega}\}$, is a continuous function.

Example 47 Consider the partition η_n, $n \in \mathbb{Z}$, over $\hat{\Omega}$ of Example 43, $W_n = G_n$, for all n, and $K = G_u$.

We consider the topology $\mathscr{W}_n$ over G_n induced by the product topology. In this way $A \in \mathscr{W}_n$ if

$$A = B \cap G_n,$$

where B is an open set on the product topology for $\hat{\Omega} \times \hat{\Omega}$.

In this way A is open on $t - \lim_{n \to -\infty} \mathscr{W}_n = \mathscr{K}$ if for all n we have that

$$A \cap X_n \in \mathscr{X}_n.$$

Denote by D the diagonal on $\hat{\Omega} \times \hat{\Omega}$ and consider the indicator function $I_D : \hat{\Omega} \times \hat{\Omega} \to \mathbb{R}$.

The function I_D is continuous over the inductive limit topology $\mathscr{K}$ over $K = G_u$.

Here G^0 will be the set $\hat{\Omega} = \overleftarrow{\Omega} \times \overrightarrow{\Omega}$. We will denote by G a general groupoid obtained by an equivalence relation R.

The measures we consider on this section are defined over the sigma-algebra generated by the inductive limit topology.

Definition 24 Given a Haar system (G, ν), where $G^0 = \hat{\Omega} = \overleftarrow{\Omega} \times \overrightarrow{\Omega}$ is equipped with the inductive limit topology, considering two continuous functions with compact support $f, g \in C_C(G)$, we define $(f \underset{\nu}{*} g) = h$ in such way that for any $(x, y) \in G$

$$(f \underset{\nu}{*} g)(x, y) = \int g(x, s)\, f(s, y)\ \nu^y(ds) = h(x, y).$$

The closure of the operators of left multiplication by elements of $C_C(G)$, $\{L_f : f \in C_C(G)\} \subseteq B(L^2(G, \nu))$, with respect to the norm topology is called the reduced C*-algebra associated to (G, ν) and denoted by $C_r^*(G, \nu)$.

Remark There is another definition of a C*-algebra associated to (G, ν) called the full C*-algebra. For a certain class of groupoid, namely the amenable groupoids, the full and reduced C*-algebras coincide. See [3] for more details.

As usual function of the form $f(x, x)$ are identified with functions $f : G^0 \to \mathbb{C}$ of the form $f(x)$.

The collection of these functions is commutative sub-algebra of the C^*-algebra $C_r^*(G, \nu)$.

We denote by $\mathbb{1}$ the indicator function of the diagonal on $G^0 \times G^0$. Then, $\mathbb{1}$ is the neutral element for the product $\underset{\nu}{*}$ operation. Note that $\mathbb{1}$ is continuous according to Example 47.

In the case there exist a neutral multiplicative element we say the C^*-Algebra is unital.

Similar properties to the von Neumann setting can also be obtained.

We can define in analogous way to Definition 13 the concept of C^*-dynamical state (which requires an unit $\mathbb{1}$) and the concept of KMS state for a continuous modular function δ.

General references on the C^*-algebra setting are [1, 17–19, 21, 22, 28, 29, 51, 53, 56].

7 Examples of Quasi-stationary Probabilities

On this section we will present several examples of measured groupoids, modular functions and the associated quasi-stationary probability (KMS probability).

Example 48 Considering the Example 2 we get that each $a \in \{1, 2, ..., d\}^{\mathbb{N}} = \overleftarrow{\Omega}$ defines a class of equivalence

$$a \times | \overrightarrow{\Omega} = a \times \{1, 2, ..., d\}^{\mathbb{N}} = (..., a_{-n}, ..., a_{-2}, a_{-1}) \times | \{1, 2, ..., d\}^{\mathbb{N}}.$$

On next theorem we will denote by G such groupoid.

Given a Haar system ν over such $G \subset \hat{\Omega} \times \hat{\Omega}$, note that if $z_1 =< a|b_1 >$ and $z_2 =< a|b_2 >$, then $\nu^{z_1} = \nu^{z_2}$. In this way it is natural do index the Haar system by ν^a, where $a \in \overleftarrow{\Omega}$. In other words, we have

$$\nu^{<a|b>}(d \ < a|\tilde{b} >) = \nu^a(d\,\tilde{b}). \tag{35}$$

Consider $V : G \to \mathbb{R}$, m a probability over $\overleftarrow{\Omega}$ and the modular function $\delta(x, y) = \frac{e^{V(x)}}{e^{V(y)}}$, where $(x, y) \in G$.

Finally we denote by $\mu_{m,\nu,V}$ the probability on $G^0 = \hat{\Omega}$, such that, for any function $g : \hat{\Omega} \to \mathbb{R}$ and $y =< a \,|\, b >$

$$\int g(y)d\mu_{m,v,V}(y) = \int_{\overleftarrow{\Omega}} (\int_{\overrightarrow{\Omega}} g(<a|b>)\, e^{V(<a|b>)}\, dv^a(d\,b))\; d\,m(da).$$

Note that $\hat{v} = e^V v$ is a G-kernel but maybe not transverse. The next theorem will provide a large class of examples of quasi-invariant probabilities for such groupoid G.

Theorem 3 *Consider a Haar System (G, v) for the groupoid of Example 48. Then, given m, V, using the notation above we get that $M = \mu_{m,v,V}$ is quasi-invariant for the modular function $\delta(x, y) = \frac{e^{V(x)}}{e^{V(y)}}$.*

Proof From (6) we just have to prove that for any $f : G \to \mathbb{R}$

$$\int\int f(x, y)e^{V(x)}v^y(dx)d\mu_{m,v,V}(dy) = \int\int f(y, x)e^{V(x)}v^y(dx)d\mu_{m,v,V}(dy). \tag{36}$$

We denote $y =< a|b >$ and $x =< \tilde{a}|\tilde{b} >$. Note that if $y \sim x$, then $a = \tilde{a}$. Note that, from (35)

$$\int (\int f\,(x, y)\, e^{V(x)} v^y(dx)\;)\; d\mu_{m,v,V}(dy) =$$

$$\int (\int f\,(<\tilde{a}|\tilde{b}>, <a|b>)\, e^{V(<\tilde{a}|\tilde{b}>)}{}_v{}^{<a|b>}(d<\tilde{a}|\tilde{b}>)\; d\mu_{m,v,V}(d<a|b>) =$$

$$\int_{\overleftarrow{\Omega}} [\int_{\overrightarrow{\Omega}} (\int\int f\,(<\tilde{a}|\tilde{b}>, <a|b>)\, e^{V(<\tilde{a}|\tilde{b}>)}{}_v{}^{<a|b>}(d<\tilde{a}|\tilde{b}>)\; e^{V(<a|b>)}\, dv^a(d\,b))\;]\, d\,m(da) =$$

$$\int_{\overleftarrow{\Omega}} [\int_{\overrightarrow{\Omega}} (\int\int f\,(<a|\tilde{b}>, <a|b>)\, e^{V(<a|\tilde{b}>)}\; e^{V(<a|b>)}{}_v{}^a(d\,\tilde{b})\, dv^a(d\,b))\;]\, d\,m(da).$$

In the above expression we can exchange the variables b and $\tilde{b}$, and, finally, as $a = \tilde{a}$, we get

$$\int_{\overleftarrow{\Omega}} [\int_{\overrightarrow{\Omega}} (\int\int f\,(<a|b>, <a|\tilde{b}>)\, e^{V(<a|b>)}\; e^{V(<a|\tilde{b}>)}{}_v{}^a(d\,b)\, dv^a(d\,\tilde{b}))\;]\, d\,m(da) =$$

$$\int_{\overleftarrow{\Omega}} [\int_{\overrightarrow{\Omega}} (\int\int f\,(<a|b>, <a|\tilde{b}>)\, e^{V(<a|\tilde{b}>)}\, dv^a(d\,\tilde{b})\;\; e^{V(<a|b>)}{}_v{}^a(d\,b)\;)\;]\, d\,m(da) =$$

$$(\int\int f\,(y,\, x)\, e^{V(x)}\; dv^y(dx)\; d\mu_{m,v,V}(dy).$$

This shows the claim. □

Example 49 Consider G associated to the equivalence relation given by the unstable manifolds for $\hat{\sigma}$ acting on $\hat{\Omega}$ (see Example 3). Let's fix for good a certain $x_0 \in \overrightarrow{\Omega}$. Note that in the case $x =< a^1|b^1 >$ and $y =< a^2|b^2 >$ are on the same unstable

manifold, then there exists an $N > 0$ such that $a_j^1 = a_j^2$, for any $j < -N$. Therefore, when $\hat{A} : \hat{\Omega} \to \mathbb{R}$ is Holder and $(x, y) \in G$ then it is well defined

$$\delta(x, y) = \Pi_{i=1}^{\infty} \frac{e^{\hat{A}(\hat{\sigma}^{-i}(x))}}{e^{\hat{A}(\hat{\sigma}^{-i}(y))}} = \Pi_{i=1}^{\infty} \frac{e^{\hat{A}(\hat{\sigma}^{-i}(<a^1|b^1>))}}{e^{\hat{A}(\hat{\sigma}^{-i}(<a^2|b^2>))}}.$$

Fix a certain $x_0 =< a^0, b^0 >$, then the above can also be written as

$$\delta(x, y) = \frac{e^{V(x)}}{e^{V(y)}} = \frac{e^{V(<a^1|b^1>)}}{e^{V(<a^2|b^2>)}},$$

where

$$e^{V(<a|b>)} = \Pi_{i=1}^{\infty} \frac{e^{\hat{A}(\hat{\sigma}^{-i}(<a|b>))}}{e^{\hat{A}(\hat{\sigma}^{-i}(<a^0|b^0>))}}.$$

Then, in this case δ is also of the form of Example 19.

In this case, given any Haar system ν and any probability m, Theorem 3 can be applied and we get examples of quasi-invariant probabilities.

The next result has a strong similarity with the reasoning of [38, 58].

Proposition 11 *Given the modular function δ of Example 20 consider the probability $M(d\,a, d\,b) = W(b)\,d\,b\,d\,a$ on $S^1 \times S^1$. Assume ν^y, $y = (a_0, b_0)$, is the Lebesgue probability db on the fiber (a_0, b), $0 \leq b < 1$, then, M satisfies for all f*

$$\int \int f(s, y)\nu^y(ds)dM(y) = \int \int f(y, s)\delta^{-1}(y, s)\,\nu^y(ds)dM(y). \tag{37}$$

Proof We consider the equivalence relation: given two points $z_1, z_2 \in S^1 \times S^1$ they are related if the first coordinate is equal.

In the case of Example 20 we take the a priori transverse function $\nu^{z_1}(d\,b) = \nu^a(d\,b)$, $z_1 = (a, \tilde{b})$, constant equal to $d\,b$ in each fiber. This corresponds to the Lebesgue probability on the fiber.

For each pair $z_1 = (a, b)$ and $z_2 = (a, s)$, and $n \geq 0$, the elements z_1^n, z_2^n, $n \in \mathbb{N}$, such that $F^n(z_1^n) = z_1 = (a, b)$ and $F^n(z_2^n) = z_2 = (a, s)$, are of the form $z_1^n = (a^n, b^n)$, $z_2^n = (a^n, s^n)$.

We define the cocycle

$$\delta(z_1, z_2) = \Pi_{j=1}^{\infty} \frac{A(z_1^n)}{A(z_2^n)}.$$

Fix a certain point $z_0 = (a, c)$ and define V by

$$V(z_1) = \Pi_{j=1}^{\infty} \frac{A(z_1^n)}{A(z_0^n)}.$$

Note that we can write

$$\delta(z_1, z_2) = \frac{V(z_1)}{V(z_2)},$$

for such function V.

Remember that by notation x_0 is a point where $(0, x_0)$ and $(x_0, 1)$ are intervals which are domains of injectivity of T.

Remark Note the important point that if $x = (a, b)$ and $x' = (a', b)$, with $x_0 \leq a \leq a'$, we get that $b_n(x) = b_n(x')$. In the same way if $0 \leq a \leq x_0$ we get that $b_n(x) = b_n$. In this way the b_n does not depends on a.

This means, there exists W such that we can write

$$\delta(z_1, z_2) = \delta^{-1}(\,(a, b), (a, s)\,) = Q(s, b) = \frac{W(s)}{W(b)},$$

where $b, s \in S^1$.

Condition (37) for y of the form $y = (a, b)$ means for any f:

$$\int \int f(\,(a, b), (a, s)\,)\ \nu^a(d\,s)\ dM(a, b) =$$

$$\int \int f(\,(a, s), (a, b)\,)\ \delta^{-1}(\,(a, b), (a, s)\,)\ \nu^a(d\,s)\ dM(a, b) =$$

$$\int \int f(\,(a, s), (a, b)\,)\ Q(s, b)\ \nu^a(d\,s)\ dM(a, b).$$

Now, considering above $f(\,(a, b), (a, s)\,)V(s)$ instead of $f(\,(a, b), (a, s)\,)$, we get the equivalent condition: for any f:

$$\int \int f(\,(a, b), (a, s)\,)\, W(s)\, d\,s\ dM(a, b) =$$

$$\int \int f(\,(a, s), (a, b)\,)\ W(s)\ d\,s\ dM(a, b).$$

As $dM = W(b) d\,b\,da$ we get the alternative condition

$$\int \int f(\,(a, b), (a, s)\,)\, W(s)\, d\,s\ \ W(b) d\,b\,d\,a\ =$$

$$\int \int f(\,(a, s), (a, b)\,)\ W(s)\ d\,s\ \ W(b) d\,b\,d\,a\,, \tag{38}$$

which is true because we can exchange the variables b and s on the first term above. □

Example 50 Consider the groupoid G associated to the equivalence relation of Example 4. In this case x and y are on the same class when there exists an $N > 0$ such that $x_j = y_j$, for any $j \geq N$. Each class has a countable number of elements.

Consider a Holder potential $A : \overrightarrow{\Omega} \to \mathbb{R}$.

For $(x, y) \in G$ it is well defined

$$\delta(x, y) = \Pi_{i=0}^{\infty} \frac{e^{A(\sigma^i(x))}}{e^{A(\sigma^i(y))}}.$$

Consider the counting Haar system ν on each class.

We say $f : G \to \mathbb{R}$ is admissible if for each class there exist a finite number of non zero elements.

The quasi-invariant condition (5) for the probability M on $\overrightarrow{\Omega}$ means: for any admissible integrable function $f : G \to \mathbb{R}$ we have

$$\sum_s \int f(s, x) dM(x) = \sum_s \int f(x, s) \, \Pi_{i=0}^{\infty} \frac{e^{A(\sigma^i(s))}}{e^{A(\sigma^i(x))}} \, dM(x). \tag{39}$$

Suppose B is such that $B = A + \log h - \log(g \circ \sigma) - c$. This expression is called a coboundary equation for A and B. Under this assumption, as $x \sim s$, we get

$$\sum_s \int f(x, s) \, \Pi_{i=0}^{\infty} \frac{e^{B(\sigma^i(x))}}{e^{B(\sigma^i(s))}} \, dM(x) =$$

$$\sum_s \int f(x, s) \, \Pi_{i=0}^{\infty} \frac{e^{A(\sigma^i(x))}}{e^{A(\sigma^i(s))}} \frac{h(x)}{h(s)} \, dM(x).$$

Take $f(s, x) = g(s, x) \, h(x)$, then, as M is quasi-invariant for A, we get that

$$\sum_s \int g(x, s) \, \Pi_{i=0}^{\infty} \frac{e^{B(\sigma^i(x))}}{e^{B(\sigma^i(s))}} \, h(x) \, dM(x) = \sum_s \int g(s, x) \, h(x) \, dM(x). \tag{40}$$

As $g(x, s)$ is a general function we get that $h(x) \, dM(x)$ is quasi-invariant for B.

Any Holder function A is coboundary to a normalized Holder potential. In this way, if we characterize the quasi-invariant probability M for any given normalized potential A, then, we will be able to determine, via the corresponding coboundary equation, the quasi-invariant probability for any Holder potential.

References [5, 8, 16, 20, 30, 48, 60, 63–67] are related to the topics discussed in the article.

References

1. Afsar, Z., Huef, A., Raeburn, I.: KMS states on C^*-algebras associated to local homeomorphisms, Internat. J. Math. **25**(8), 1450066, 28 pp (2014)
2. Anatharaman-Delaroche, C.: Ergodic Theory and Von Neumann Algebras: an Introduction. preprint Univ d'Orleans (France)
3. Anantharaman-Delaroche, C., Renault, J.: Amenable groupoids, Monographs of L'Enseignement Mathématique, 36, p. 196. L'Enseignement Mathématique, Geneva (2000)
4. Araki, H.: On the equivalence of KMS and gibbs conditions for states of quantum lattice systems. Commun. Math. Phys. **35**, 1–12 (1974)
5. Banakh, T.: Direct Limit topologies and a characterization of LF-spaces
6. Baraviera, A., Cioletti, L.M., Lopes, A.O., Mohr, J., Souza, R.R.: On the general XY model: positive and zero temperature, selection and non-selection. Rev. Math. Phys. **23**(10), 1063–1113 (2011)
7. Bissacot, R., Kimura, B.: Gibbs Measures on Multidimensional Subshifts, preprint (2016)
8. Blackadar, B.: Operator algebras. Theory of C^*-algebras and von Neumann algebras. Encyclopaedia of Mathematical Sciences, 122. Operator Algebras and Non-commutative Geometry, III. Springer-Verlag, Berlin (2006)
9. Bratteli, O., Robinson, D.: Operator Algebras and Quantum Statistical Mechanics I and II. Springer
10. Cioletti, L., Denker, M., Lopes, A.O., Stadlbauer, M.: Spectral properties of the ruelle operator for product type potentials on shift spaces. J. London Math. Soc. **95**(2), 684–704 (2017)
11. Cioletti, L., Lopes, A.O.: Interactions, specifications, DLR probabilities and the Ruelle operator in the one-dimensional lattice, Discrete and Cont. Dyn. Syst. Series A **37**(12), 6139–6152 (2017)
12. Connes, A.: Sur la Theorie non commutative de l'integration, Lect. Notes in Math. 725, Seminaire sur les Algebres d'Operateurs, Editor P. de la Harpe, pp. 19–143 (1979)
13. Connes, A.: Noncommutative Geometry. Academic Press (1994)
14. Deaconu, V.: Groupoids associated with endomorphisms. Trans. Amer. Math. Soc. **347**(5), 1779–1786 (1995)
15. DellAntonio, G.: Lectures on the Mathematics of Quantum Mechanics II, Atlantis Press (2016)
16. Diximier, J.: Von Neumann Algebras, North Holland Publishing (1981)
17. Exel, R.: A new look at the crossed-product of a C^*-algebra by an endomorphism. Ergod. Theo. Dyn. Syst. **23**(6), 1733–1750 (2003)
18. Exel, R.: Crossed-products by finite index endomorphisms and KMS states. J. Funct. Anal. **199**(1), 153–188 (2003)
19. Exel, R.: KMS states for generalized gauge actions on Cuntz-Krieger algebras. Bol. Soc. Brasil. Mat, An application of the Ruelle-Perron-Frobenius Theorem) (2004)
20. Exel, R.: Inverse semigroups and combinatorial C^*-algebras, Bull. Braz. Math. Soc. (N.S.), **39**, 191–313 (2008)
21. Exel, R., Lopes, A.: C^*-algebras, approximately proper equivalence relations and thermodynamic formalism. Ergod. Theo. Dyn. Syst. **24**, 1051–1082 (2004)
22. Exel, R., Lopes, A.: C^*- Algebras and thermodynamic formalism. Sao Paulo J. Math. Sci. **2**(1), 285–307 (2008)
23. Feldman, J., Moore, C.: Ergodic equivalence relations, cohomologies, von Neumann algebras I. TAMS **234**, 289–359 (1977)
24. Feldman, J., Moore, C.: Ergodic equivalence relations, cohomologies, von Neumann algebras II. TAMS **234** (1977)
25. Hahn, P.: The regular representations of measure groupoids. Trans. Amer. Math. Soc. **242**, 35–72 (1978)
26. Haydn, N.T.A., Ruelle, D.: Equivalence of Gibbs and Equilibrium states for homeomorphisms satisfying expansiveness and specification. Comm. Math. Phys. **148**, 155–167 (1992)
27. Kastler, D.: On Connes' noncommutative integration theory. Comm. Math. Phys. **85**, 99–120 (1982)

28. Kessebohmer, M., Stadlbauer, M., Stratmann, B.: Lyapunov spectra for KMS states on Cuntz-Krieger algebras. Math. Z. **256**(4), 871–893 (2007)
29. Kumjian, A., Renault, J.: KMS states on C^*-Algebras associated to expansive maps. Proc. AMS **134**(7), 2067–2078 (2006)
30. Kishimoto, A., Kumjian, A.: Simple stably projectionless C^*-algebras arising as crossed products. Canad. J. Math. **48**(5), 980–996 (1996)
31. Landsman, N.: Lecture Notes on C^*-Algebras and Quantum Mechanics, Univ. of Amsterdam (1998)
32. Ledrappier, F., Young, L.-S.: The metric entropy of diffeomorphisms: Part I: characterization of measures satisfying Pesin's entropy formula. Ann. Math. **122**(3), 509–539 (1985)
33. Lopes, A.O., Mantovani, G.: The KMS Condition for the homoclinic equivalence relation and Gibbs probabilities. Sao Paulo J. Math. Sci. **13**(1), 248–282 (2019)
34. Lopes, A.O., Oliveira, E.: Continuous groupoids on the symbolic space, quasi-invariant probabilities for Haar systems and the Haar-Ruelle operator. Bull Braz. Math. Soc. **50**(3), 663–683 (2019)
35. Lopes, A.O., Mengue, J.K.: Thermodynamic Formalism for Haar systems in Noncommutative Integration: transverse functions and entropy of transverse measures, to appear in Ergod. Theo. Dyn. Syst. **41**, 1835–1863 (2021)
36. Lopes, A., Mengue, J.K., Mohr, J., Souza, R.R.: Entropy and variational principle for one-dimensional Lattice systems with a general a-priori probability: positive and zero temperature. Erg. Theory Dyn Syst. **35**(6), 1925–1961 (2015)
37. Lopes, A., Mengue, J.: On information gain, Kullback-Leibler divergence, entropy production and the involution kernel, arXiv (2020)
38. Mantovani, G.: Teoria não comutativa de integração e dinâmica hiperbólica. Dissertação de Mestrado, ICMC-USP - São Carlos - Brasil (2013)
39. Matsumoto, K.: Continuous orbit equivalence of topological Markov shifts and KMS states on Cuntz–Krieger algebras, arXiv (2019)
40. Muir, S., Urbanski, M.: Thermodynamic formalism for a modified shift map. Stochastics Dyn. **14**(02), 1350020 (2014)
41. Miller, B.: The existence of measures of a given cocycle. I. Atomless, ergodic α-finite measures. Ergodic Theory Dyn. Syst. **28**(5), 1599–1613 (2008)
42. Miller, B.: The existence of measures of a given cocycle. II. Probability measures. Ergodic Theory Dynam. Syst. **28**(5), 1615–1633 (2008)
43. Meyerovitch, T.: Gibbs and equilibrium measures for some families of subshifts. Ergod. Th. Dynam. Syst. **33**, 934–953 (2013)
44. Mohr, J.: Product type potential on the XY model: selection of maximizing probability and a large deviation principle, arXiv (2019)
45. O'Uchi, M.: Measurable Groupoids And Associated Von Neumann Algebras. Ehime University Notes (1984)
46. Panangaden, J.: Energy Full Counting Statistics in Return-to-Equilibrium Jane Panangaden, thesis McGill Univ (2016)
47. Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. Astérisque, 187–188 (1990)
48. Pollicott, M., Yuri, M.: Dynamical Systems and Ergodic Theory. Cambrige Press (1998)
49. Pedersen, G.K.: C^*-Algebras and and their Automorphism Groups. Acad Press (1979)
50. Putnam, I.: Lecture Notes on Smale Spaces. University of Victoria - Canada (2015)
51. Putnam, I., Spielberg, J.: The Structure of C^*-Algebras associated with hyperbolic dynamical systems. J. Funct. Anal. **163**(2), 279–299 (1999)
52. Putnam, I.: Lecture Notes on C^*-Algebras (2019)
53. Renault, J.: A Groupoid approach to C^*-algebras. Lecture Notes in Mathematics, vol. 793. Springer (1980)
54. Renault, J.: AF equivalence relations and their cocycles, Operator algebras and mathematical physics (Constanca, 2001), 365–377. Theta, Bucharest (2003)

55. Renault, J.: The Radon-Nikodym problem for appoximately proper equivalence relations. Ergodic Theory Dyn. Syst. **25**(5), 1643–1672 (2005)
56. Renault, J.: C^*-Algebras and Dynamical Systems, XXVII Coloquio Bras. de Matematica—IMPA (2009)
57. Ruelle, D.: Noncommutative algebras for hyperbolic diffeomorphisms. Inv. Math. **93**, 1–13 (1988)
58. Segert, J.: Hyperbolic Dynamical Systems and the Noncommutative Theory of Connes. Princeton University, Department of Physics (1987). Ph.D. Thesis
59. Smith, L.: Chaos: A Very Short Introduction. Cambridge University Press
60. Takesaki, M.: Tomita's Theory of Modular Hilbert Algebras and its Applications. Lecture Notes in Mathematics, vol. 128. Springer (1970)
61. Thomsen, K.: The homoclinic and heteroclinic C^*-algebras of a generalized one-dimensional solenoid. Ergodic Theory Dyn. Syst. **30**(1), 263–308 (2010)
62. Thomsen, K.: C^*-algebras of homoclinic and heteroclinic structure in expansive dynamics. Mem. Amer. Math. Soc. **206**(970) (2010)
63. Thomsen, K.: On the C^*-algebra of a locally injective surjection and its KMS states. Comm. Math. Phys. **302**(2), 403–423 (2011)
64. Thomsen, K.: Phase transitions on O_2. Comm. Math. Phys. **349**(2), 481–492 (2017)
65. Thomsen, K.: KMS-states and conformal measures. Comm. Math. Phys. **316**, 615–640 (2012)
66. Thomsen, K.: KMS weights on groupoid and graph C^*-algebras. J. Func. Anal. **266**, 2959–2998 (2014)
67. Thomsen, K.: KMS weights on graph C^*-algebras II. Factor types and ground states, Arxiv
68. Weinstein, A.: Groupoids: unifying internal and external symmetry. A tour through some examples. Notices Amer. Math. Soc. **43**(7), 744–752 (1996)

Mixed Compression Air-Intake Design for High-Speed Transportation

Can Çıtak, Tekin Aksu, Özgür Harputlu, and Gerhard-Wilhelm Weber

Abstract Main objective of this book chapter is to explain design procedure of a mixed compression supersonic air intake. Conceptual design of air intake, modelling and simulation of supersonic flow, design of sub-components, design point of view, the objectives, constraints, phenomenon observed during the design period are mentioned. "Optimum" supersonic air-intake configuration for high speed transport aircraft are introduced briefly. Furthermore, strategy for the air-intake design of previous projects and design methodology are mentioned. Although it seems the performance requirements is enough for the successful design; operability, endurance and low-cost of the air intake are other important design goals. The aim of achieving that integration of the air-intake configuration to the propulsion system is investigated. Employing wind tunnel tests for all design alternatives makes the design period extremely long and inefficient. Instead, a simulation method of the air intake is proposed. Comparison of the results of the simulations and wind tunnel experiments are represented.

Keywords Supersonic inlet · Computational fluid dynamics · Compressible flow · Optimization

1 Introduction

High speed transport for civilian purposes has been studied for years. Different companies at various countries have tried to achieve fuel-efficient, reliable supersonic transport. Civilian supersonic flight has not been observed since 2003. Cost of the

C. Çıtak (✉) · T. Aksu · Ö. Harputlu
Roketsan Missiles Co., Ankara, Turkey
e-mail: can.citak@roketsan.com.tr

G.-W. Weber
Faculty of Engineering Management, Department of Marketing and Economic Engineering, Poznan University of Technology, Poznań, Poland
e-mail: gerhard.weber@put.poznan.pl

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_6

program and challenges such as excessive loading at sonic region, propulsion system design make the high-speed transportation difficult to carry through. There are too much controversy in supersonic transportation in these years. Engines are highly inefficient and unreliable in both fuel consumption and security reasons. After these controversies, Tupolev Tu-144 crashed at Paris Air Show in 1973 with that accident Tupolev's supersonic jet program is retired.

Furthermore, after that disaster, Concorde Supersonic Jet is crashed cause of engine explosion in 2000. This supersonic era is retired by this accident. After these inefficient designs in terms of energy efficiency and engineering, field is developed by time, which can be investigated as thermodynamics and aerodynamics, and the supersonic jet trend become popular again with supersonic business jets. Supersonic business jet development leads to a lot of new developments. On the other hand, time is so precious in today's world. Supersonic transportation concept will shorten travelling time. Even with time this concept will widely spread in to the entire world. Several concept business jets are designed already. Some of them are Aerion 2, Sukhoi Gulfstream S-21 and Jaxa NEXST. Aerion 2 is a supersonic business jet which is designed by Aerion Corporation whose motto is "It is about time". Aerion 2 is adopted with a substantial technology called Supersonic Natural Laminar Flow which is a conjugated transonic wind tunnel test and supersonic flight test. National Aeronautics and Space Administration and Aerion Corporation are working together, whereby they demonstrated the viability of supersonic natural laminar flow and showed that when this method is applied into aircrafts it is obvious they generate great efficiency [1]. Sukhoi Gulfstream S-21 was Russian – American Association Supersonic Business Jet. Sukhoi Design joined an undertaking with Gulfstream Aerospace in order to design a supersonic business jet, but because of doubtful market demand project was delayed [2, 3]. Japan Aerospace Agency developed Next Generation Supersonic Transport Jet [4]. It explained the need of supersonic transportation as "Japan is distant from Europe and United States, if the travel is shortened economic activities would be bolstered". Jaxa emphasizes that high technical potential of Japan aircraft technology will be presented with NEXST. Additionally, view of Aerion 2 [1], Sukhoi Gulfstream S-21 [2] and Jaxa Nexst [4] are shown in Fig. 1.

In order to design, these types of aircraft, propulsion system has substantial importance of aircraft performance. Especially, at supersonic speeds air-intake system design is key component of propulsion system. For years, many researches have been conducted with both experimental and computational. Design of air-intake system has many crucial problems such as inlet buzz, instabilities, air-intake system performance and mass capture and unstart of the intake system. Moreover, the cowl lip, bleed system design and diffuser angle is also key components of air-intake systems. Design of air-intake system is coupled process with computational and experimental efforts. Optimization types and methods are key elements of this process. Slater [5], developed computational tool, which named as SUPIN, for supersonic inlet design. SUPIN uses empirical, analytical and numerical methods for design and extends its fidelity by using a CFD software. CFD method Computational domain that used for CFD simulations consist and air-intake system and parametric nozzle that provide

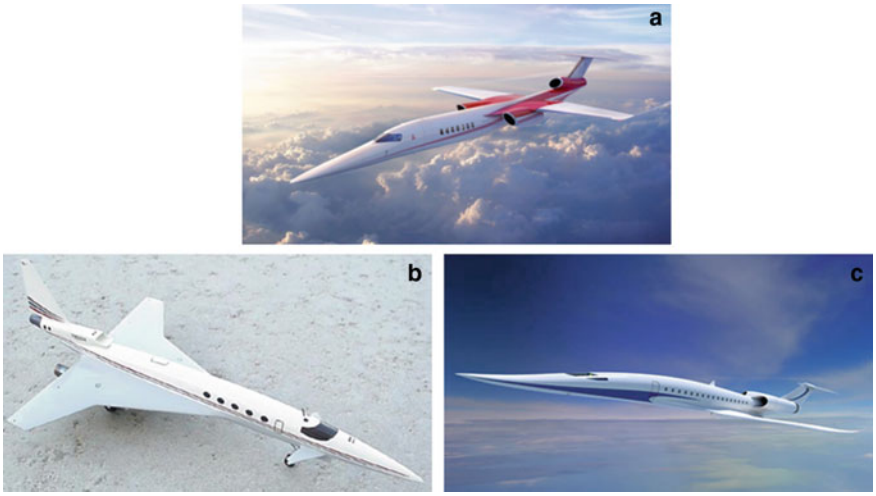


Fig. 1 Supersonic Jets **a** Aerion 2, **b** Sukhoi Gulfstream S-21, **c** Jaxa Nexst

required compressor pressure at the back of the air-intake system. From design process with and without bleed configuration air-intake systems are discussed. Air-intake system with 1% bleed, has a 4.3% better total pressure recovery but 2% lower mass ratio. Choe and Kim [6], is investigated air-intake system by using Computational Fluid Dynamics (CFD) simulations. Axisymmetric inlet configuration is used and performance curves are examined with Mach numbers 1.8 to 3.0 and with different throat ratios. Buzz and flow regimes are visualized. RANS equations is used with SST $k - \omega$ turbulence model. Air-intake system is optimized with Genetic Algorithm (GA). In addition, with and without bleed configurations are analyzed. For Mach number of 2.0 with bleed configuration TPR is enhanced by 12%. GA with CFD simulations is great way to design an air-intake system. Another substantial parameter for air-intake systems is intake unstart. Intake unstart is examined with experimental efforts. Wagner et al. [7], investigates intake unstart at Mach Number of 5 by wind tunnel experiments. Air intake back pressure is controlled by flapping nozzle. Measurements are observed with schlieren imaging with wall pressure data. Two different unstart flows, which are low and high amplitude unstarted flow, are emphasized. Pressure oscillations and frequencies are observed.

Vivek and Mittal [8], use a two-dimensional stabilized finite element method to study unsteady flows in a mixed-compression supersonic intake. Their research is focused on effectiveness of various bleed configurations in terms of starting of intake and buzz instability. Chung [9], performs computational analysis for flow through mixed compression variable diameter supersonic inlet. Time accurate Navier-Stokes solver PARC is used to simulate dynamic interactions between the inlet and the engine. Characteristics of flow with increasing back pressure is studied. Due to the CPU time and stability requirements of unsteady terms, the computations has been limited to inviscid model. Therefore, oscillations due to shock-wave boundary-layer

interactions are not investigated. Lim et al. [10], conduct a CFD analysis to investigate the effect of a three-dimensional bump installed in the supersonic inlet as an effective compression surface and a boundary-layer removal system. Their flow model includes the Reynolds-averaged Navier-Stokes (RANS) equations and Shear Stress Transport (SST) turbulence model. Moerel et al. [11], perform two and three dimensional CFD studies to generate air-intake performance data. They conduct wind tunnel tests to validate the numerical results. In two dimensional calculations bleed system is modeled as porous wall by applying static pressure boundary condition. This approach is insufficient when the normal shock passes over the bleed hole. Therefore it is not appropriate for investigating unstart characteristics. Their three-dimensional results are limited to a single point with completely closed bleed at supercritical operation. Intake performance curve and bleed effects are not investigated with three-dimensional calculations. Domel et al. [12], conduct a study on performance and stability bleed sections in mixed compression inlets, which reside on converging and diverging sections of inlet, respectively. Their work also includes the passive flow control devices as an alternative to bleed system. CFD solvers internally developed at Lockheed Martin Aeronautics Company were used examine the shock –included separation and the effect of passive devices. They perform steady RANS simulations and bleed holes were modeled as boundary conditions.

In this chapter, the system solution of the air intake for high speed transport aircraft at supersonic flow regime is explained. Design methodology and brief information about supersonic air intake is presented in the next sections. Eventually, simulation methods which take an important role during design process, are explained.

2 Design Methodology

Air-intake system is the heart of the air breathing propulsion system. It should operate between required flight envelope and given restrictions. Moreover, the amount of air and air properties that are supplied from the air-intake system is substantial. Air-intake system performs different missions at different Mach numbers and flight altitudes, and it should ensure requirements at minimum in different flight conditions but provide maximum at design point. In order to decide on design point and air-intake system performance characteristics will be explored in depth.

The aspect of design of the supersonic air intake includes both propulsion system performance and aerodynamic efficiency of the overall system. Although the effects of these two approaches seem separated, the objective function of the system must provide a coupled requirements written subsequently:

- The performance requirements inside the flight envelope, highest performance at the design point,
- Minimum drag force contributed to the overall system,
- Unstart of Air-Intake System,
- Buzz Limit,

- Required mass capture ratio,
- Solid structure during operation,
- Low cost during the production and maintenance.

Treatment of the design optimization of the air intake is not unique for all flight regimes and performance requirements. For this reason, direct optimization techniques might not give the shortest path to the "best" configuration to reach the objectives. Additionally, the physics related to supersonic air intake needs attention for the flow simulation. As a result, the design space of the air intake can be restricted by taking advantage of the literature and experiments completed before. Brief information about the definition of performance characteristics and types of supersonic air intakes, and "rule-of-thumbs" for the flow simulation techniques can be obtained and applied for the current study.

The efficiency of the propulsion system is a key parameter for the air-intake design. The drag force contribution to the overall system is an additional objective to obtain a longer range of high-speed transport aircraft. There are several solutions for the fast and long-range transportation. Supersonic air intake can be characterized by change of total pressure recovery with respect to mass capture ratio. Total pressure recovery is the ratio of the total pressure at combustion chamber to freestream total pressure:

$$Maximum\ Contraction\ Ratio\ (MCR) = \frac{A_0}{A_t},$$

$$Total\ Pressure\ Recovery\ (TPR) = \frac{P_{t4}}{P_{t0}}.$$

It is desired that a supersonic air intake which is employed for high speed transport aircraft has high total pressure recovery values with high mass capture ratio. Furthermore, low drag force due to air intake, is able to operate at various altitude and Mach number. Supersonic air intakes can be classified according to these objectives mentioned above. Three types of air intake can be talked about; namely, internal, external and mixed compression supersonic air intakes may be defined. Internal compression air intakes compress the freestream flow by series of oblique shocks inside; finally, normal shock at the inlet throat makes the flow subsonic. Typical flow structure inside the air-intake system is represented in Fig. 2.

Reflection of the flow inside the internal compression air intake is always directed in axial location. The need of re-direction of the flow vanishes which reduces the intake drag force. Since the strength of the shock waves are strong, boundary layer

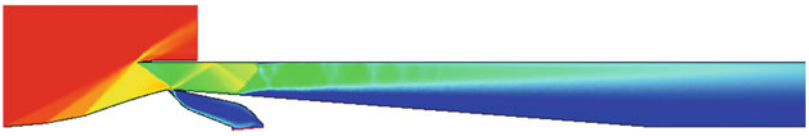


Fig. 2 Typical flow structure and shock train in air-intake system

on the internal faces of the intake can be separated easily which results in an unstart situation of air intake. In order to get subsonic flow before combustion chamber entrance, internal compression types of air intakes need longer geometries which is not an efficient solution in cases beyond Mach 2.

Compression includes one more oblique shock wave at the external part of the air intake. Aerodynamic throat (smallest section of supersonic air intake) is close to cowl lip. Equating the strength of the shock waves gives an optimum air-intake geometry for a given design point (Mach number and altitude). Increasing the number of oblique shock waves makes the air-intake efficient due to reducing total pressure loss. Efficiency of external compression air intake reduces if we are above Mach 2.5 with increasing design Mach number.

Mixed compression aims at the use of advantages of internal and external compression air intakes. In spite of seeming perfect, there are some disadvantages, too. It has higher total pressure recovery ratios at high Mach numbers above Mach 2.5. On the other hand, the ability of starting of the mixed compression air intakes is the main problem to be solved. Since the compression is separated with external and internal shock waves, the contour design of mixed compression air intake is more complex than other types of air intakes; this makes design procedure difficult. Variation of total pressure recovery with design Mach number for different types of air intakes is presented in Fig. 3.

Type of air-intake system is considered according to the flight envelope and flight regime. In addition, performance characteristics of air-intake systems are investigated by performance curves. A typical performance curve is shown in Fig. 4.

Air intakes can be differentiated by their geometric shapes as axial-symmetric, two-dimensional and three-dimensional air intakes. Axial–symmetric air intakes can be employed one at central or two at the sides of high-speed aircraft. A typical

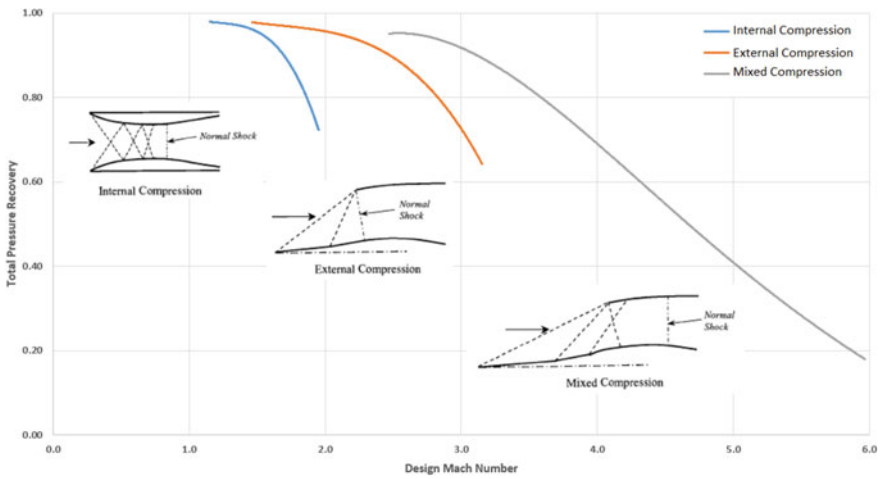


Fig. 3 Pressure recovery for different types of intakes

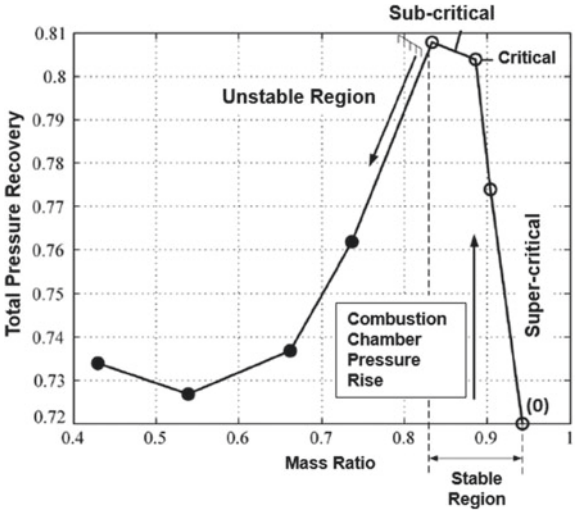


Fig. 4 Regions of performance curve for air-intake system

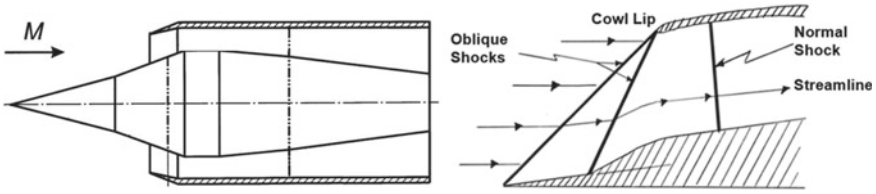


Fig. 5 Axial-symmetric and two-dimensional air-intake configurations

representation of these air-intake systems is represented in Fig. 5. Robustness and "easy-to-design" are the advantages of them. Obtaining the highest performance at zero angle of attack is the main disadvantage of axial-symmetric type of air intake. Two-dimensional air intake is employed for many high speed aircrafts. Concorde and Tu-144 are shown in Figs. 6 and 7.

Improving performance with increasing angle of attack makes easier to climb after take-off. Three dimensional air intakes are employed at the sides of the nacelle without using diverter generally. Diverterless supersonic air intakes controls the boundary layer by using aircraft surface. Locations of the air intake(s) on the high speed aircraft are decided by considering the effect on propulsion system, drag force contributed to overall system and operability.

Fig. 6 Concorde and air-intake representation [13]

Fig. 7 Tupolev Tu-144 and air-intake representation [14]

2.1 Design Parameters

Reaching the most suitable supersonic air-intake configuration for the high-speed transport aircraft at reasonable time period needs an efficient design process. In that time period, the air intake is designed by the parameters.

Design parameters:

- Number of ramps,
- Ramp lengths,
- The length of air intake and diffusor,
- Bleeding system type,
- Compression type,
- Air-intake geometry,
- Cowl-lip angle.

Investigation of performance characteristics of an air-intake inside the flight envelope makes the wind tunnel tests compulsory. However, wind tunnel tests are quite expensive. Thereby, investigation of all the design parameters for the flight envelope by wind tunnel tests is not the best option to design a procedure. Computational Fluid Dynamics (CFD) tools are used for the simulation of the supersonic flow related to supersonic air-intake design. Information about validation of CFD tool with similar studies and experimental results, settings for numerical schemes are given in the next sections.

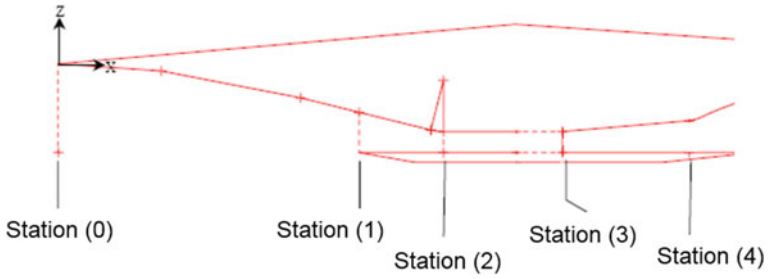


Fig. 8 Air-intake system stations

2.1.1 Air-Intake Stations

An air-intake system is considered with notation included as in Fig. 8.

- Station 0: Free flow region,
- Station 1: Start of internal compression region,
- Station 2: Aerodynamic throat,
- Station 3: End of aerodynamic throat,
- Station 4: Combustion chamber entrance.

Air intake meets the supersonic flow at Station 0, starts to decelerate and compress until Station 2. Station 2 which is also characterized as "aerodynamic throat", defines the sub-sections of the performance curves geometrically. Normal shock stays at Station 2 on the "critical point". Compressed and decelerated flow to subsonic speed passes through Station 4 before entering the combustion chamber.

2.1.2 Design Criteria and Air-Intake System Characteristics

In order to assess air-intake system configurations some definitions are shown in Fig. 9. The air-intake system mass capture ratio and contraction ratios are the impor-

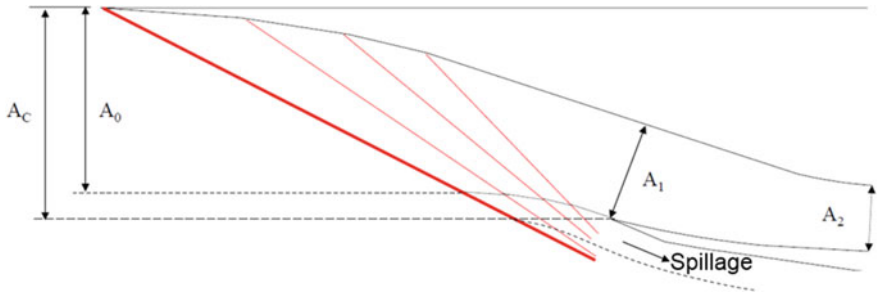


Fig. 9 Sectional area definitions of the air-intake

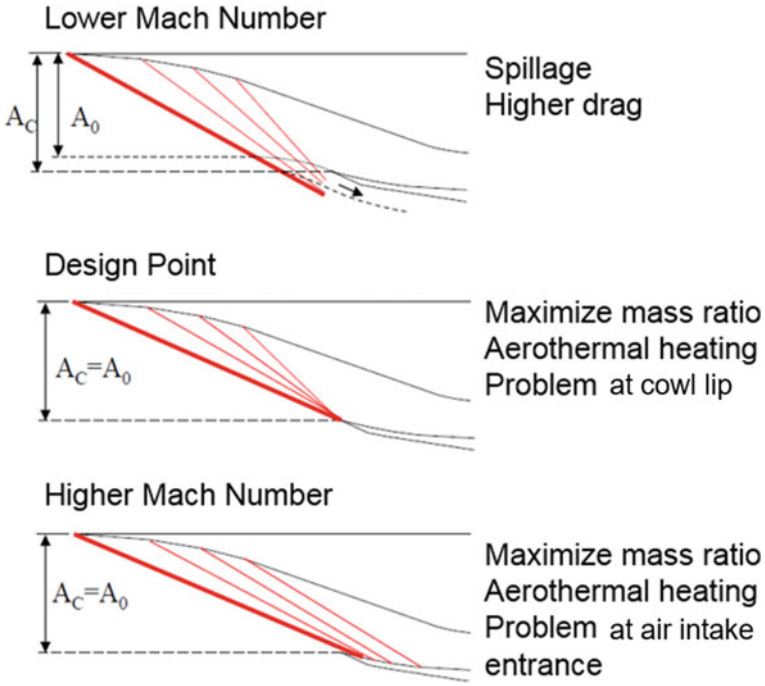


Fig. 10 Aerodynamic structure of the shock wave at different mach numbers

tant parameters. A_0 is the captured area, A_c is the maximum area that can be captured by system. The aerodynamic throat is implied as A_2, the cowl lip area is also denoted by A_1.

The air-intake system can capture any area in different flow regimes. From low supersonic to high supersonic shock-wave characteristics are shifting through the aerodynamic throat. For fixed geometry air-intake system, the captured area is changing by freestream Mach number. On design point shockwaves are unite at cowl lip but off design points for lower Mach number spillage occurs. For higher Mach numbers shock waves proceed into the air intake which leads to high thermal loading. A typical figure of shock waves for different operating points is shown in Fig. 10.

Aerodynamic throat ratio (A_0/A_t) is very important parameter for air-intake system. This parameter varies with freestream Mach number. This ratio affects both general performance characteristics of the system and self-start condition. Mahoney [16] implies self-start throat variables but that can be extended by using an efficient bleed system. This study compromises both analytical and experimental work.

Fixed compression surface air-intake systems are designed according to on-design condition. Therefore, off-design points' performance are maximized and subsidiary to the on-design performance. Soltani and Daliri [15] emphasize an air-intake system that has an design point for Mach number of 2 at Fig. 11. When performance curves are discussed, if the Mach number increased from 2.0 to 2.2, then the total pressure

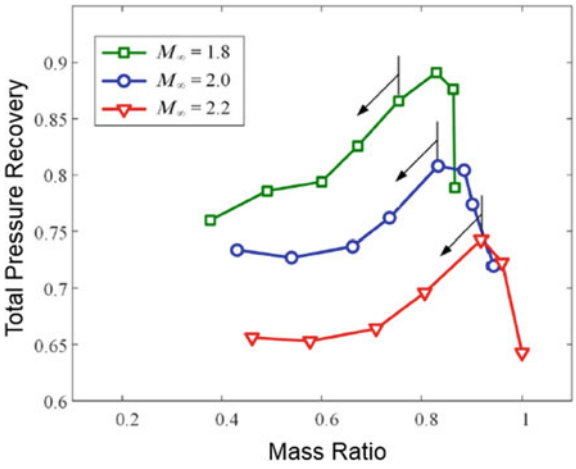


Fig. 11 Performance curves of the air-intake configuration [6]

recovery factor is diminished by ten percent. In theory, the total pressure recovery factor is decreased by increasing the Mach number. A characteristic performance curve for different Mach numbers is shown in Fig. 11.

Aerodynamic throat should be determined according to a design point that has minimum Mach number to be able to self-start. However, by choosing a minimum Mach number as design point, total pressure recovery factor cannot reach the maximum total pressure recovery factor in dynamic compression surface of the air-intake systems.

Computational Fluid Dynamics are viable tool for design of air-intake systems. In terms of CPU time first one dimensional shock relations can be solved for initial geometry. Shock relations are isentropic and, therefore, no viscous effects are considered. Ramp angle, altitude, geometric dimensions, deflection angle, cowl lip angle, bleed system, shock boundary-layer interactions should be taken into account for air-intake systems. These parameters can optimized with CFD calculations.

2.1.3 External Compression Shock Wave Characteristics

The general aim of the air-intake system is to maximize the efficiency while capturing and decelerating shock waves by external compression surface. Compression surface is decelerating shockwaves by using compression ramps. Typical compression of external surface is illustrated in Fig. 12.

Ramp number and ramp angle change according to the configurations and the flight envelope. While ramp number is increasing, compression maximizes efficiency and total pressure recovery. However, geometric and production restrictions restrain the number of ramps. One of the important design parameters is transition between ramps. The limit is decelerating the flow speed by 25% in order to avoid flow separations [16].

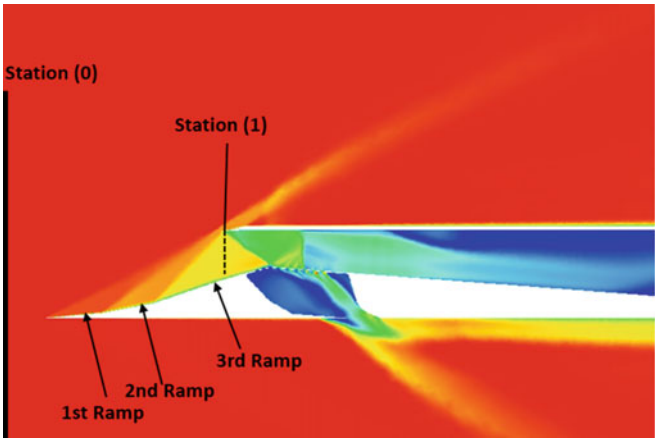


Fig. 12 External compression surface for air-intake system

2.2 Computational Fluid Dynamics Simulations

Commercial CFD tool solves Reynolds Averaged Navier Stokes (RANS) equations. Transient and steady-state approaches are enabled depending on the focus of the simulation. RANS with $k-\omega$ turbulence model equations are solved implicitly. These equations are represented below. Continuity, momentum and energy equations are shown as Eqs. (1), (2) and (3). In Eq. (1), density is implied with ρ, subscript t is time and x is space. In addition, velocity is emphasized with u. For our momentum equation p is pressure, $\overline{\tau_{ij}}$ is viscous strain tensor:

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial (\bar{\rho}\widetilde{u_j})}{\partial x_i} = 0, \tag{1}$$

$$\frac{\partial (\bar{\rho}\widetilde{u_i})}{\partial t} + \frac{\partial}{\partial x_j}\left(\bar{\rho}\widetilde{u_i}\widetilde{u_j}\right) = \frac{\partial \bar{\rho}}{\partial t} + \frac{\partial}{\partial x_j}\left(\overline{\tau_{ij}} + \bar{\rho}\overline{u_i'' u_j''}\right), \tag{2}$$

$$\frac{\partial (\bar{\rho}\widetilde{H})}{\partial t} + \frac{\partial}{\partial x_j}\left(\bar{\rho}\widetilde{u_j}\widetilde{H}\right) = \frac{\partial}{\partial x_j}\left(\bar{\rho}\alpha\frac{\partial \widetilde{H}}{\partial x_j} + \bar{\rho}\overline{u_j'' H''}\right), \tag{3}$$

$$\alpha = \frac{k}{\bar{\rho} C_p}, \tag{4}$$

$$\bar{\rho}\overline{u_i'' u_j''} = u_t\left(\frac{\partial \widetilde{u_i}}{\partial x_j} + \frac{\partial \widetilde{u_j}}{\partial x_i}\right) - \frac{2}{3}u_t\frac{\partial \widetilde{u_k}}{\partial x_k}\delta_{ij} - \frac{2}{3}\bar{\rho}\delta_{ij}, \tag{5}$$

$$\bar{\rho}\overline{u_j'' H''} = \frac{u_t}{Pr_t}\frac{\partial \widetilde{H}}{\partial x_j}. \tag{6}$$

In Eq. (3), H means total enthalpy and α is thermal diffusivity. In Eq. (4), k is the thermal conductivity and c_p is specific heat. In order to solve Favre averaged equations, turbulence related $\bar{\rho}\overline{u_i'' u_j''}$ and $\bar{\rho}\overline{u_j'' H''}$ are going to be solved with additional models. In Eq. (2) $\bar{\rho}\overline{u_i'' u_j''}$ term is Reynolds stress tensor, u_j'' and u_i'' is the deviation from the averaged velocity in any direction. These terms are solved by Boussinesq approach. According to that approach turbulence related terms are solved with Eqs. 5 and (6). In Eq. (5), μ is dynamic viscosity and δ_{ij} is the Kronecker delta.

2.2.1 Standard $k-\omega$ Turbulence Model

Standard $k-\omega$ turbulence model is found by Wilcox [17]. Wilcox simplify equations that were introduced by Kolmogorov, and turbulent viscosity terms are calculated from Eq. (7):

$$\mu_t = \frac{\bar{\rho}k}{\overset{\infty}{\omega}}, \tag{7}$$

$$\overset{\infty}{\omega} = \max\left[\omega, 0.875\left(\frac{2\overline{S_{ij}}\,\overline{S_{ij}}}{\beta^*}\right)^{0.5}\right], \tag{8}$$

$$\overline{S_{ij}} = S_{ij} - \frac{1}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}. \tag{9}$$

In these equations, ω is the turbulent dissipation rate, $\overline{S_{ij}}$ is the strain rate and β^* a model coefficient. Turbulence kinetic energy, k, and turbulence dissipation rate, ω, is calculated from Eqs. (10) and (11):

$$\frac{\partial(\bar{\rho}k)}{\partial t} + \frac{\partial\left(\bar{\rho}\,\overset{\infty}{u_j}k\right)}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\left(\mu + \sigma^*\mu_t\right)\frac{\partial k}{\partial x_j}\right) + \bar{\rho}\tau_{ij}\frac{\partial u_i}{\partial x_j} - \bar{\rho}\beta^*\omega k, \tag{10}$$

$$\frac{\partial(\bar{\rho}\omega)}{\partial t} + \frac{\partial\left(\bar{\rho}\,\overset{\infty}{u_j}\omega\right)}{\partial x_j} = \alpha\frac{\omega}{k}\bar{\rho}\tau_{ij}\frac{\partial u_i}{\partial x_j} + \frac{\partial}{\partial x_j}\left((\mu + \sigma\mu_t)\frac{\partial\omega}{\partial x_j}\right) + \sigma_d\frac{\bar{\rho}}{\omega}\frac{\partial k}{\partial x_j}\frac{\partial\omega}{\partial x_j} - \bar{\rho}\beta\omega^2. \tag{11}$$

Specific Reynolds stress tensor, τ_{ij}, is implied as in the following Eq. (12).

$$\tau_{ij} = 2\frac{\mu_t}{\bar{\rho}\overline{S_{ij}}} - \frac{2}{3}\delta_{ij} \tag{12}$$

Table 1 Standard $k-\omega$ model experimental coefficients

α	β	β^*	σ^*	σ_d	σ
0.52	0.07	0.09	0.6	0.125	0.5

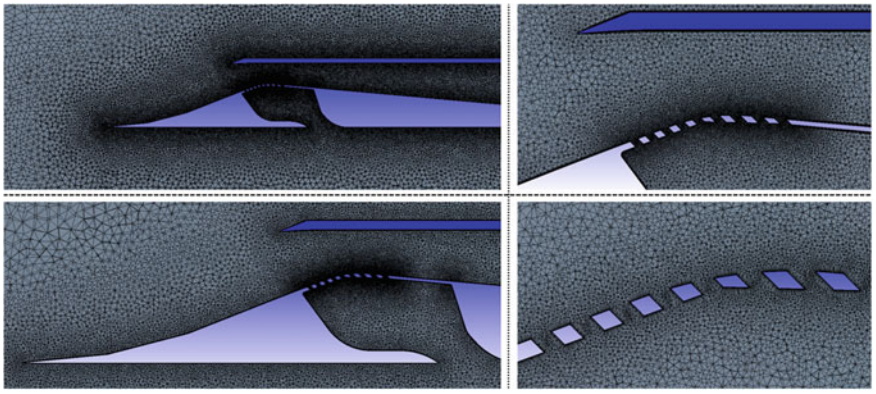


Fig. 13 Section view of mesh used in CFD simulations

Experimental coefficients for standard $k-\omega$ turbulence model are shown in Table 1.

Density-based Roe flux splitting scheme is employed. Maximum Courant number is 8, and no-slip boundary condition is applied for all air-intake walls. The unstructured grid was used for both 2- and 3-dimensional domains.

Computational Fluid Dynamics (CFD) solver is employed to simulate flow parameters inside the flight envelope of the air intake. Since the flight envelope consists of the interval of Mach number and freestream pressure, the flight path of the air intake is partitioned by considering the design points. The main objective of these simulations is to calculate performance characteristics of supersonic air intake for all discretized design points. A computational grid representation of an air-intake system is shown in Fig. 13. The average cell number for simulations is 10 million tetrahedral cells with structured grid at near wall for resolving the boundary layer.

Calculation of the air-intake performance is conducted with CFD simulations by changing backpressure (pressure at combustion chamber entrance). Changing backpressure mechanism can affect the numerical computation of the performance curves. The terminal shock moves from the combustion chamber entrance to the cowl lip by increasing back pressure. Air-intake system components are shown in Fig. 22. It is observed that speed of the movement is affected by the pressure gradient near walls. Here, the performance of air intake is computed less than if the pressure increment is modelled by boundary directly. For this reason, a parametric nozzle is used for the back pressure increment. By doing this, flow is choked by the help of the "virtual" nozzle. To simulate the different back pressure, the nozzle throat opening is modified. Herewith, the internal part of air intake is pressurized naturally.

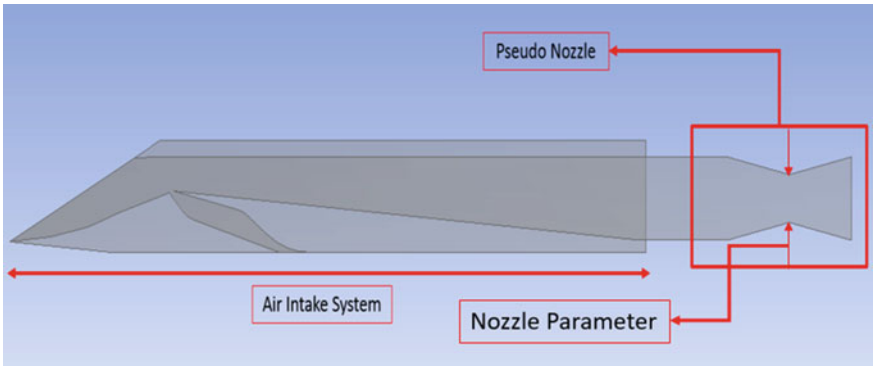


Fig. 14 Pseudo nozzle for the CFD simulations and nozzle parameter for back pressure adjustment

2.3 Performance Assessment of Air-Intake System

For our research, a performance map for the air-intake system is a must. Performance maps are investigated by using a pseudo nozzle at back of the air-intake system. The nozzle throat is changed by parametric study and the combustion chamber pressure is simulated. Testing each configuration of an air-intake system is very expensive design process.

That is why CFD simulations is a viable tool for the performance map. The computational domain for performance assessment is shown in Fig. 14. The nozzle throat is diminished and the performance curve is obtained.

2.4 Validation of CFD Simulations and Numerical Setup

Wind-tunnel tests are conducted for the same air-intake model. The aim of the process is the compare the results of the CFD simulations; thereby, we validate the simulation results. Wind-tunnel tests are completed for various flight conditions which belong to flight zone. The performance curves obtained from tests, and simulations for the same flight conditions are compared. As a result, CFD simulations take more place from wind-tunnel tests which brings out an efficient, fast and inexpensive design process.

A 3-dimensional unstructured mesh is employed for the air-intake simulations. The points inside the flight zone are chosen to simulate flight conditions at the relevant discrete time zone. Flight conditions are simulated at the wind-tunnel tests by adjusting Mach number, freestream pressure and temperature. The results are compared to the change of total pressure ratio with respect to mass capture ratio, called performance curve. It is the most important parameter to define the performance

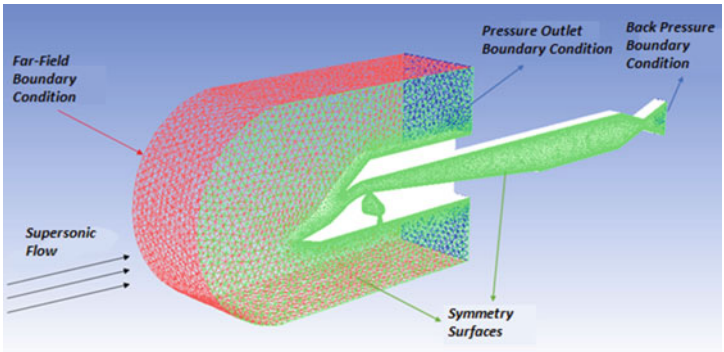


Fig. 15 Boundary condition of the air-intake simulation model

characteristics. In detail, flow at supersonic speed is compressed and reaches the combustion chamber after passing through the air intake. Mission of the air intake is to decelerate the flow at high efficiency in terms of total pressure. In other words, the efficiency and ability to handle the total pressure at various flight zones of the supersonic air intake can be defined by the performance curves. "Safe zone" for the air-intake operability according to performance curves defines limitations of the overall propulsion system.

"CFD simulation method", generated by the air-intake design team, is employed for the validation test cases. Performance curve is obtained by changing throat opening of the virtual throat. Herewith, the pressure at the end of the air intake is modified without disturbing the boundary layer additionally. Figure 15 shows the boundary condition of the simulation model. Flow parameters for the far-field and the pressure outlet boundaries are obtained from outstanding air breathing propulsion system modelling results.

Four different wind-tunnel tests are completed. The flow parameters of the tests are represented in Table 2.

Performance curves for various Mach numbers are simulated with discrete virtual nozzle throat openings. In this way, the pressure at the end of the air intake is controlled for supercritical, subcritical and critical zones [18]. Mach contour of the symmetry plane of the air intake for the second test is shown in Fig. 16.

Table 2 Flow parameters employed for CFD simulations

Test number	Mach number[a]	$T_{t0}[K]$	$P_{\infty}[bar]$	$T_{\infty}[K]$
1	0,71	292	0,22	130
2	0,86	290	0,16	104
3	1.00	288	0,12	83

[a] Flow Mach numbers are represented non-dimensionally due to confidentiality considerations

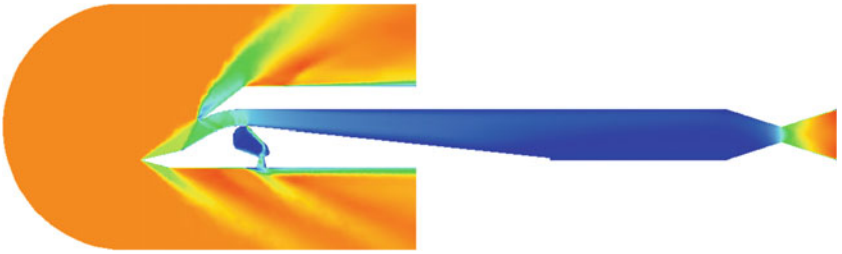


Fig. 16 Mach contour of the air-intake

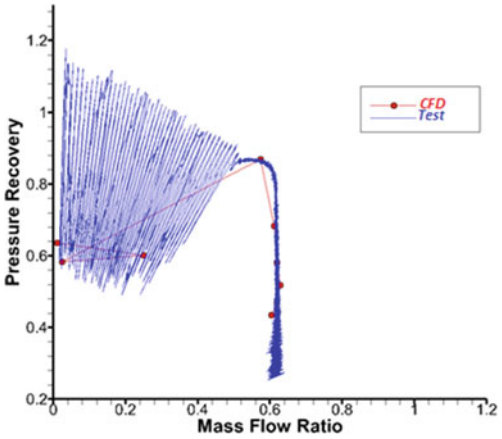


Fig. 17 Comparison of the performance curves—first test

Performance curves obtained from the wind-tunnel tests and CFD simulations are compared below. As expected, the overall performance of the air intake reduces with increasing Mach number. Additionally, the stable region margin, which air intake can operate safely, changes with the freestream flow parameters.

Operability zone of the air intake is limited by considering that unstable region (high frequency variation of the flow parameters). Despite the limitation, test results from unstable region are used for the validation of CFD simulations. Figure 17 shows our comparison of the performance curves related to the first test.

A comparison for the second test is presented in Fig. 18. Flow passing through the air intake is compressed by the oblique shock waves. Since these shock waves hit the boundary layer, disturbed boundary layer may cause separation. The result of the boundary-layer distortion decreases the effective aerodynamic throat size. It leads the lower performance than the capacity of the flow compression. The resultant of the shock-wave boundary-layer interaction can be seen at the performance curves which provide reliable comparison parameters.

Performance curves for the last test are shown in Fig. 19. The maximum of the total pressure recovery decreases with increasing Mach number, as expected. The capability of the compression by oblique shock waves is lower at faster freestream

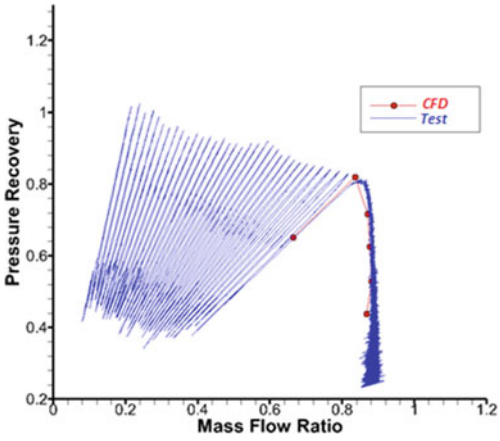


Fig. 18 Comparison of the performance curves—second test

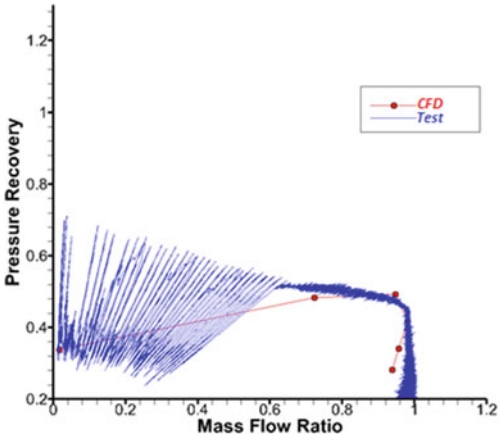


Fig. 19 Comparison of the performance curves—third test

flow speed. In addition, "distance" between the design point and the actual point inside the flight zone is a strong parameter. The performance of the air intake rises as the distance shortens. Flow parameters of the third condition are the furthest to design point, which has the lowest performance.

Schlieren pictures also help to validate the CFD simulation method by visualizing the shock waves. Numerically computed and experimental Schlieren contours are compared for the same back pressure values. A comparison of Schlieren contours for the first test is represented in Fig. 20.

Location and the shape of the shock waves, which has a darker color on the contours, are quite similar. The shape of the external compression waves on the experiment is sharper than the numerical Schlieren contour. Numerical approach, mesh quality and convergence criteria are effective with regard to the shape of the shock waves. However, the total outcome of these can be realized on the performance

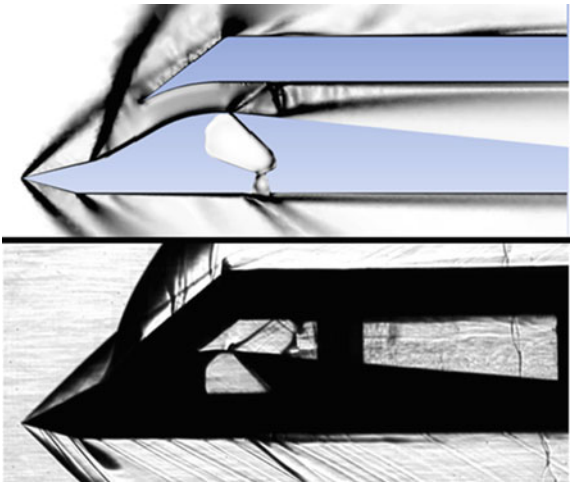


Fig. 20 Schlieren contours of first test (Numerical—above, Experimental—below)

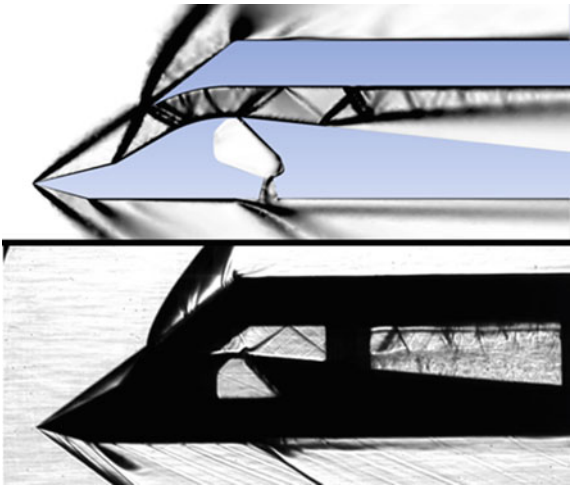


Fig. 21 Schlieren contours of second test (Numerical—above, Experimental—below)

curves implicitly. Figure 21 represents the Schlieren contours of the second test at the same back pressure. The shock wave trains for both approaches are similar.

To sum up, results of the experimental and CFD simulations is quite similar. CFD simulations are able to be employed to obtain characteristics of the air intake. It leads to a reduction in number of wind-tunnel tests. Although CFD solver does not simulate flight, it results in highly accurate results on supersonic air-intake simulations.

3 Component Design

3.1 Cowl Lip

The unstart phenomenon is one of the main issues encountered in mixed compression air intakes that seriously affect the properties and mass flux of the flow delivered to the combustion chamber. Thickening of the boundary layer inside the intake due to shock boundary-layer interaction causes separation. The separation process drives flow oscillations and expulsion of shocks. When the shock system is expelled outside of the intake unstart, process is observed. Unstart leads decrease of total pressure and mass flow entering the combustion chamber which results inefficient propulsion system cycle. Due to the shock-wave boundary-layer interaction flow separation occurs. Separated flow decreases the effective aerodynamic throat. Compression efficiency of the air intake decreases. Shock waves can blow out and air intake might unstart [19]. Cowl lip is one of the sub-components of the air intake. The mission of which is to reflect the shock waves coming from external compression ramps. An air intake is represented in Fig. 22.

Cowl-lip angle governs the strength of the reflected shock wave and reflection angle. According to cowl-lip angle, reflected shock-wave gradient might change on the lower surface. Since it has an effect on the shock train system and gradient

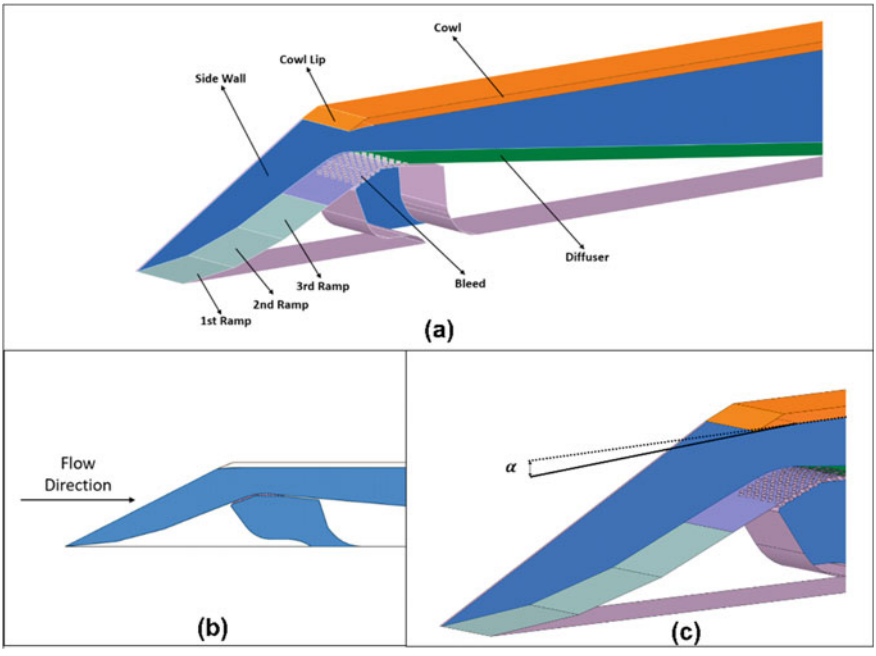


Fig. 22 **a** Subcomponents of air intake; **b** Side view of the air intake; **c** Cowl-lip angle

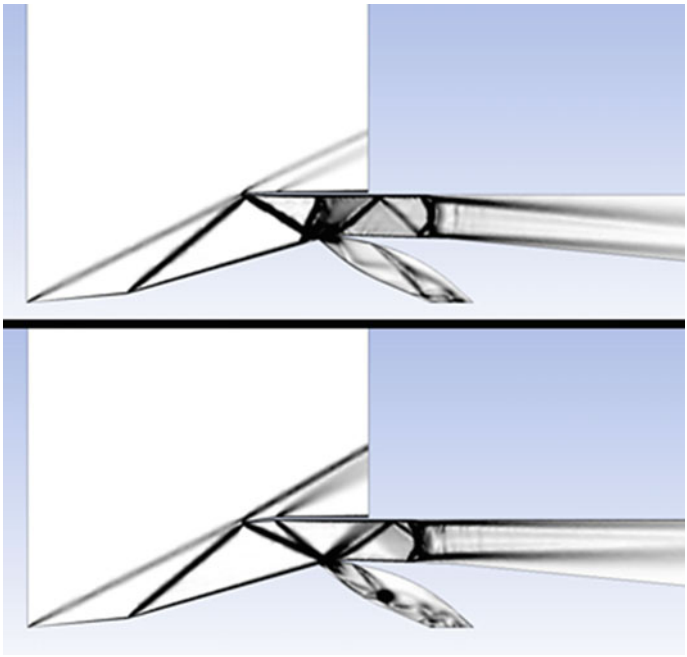


Fig. 23 Numerical schlieren contours on the symmetry plane (top-0°, bottom-6° cowl-lip angle models)

location, it can be chosen to have optimum (maximum) performance values. Initially, CFD simulations are completed for 0° and 6° cowl-lip angle models. It is seen that 6° cowl-lip angle model has a higher total pressure recovery limit. Figure 23 represents the numerical Schlieren contours of two different cowl-lip angle models.

Cowl-lip angle provides a reduction of the reflected shock wave strength. Since the shock strength is reduced, the flow has higher a Mach number than the zero cowl-lip angle model. Two approaches can be followed with cow-lip angle: narrowing down the throat area and/or increasing external compression ramp angle(s). Since the shock strength is reduced by the cowl-lip angle, more compression can be done until throat-area location. On the other hand, the compression level can be kept constant by increasing ramp angles. This process enables us to shorten the air intake. Increasing the total pressure limit is not the only viewpoint on the design. A high mass-capture ratio is desired too [20]. It is realized that 6° cowl-lip angle model has lower mass capture ratio than expected. Air-to-fuel ratio is not satisfied for the combustion process. The 6° cowl-lip angle model is investigated both for total pressure recovery and mass capture ratio. Mach contours for 0° and 3°cowl-lip angle models are represented in Fig. 24.

Maximizing total pressure recovery or mass capture ratio only is not the design strategy. At the similar view; the 3° cowl-lip angle model is superior in a multi-objective optimization perspective [21]. Numbers of ramps, ramp angles and length

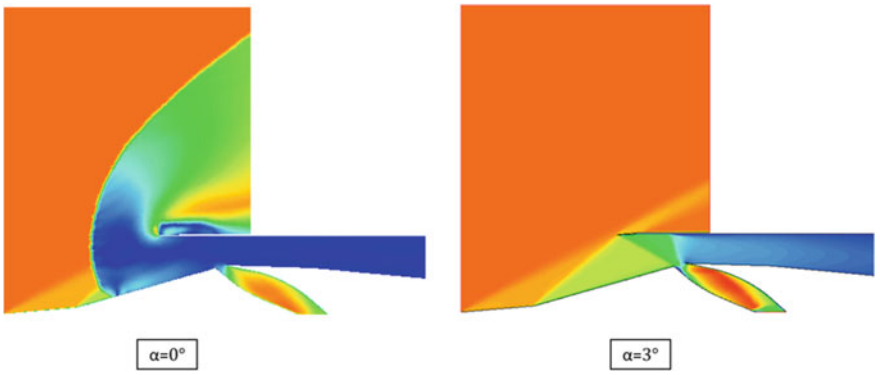


Fig. 24 Mach contours of 0° (left) and 3° (right) cowl-lip angle models

are the other parameters which directly affects the compression system. The design process starts with the isentropic shock wave relations. Conceptual design of the air intake can be finalized by using optimization of angle and length of the ramps. However, isentropic shock-wave relations do not take into account viscous forces [22]. In other words, flow separation due to shock-wave boundary-layer interaction cannot be observed. Therefore, effective aerodynamic throat-area computation is impossible. CFD simulation is needed for this purpose. The nonlinear relationship between design parameters and objectives makes the design process difficult and challenging [23].

3.2 Bleeding System

Bleeding systems are used to thin the boundary layer, which is thickened due to shock-wave boundary-layer interactions. The system provides the size of the aerodynamic throat keep constant. Otherwise, aerodynamic throat could be shrunk due to shock waves which leads a lower performance than expected. Despite seeming mass flow has been thrown away instead of being used for the combustion, it is proven that 1% of mass flow ejected out by the bleeding system enhances 10% total pressure recovery, which is a better choice for the overall propulsion system [24]. Two types of the bleeding system are investigated. Namely, perforated and slot bleeding systems. Slot bleeding system has a high effect on the design point. On the other hand, at "off-design points" it provides no improvement on elimination of shock-wave boundary-layer interaction effects. Perforated bleeding system is superior to slot one. It is able to ensure the amortization of thick boundary layer inside the air intake. Figure 25 represents the types of bleeding systems.

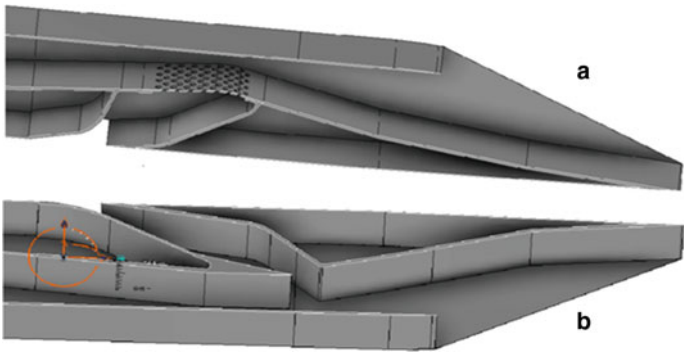


Fig. 25 Bleed Types: **a** Perforated and **b** Slot-shaped bleed system

4 Conclusion

In this study, design procedure of a mixed compression supersonic air intake is introduced. Important design requirements related to performance and operatability during mission are emphasized. Key design parameters and their effect on the intake performance are mentioned. Experimental and numerical methods used for obtaining performance characteristics are explained. Air-intake system is mainly discussed in the view of performance characteristics and performance maps. Importance of total pressure-recovery factor and mass-capture ratio is emphasized. Typical performance data for CFD and test results are shown. In addition to the discussed parameters, inlet unstart and inlet buzz are substantial topics for air-intake systems. Supersonic transportation is a very challenging subject and the air-intake systems' performance directly affects the general performance of the engine. In future studies, for increasing the fidelity of the CFD simulations, buzz and unstart characteristics will be explored in depth.

References

1. Garzon, G.A., Matisheck, J.R.: Supersonic testing of natural laminar flow on sharp leading edge airfoils. Recent Experiments by Aerion Corporation, 42nd AIAA Fluid Dynamics Conference and Exhibit, New Orleans (2012)
2. Chudoba, B., Coleman, G., Roberts, K., Mixon, B., Mixon, B., Oza, A., Czysz, P.A.: What price supersonic speed?–a design anatomy of supersonic transportation–Part 1. Aeronaut. J. **112**(1129), 141–151 (2008)
3. Chudoba, B., Coleman, G., Huang, X., Huizenga, A., Czysz, P. A., Butler, C. M.: A feasibility study of a supersonic business jet (SSBJ) based on the Learjet airframe. 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno (2006)
4. Furukawa, T., Makino, Y.: Conceptual design and aerodynamic optimization of silent supersonic aircraft at JAXA. 25th AIAA Applied Aerodynamics Conference, Miami (2007)

5. Slater, J.W.: SUPIN: A computational tool SUPIN: a computational tool. 54th AIAA Aerospace Sciences Meeting AIAA SciTech Forum, San Diego (2016)
6. Choe, C.K.Y.: Numerical investigation of bleed effects on supersonic inlet under various bleed and inlet conditions. 34th AIAA Applied Aerodynamics Conference AIAA Aviation Forum, Washington D.C. (2016)
7. Wagner, J.L., Yuceil, K.B., Valdivia, A., Clemens, N.T., Dolling, D.S.: Experimental investigation of unstart in an inlet/isolator model in mach 5 flow. AIAA J. **47**(6), 1528–1542 (2009)
8. Vivek, P., Mittal, S.: Buzz instability in a mixed-compression air intake. J. Propulsion Power **25**(3), 819–822 (2009)
9. Chung, J.: Numerical simulation of a mixed-compression supersonic inlet flow. 32nd Aerosp Sci Meeting Exhibit. Reno, NV (1994)
10. Lim, S., Koh, D.H., Kim, S.D., Song, D.J.: A computational study on the efficiency of boundary layer bleeding for the supersonic bump type inlet. 47th AIAA Aerospace Sciences Meeting Including the New Horizions Forum and Aerospace Exposition. Orlando, FL (2009)
11. Moerel, J.-L., Veraar, R.G., Halswijk, W.H.C., Pimentel, R., Corriveau, D., Hamel, N., Lesage, F., Vos, J.B.: Internal flow characteristics of a rectangular ramjet air intake. 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Denver, CO (2009)
12. Domel, N.D., Baruzzini, D., Miller, D.N.: A perspective on mixed-compression inlets and the use of CFD and flow control in the design process. 50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Nashville, TE (2012)
13. Cain, T.: Ramjet Intakes, RTO-EN-AVT-185 (2010)
14. McLean, F.E.: Supersonic Cruise Technology, NASA SP; 472 (1985)
15. Soltan, M.R., Younsi, J., Daliri, A.: Performance investigation of a supersonic air intake in the presence of the boundary layer suction. Proc. Inst. Mech. Eng. **229**(8), 1495–1509 (2014)
16. Mahoney, J.: Inlets for Supersonic Missiles, AIAA Education Series (1991)
17. Wilcox, D.C.: Turbulence Modelling for CFD, ed., California: DCW Industries Inc. (1998)
18. Park, W., Park, S.: Optimal terminal shock position under disturbances for ramjet supercritical operation. J. Propul. Power **29**(1), 238–248 (2013)
19. Das, S., Prasad, J.K.: Starting characteristics of a rectangular supersonic air-intake with cowl deflection. Aeronaut. J. **114**(1153), 177–189 (2010)
20. Williams, A.S.J.: Computational prediction of subsonic intake spillage drag. 24th Applied Aerodynamics Conference, San Francisco (2006)
21. Allison, D., Morris, C., Schetz, A.: A multidisciplinary design optimization framework for design studies of an efficient supersonic air vehicle. 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Indianapolis (2012)
22. Flock, K., Gülhan, A.: Viscous effects and truncation effects in axisymmetric Busemann scramjet intakes. AIAA J. **54**(6), 1881–1891 (2016)
23. Torp, H.: Favourable Design of the Air Intake of a Ramjet, Norwegian University of Science and Technology, Trondheim, Master Thesis (2016)
24. Meerts, C., Steelant, J.: Air intake design for the acceleration propulsion unit of the LAPCAT-MR2 hypersonic aircraft. European Space Research and Technology Centre ESTEC-ESA, Munich (2013)

Social Entrepreneurship Business Models for Handicapped People—Polish & Turkish Case Study of Sharing Public Goods by Doing Business

Dominik Czerkawski, Joanna Małecka, Gerhard-Wilhelm Weber, and Berat Kjamili

Abstract Entrepreneurship is a complex issue, having an interdisciplinary character, which definitely constitutes a set of personality traits expressed in the entrepreneur's attitude. This background is the basis for reflection and an to take an attempt to present the definition of entrepreneurship, which multitude in the literature review exists as the scientific subject. Hence, the aspect of conducted research is individual conditions and direct interviews, allowing to describe the social-economic environment in which current and potential business owners grow. Almost 300 years have passed since the publication of R. Cantillon's theory lead economy to the creation of works in the field of social entrepreneurship. The entrepreneurship manifestations described by specific examples—coming from the commercial market—made it possible to put forward the thesis that all symptoms are a consequence of achieved profits, or are a derivative of the "road to success", whose main motive is the desire to make a profit. The issue of entrepreneurship, however, takes on a different connotation when one additional feature of the entrepreneur is added under consideration: his physical disability. Then, social commitment, passion, openness and empathy take on new connotations, from which once again one feature that determines the whole process cannot be selected, only their team. It should be agreed, however, that these processes will seriously imply innovations, which are the "key" to more effective, lasting and with consequences solutions. The projects were carried out separately in Poland and Turkey. Social entrepreneurship described in accordance with the *Nshareplatform* (NSP) was created for people with physical disabilities in order to create better living conditions and integration with society as active individuals in a dynamically changing economic environment. A support system has been created

D. Czerkawski (✉) · J. Małecka · G.-W. Weber
Faculty of Engineering Management, Poznan University of Technology, Poznan, Poland
e-mail: dominik.sla.czerkawski@doctorate.put.poznan.pl

J. Małecka
e-mail: joanna.malecka@put.poznan.pl

G.-W. Weber
Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey

B. Kjamili
Department of Economics, Middle East Technical University, Ankara, Turkey

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_7

that enables both networking and reaping benefits for able people who can support disabled people by achieving their goals. The Turkish experience was based on the *Liberated Social Entrepreneur* and *Migport* project, which meet the needs of many migrants in Turkey, improving their living conditions, activating socially needed activity, which in turn affects social peace in a country that is burdened with many risks—including political, belonging to the range of basic determinants limiting the economic development of this country.

Keywords Social entrepreneurship · Social impact · Business modeling · Handicapped people · Public goods

1 Introduction

One of the 17th century philosopher—Pascal, stated: "*Movement is life, immobility is death*" [5]. Movement is both the natural need of every human being and the means of expression. Because of the movement, the expression of ourselves could exist and conversion of human feelings or emotions. In addition, the movement is the basis of human activity manifesting itself even in the activities of everyday life, while performing professional, artistic and sport activities. It is therefore extremely important that this movement is the quintessence of human being lives [10].

Let's use playing tennis as an example of research problem concerning disabled people which will be considered. Tennis is one of the most popular sports in the world. It is gaining more and more popularity among people with visual impairments [8]. However, for these people to enjoy the game of tennis, be am member of tennis clubs, should be adapted to their capabilities and needs in terms of both technical and organizational aspects [15, 37]. In Poland, only a few clubs allow blind players to play tennis. Thanks to the recent victories of Polish sightless players at international tournaments, more and more people from different cities want to play this sport. Adaptation of sports facilities to the needs of people with visual disabilities is an extremely important aspect in the context of integration of this social group. Playing sports by people with disabilities has a positive effect on their physical and mental condition. People with mobility impairments are especially susceptible to isolation and social exclusion. Quality of their lives depends on accessibility of means of public transport which is one of the factors affecting their possibility of having active life, learning and working. Increasing accessibility of trains, infrastructure and transport services cause this kind of isolation to fade. Big Data Analytics tools can be useful for analysis current situation and create a friendly space to living and social activity (see also: [15, 37]).

Blind tennis was created in Japan, in Kawagoe in 1984 by Miyoshi Takei. He was a blind student who wanted to play tennis and together with a physical education teacher he tried to "adapt" to this sport [53]. After several years of work on the ball and the rules of the game, more and more people began to play tennis. The first tournament for the blind was held in October 1990 in Japan [3].

A person with disabilities, due to their individual difficulties, needs additional help in many situations to be able to function actively in society. Unfortunately, due to the numerous mental and technical barriers, these people often spend most of their time at home and do not integrate at all or less (in relation to people with disabilities). Polish law understands participation in social life as an opportunity to fulfill social roles and overcome barriers—in particular, ones of psychological or architectural kind, ones in urban planning, transport and communication [3]. Therefore, people with disabilities often need social rehabilitation, which includes making personal resourcefulness and stimulating social activity of a disabled person, developing the ability to fulfill social roles independently, eliminating barriers and shaping the right attitudes and behaviors in a society that are conducive to integration with people with disabilities. One of the important points that must be meet is the adaptation of public transport, including rail transport, to the needs of people with disabilities, who should use it in a manner least dependent on third parties [40].

In the presence of a still growing number of public goods, adapted for people with movement and/or eye dysfunction, it can still be observed a rather small percentage of people who use it [20, 40]. This group of people has got a special need. Mostly they need a help in doing a daily activity, and they have got this—the regulation from the government is taking care of it [40]. But they also have got a *dream*—like being *a part of a social community*.

There are many organizations helping disabilities person, because the problem is global [47], e.g.,

(1) student's organization, local organization
(2) world organizations
(3) sport association
(4) volunteers
(5) dedicated people, who take money for this job.

However: does their help have got a social impact? Maybe they do just a mechanical work? There are one of the questions on witch authors try to find answers by leading investigation between handicapped people in Poland and Turkey. The main goal was—as a result, of research questionnaire and individual interviews—create a new method which could provide to the answer on whether to start a social business in order to introduce sustainable impacts in developing countries, that would be sustainability solution with possibilities to implementation in few services as a small business idea, in term of entrepreneurial attitude. How pick up the activity of people with disabilities, including the dysfunction of movement and/or eye in social life? As a secondary issue: how to make better of communication: a non-disabled person—a person with a disability? The authors try to find, create and test—in two countries: Poland—as a member of UE and Turkey—outside from UE—models of leading profitable businesses, according to the rules of social economy.

2 Entrepreneurship as a Part of Social Economy

Entrepreneurship, which is the basis of discourse in the literature on the subject, goes back to the considerations of Aristotle and Xenophon. The essay by R. Cantillon—now called the father of entrepreneurship—published in 1755—has led to many considerations, dividing those who treat the issue of entrepreneurship as part of the field of scientific management and those who believe that it created economics [11, 32]. One cannot ignore the classic approach, in which, in contemporary considerations of A. Smith's successors, he appears as the fourth factor next to land, capital and labor [32, 48, 49].

The aspect of social systems that indisputably have a direct impact on the development of all micro and macroeconomic indicators that cause the development of world economies was considered by T. Parsonas and N.J. Smelers, whose praxeological approach and the importance of internationalization processes form the basis of research assumptions—in the field of each economic activity whose social and cultural processes are supported by sociological and psychological conditions [41].

Undoubtedly, the issues of morality itself cannot be ignored in the considerations regarding economics, management or entrepreneurship. Because quick material benefits and only profit-oriented approach do not fully define *modern homo oeconomicus* [16, 32, 48, 49]. Undoubtedly, in the review of economic literature, there are numerous not only connotations, but even synonymous aspects of entrepreneurship and activities "for and by" micro, small and medium-sized enterprises (SMEs) [32, 48, 49, 51–53], (see also [30, 34, 46]). This economic activity of individual units means the basis for the socio-economic development of all economies of the world, in which M. Weber's approach in terms of value-rational and ideological is becoming increasingly important [27, 54, 55]. This aspect probably plays an important role in the analyzed countries: Turkey and Poland, whose diversity of denominations and practiced religions, by definition conditions the pluralism of the approach and considerations on conducting economic activity—not only in the local, but also regional and global scope (see also [29]).

The aspect under consideration, however, implies the issue of entrepreneurship with the personality traits of the entrepreneur himself. Hence the indication that entrepreneurship is an interdisciplinary issue, taking into account the achievements of many sciences, of which the leading should be social sciences: psychology and sociology, next to economics and law, taking into account local institutional and legislative conditions as well as the ability to manage and run finances, which should be considered in aspect of competence of the managing person—entrepreneur [32].

These elements can be found in the theories of JM Keyns, J.B. Say or D. Ricard, in which money circulation creates the theory of absolute income. Entrepreneurship is one of the eight main motives for saving, so it is not a separate issue in itself, but is an attitude, represented by a specific person with specific traits and attitudes, elements of which can also be found in the works of Kirzner and Knith [21–24].

Entrepreneurship will therefore be defined as a set of personal traits that are expressed through the entrepreneur's attitude, readiness to set up, lead and set a direction for the development of own enterprise and as further research in this field is conducted, it will probably be evaluated [32].

On the other hand, the issue of social entrepreneurship—apart from non-profit organizations, which should be covered by a separate study—can be considered in five key dimensions: (I) social mission, (II) social innovation, (III) social change, (IV) entrepreneurial spirit and (V) personality [7, 42].

The social mission (I) should exert an impact on the whole community, it should introduce changes that are favorable both locally and globally. It often addresses health, educational, economic, political and cultural problems associated with chronic poverty [1, 2]. Among the issues of social entrepreneurship, disability occupies the fourth (4) most discussed issue contained in the mission, just before (1) aging, (2) addiction to chemical substances, (3) children with special needs, and before: (5) discrimination against minorities, (6) education, (7) lack of access to information and communication technologies, (8) energy production and distribution, (9) environment, (10) health, (11) homelessness, (12) concern for peace and conflict resolution, (13) poverty, (14) creating sanitary conditions in rural communities, (15) street children, (16) renewable energy, (17) trafficking in women and children, (18) unemployment, (10) equality for women rights [42].

Innovations (II)—in this aspect—should be not only effective, but also lasting and having consequences—an impact that on the principle of "domino effect" will affect the attitudes of the whole society. It's a new way of doing things, "new ideas that have a raison d'etre", highlighting the difference between real innovation and improvement [38]. Social innovation helps achieve social goals.

The consequence of social innovation is social change (III), which has far—reaching consequences, often accompanied by the side effects of change—which may be more significant than the change itself [44]. Citing Kramer [28]: "social entrepreneurs set themselves a goal that non-profit organizations don't even dream about—they want not only to serve the local community or create a national network, but also to introduce lasting changes in the behavior of the entire nation or even the whole world so that raise the standard of living for millions of people"[1] [28].

Whereas the entrepreneurial spirit (IV) is determined by the individual characteristics of the entrepreneur, to which the authors attach the most important, interdisciplinary character by creating their own definition of entrepreneurship. Just creativity or leadership qualities as well as management skills and the so-called resourcefulness in life will not be sufficient determinants of entrepreneurship, because only perseverance in the pursuit of exerting influence on the entire field in various social and geographical aspects will testify to the completeness of the action. The personality (V) is a separate, fifth dimension of social entrepreneurship that includes creativity—an innovative approach that goes beyond existing frameworks and existing conventions. A detailed division was made by the Ashoka organization (International Organization of Social Innovators "Ashoka Innovators for the Public Good,) according to which,

[1] Own translation from Polish language.

out of 10 million entrepreneurs, only 1 meets all criteria [39]. Hence the functional approach to management, which is not typical for SMEs, and the creation of mission, vision and values can also be the key to efficiency and success (see also: [45]).

The authors have attempted to characterize the business models created in Poland and Turkey, based on social entrepreneurship, addressed to people with disabilities who are often excluded and, although they remain fully mentally fit, their physical disability prevents them from full social activity.

3 Entrepreneurial Barriers

When systematizing the enterprise market, the size of the company is usually taken into consideration—quantitative criteria regarding the average number of employees, the amount of income and the value of fixed assets. The most dynamically developing companies include enterprises from the SME sector, which at the same time constitute 99.8% of all companies in the EU-28 [32, 33, 36]. Undoubtedly, motivation and commitment occurring in both employees and employers are conducive to the development of a learning organization in a continuous process, and the popularization of knowledge management is conducive to the investments made.

However, development barriers exist and significantly affect and limit entrepreneurship. Distractions in this regard can be carried out in terms of both the enterprise and entrepreneur. Hence the discourse about competences, perceiving the opportunities and market opportunities of the managing person. The five most important determinants mentioned in the relevant literature review include:

(1) financial
(2) legal
(3) administrative and bureaucratic
(4) social
(5) the gray area
(6) price competitiveness
(7) low economic potential
(8) concentration of development on the local market and low export capacity
(9) small innovative and innovative character of the products
(10) majority of micro enterprises
(11) low level of cooperation between enterprises
(12) lack of experience of entrepreneurs
(13) low level of knowledge in management and marketing
(14) no strategy, focus on ongoing operations
(15) low investment
(16) failure to recognize the intellectual potential of employees,

which are visible in both countries where the research was conducted: Poland and Turkey.

Due to the research aspect, the most important will be the social barrier—increasingly appearing as very important, especially in the knowledge-based economy in which the human factor is a priority [33]. Generally, social barriers are obstacles or limitations of any kind that reduce access to specific social groups. Most often, these barriers appear towards people with disabilities or those who are environmentally excluded who cannot be active at the expected level. In post-transformed societies, there is also a limited acceptance of self-employment as a social barrier. However, according to the Industry 4.0 concept, this barrier can also be considered in terms of implementation: (1) resistance to change, (2) users' approach to the "new", (3) fear of responsibility, (4) lack of motivation, (5) lack of integrated thinking .

Change is an inseparable part of development, and yet, it is still associated pejoratively and evokes a sense of danger. The "new" is usually associated with the elimination of the "old", therefore there is a natural tendency to reject any restrictions incomprehensible or requiring the acquisition of new qualifications and competences. The old methods are proven methods that show efficiency at a known level. The new one requires physical and intellectual effort, at the cost of which employee teams are reluctant. The difficulties that arise are usually suppressed so that the errors that occur are not a basis for exercising personal responsibility, and the attitude towards new issues is rather particular, without taking into account the wider perspective of the entire enterprise.

Due to the approach to conducted research, attention should also be paid to the financial barrier, which can be considered in the aspect of investing and implementing innovations that are necessary to survive in a turbulent and competitive global commercial market, but also in terms of obtaining sources for start—to start a business. One of the proposed models is the development of an enterprise based on start-up financing, which belongs to the offer of private equity solutions. They finance the activities of private enterprises that are not listed on the public market (stock exchange) and encounter limited opportunities to raise capital for development from the money market (e.g. bank credit, encountering the phenomenon of so-called credit discrimination) [17, 18, 36]. Private equity (PE) funds make their offer dependent not so much on the size of the enterprise as on its age, adjusting the forms and structures of individual instruments (so-called early-stage funds, later stage venture found and balanced found). The younger the enterprise, the more of early stage venture could use (seed or start-up), even if the stage it is planning phase. The design of products (no interest, commissions, collateral or pledge) contribute to the economic development and economic growth of entire regions of the world [35]. It is worth emphasizing that the funds, providing financial resources for the development of other enterprises, at the same time create new jobs and constitute an important forge of innovation, often belonging—in the aspects of national business registration—to the group of small and medium-sized enterprises (Fig. 1).

Although private equity has several possible forms of financing, capital raised through it is often referred to as risk capital. The name—as mentioned earlier—thus refers to the translation of the phrase venture capital understood as a form of investing in high-risk projects and at the same time the possibility of collecting high profit, assuming success in business. The fund's life cycle usually consists of four stages:

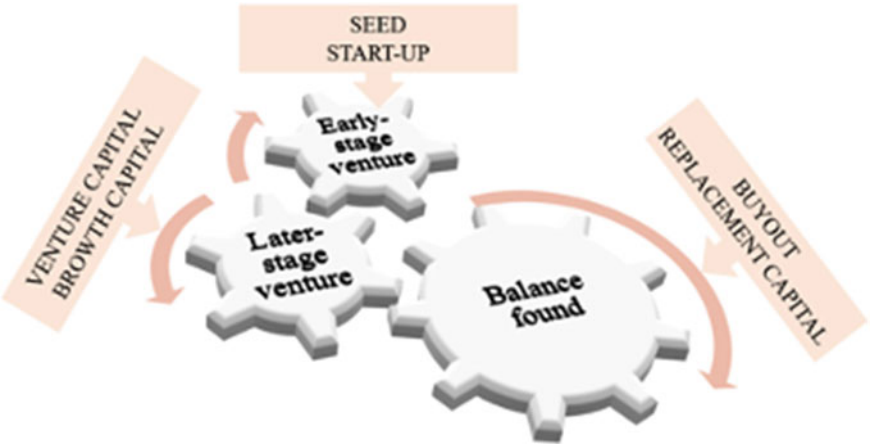


Fig. 1 Private equity financing by company development phase

(1) equivalent actions to find interesting investment opportunities and mobilize capital
(2) commencement of the fund's activity, at the stage of which the investment objective is selected by analysing business plans, measuring the investment sample, undertaking negotiation activities and gathering all information of interest to investors
(3) building an investment portfolio consisting of the value of individual portfolio companies—determining the relationship between risk and rate of return, consisting in minimizing the risk measure (so-called variance) at the assumed rate of return or maximizing the value of reimbursement (rate of return) at a given risk value (variance)
(4) the "exit from investment" stage—usually after 5–7 years in the form of listing companies, resale, merger, decapitalization, sale of non-core assets, and in the future a secondary buy-out.

Private equity funds are created in a form dedicated to individual markets. However, their activity in Central and Eastern Europe still does not match the capital commitment in the rest of the continent (Fig. 2).

Poland is often called the country with the most intensively developing capital market, because it has a developed system of financial instruments, as well as the most dynamically developing stock exchange in this region—the Warsaw Stock Exchange (WSE). The Istanbul Istanbul Stock Exchange (BIST) has a much higher capitalization, however, the political risk in this country causes a lot of upsetting, which makes the Turkish market less attractive also for the PE solutions (WSE—449 companies with capitalization 1 121 880. 78 million PLN—https://www.gpw.pl/statystyka-gpw data for 27.12.2019; data enabling comparison: WSE: 432 companies with capitalization USD 125 billion versus BIST- 300 companies with a total

Fig. 2 Fundraising geographic breakdown of venture capital [% of total amount]. *Source* Małecka, 2016 [35]

capitalization of approx. USD 200 billion—https://www.bankier.pl/consciousosc/Turecka-gielda-tylko-dla-inwestorow-o-mocnych-nerwach-7441545.html data for 22.07.2016, see also [29].

4 Social Entrepreneurship Business Models

The literature review shows few approaches of investigations to reach the aim and create benefits with working with analyzing tools, as data mining, and analytics and artificial intelligence: CMARS, RMARS and RCMARS (Table 1) [4].

As a process of creating new models to understand the problem, specially rising money for success rate of the startups way, two models were created: at Polish and Turkish background. First of them, the *Nshareplatform*, was created by handicapped member of this organization, who is also a President of this association since 2011. The value of described method is based on the direct interviews with people engaged in this organization. Similar situation concern the *Migport Program*, but in this case model has own results as a leading business, with measurable effects and noticeable impact on refugees in Turkey.

4.1 NSHAREPLATFORM as a Business Model

The analysis from the interview with fifty disable people from Poland from Wielkopolskie region shows that persons who work for public as a support for them

Table 1 Benefits with working with Big Data Analytics tools (*source* [4])

Level: Grade	Gap	Market $	Competitors	Network	Entrepreneur
10	Most painful killing headache	Over trillions	Potential strong competitors	Existing top clients	Dedicated; convincing; finds a way; most key factor; never give up
9	Strong serious headache	Trillions	Strong serious competitors	Network to be introduced top clients	Dedicated; convincing; finds a way; most key factor
8	Serious headache	Billions	Serious competitors	Membership with top clients	Dedicated; convincing; finds a way
7	Weekly headache	A few Billions	Up to 10 Competitors	Knows a few top clients	Dedicated; convincing
6	Daily headache	Millions	A few competitors	Knows a few middle people to reach top clients	Dedicated
5	Early headache	A few Millions	Up to 3 strong competitors	Have some clients	Finds a gap but doesn't pivot
4	Symptom	Hundred thousand's	Up to 1 Strong competitor	Meeting with a few clients	Tries to find a gap
3	Early symptoms	A few Hundred Thousand's	Up to 3 not very strong competitor	Submit pivots to people met	More than only will intends to start
2	Symptoms examination	$10.000–$100.000	Up to 1 not very strong competitor	Meet with people and follow up	Have a will to be an entrepreneur
1	Thinks have a headache	$1000–$10.000	Ordinary a few competitors	Meet with people	Have a will but doesn't evet start
0	No pain	<$1000	No Competitors	No network	Not entrepreneur
The level	10	10	10	10	10

more often doesn't know the real needs of people with disabilities. This staff is trained by public sector for general cases and needs. More often, the assistant (this is the formal name for such persons, works in public sector) doesn't even have similar interests as the person to whom the help is addressed. This makes the support provided is not sincere and amounts to mechanical work. This way of thinking gave opportunity to create the main problem of global culture in four points:

(1) many ways of helping are not useful
(2) ways of helping aren't focused of their social needs
(3) there must be a social impact of helping
(4) need to personalize a guardian, who is a special person helps disabilities (see also [6]).

At this way the *Nshareplatform* (NSP) was created—connect the ideas of both sides and create system to help with benefit. It will be creating a friendly public space for people with disabilities. In case of still growing number of public goods, adapted for people with movement and/or eye dysfunction. The reason may be, that people with disabilities will need for an assistant—a special person, who help them. Based on observation and experience, authors suspects that the fear of people with disabilities in using the space dominated by non-disabled people may be the reason. Non-disabled people who will be a part of the NSP project they provide will have the opportunity to learn the real needs of people with disabilities depending on the situation. They will know how to help effectively. NSP responds to the main problem: how pick up the activity of people with disabilities, including the dysfunction of movement and/or eye in social life? As a secondary issue: how to make better of communication: a non-disabled person—a person with a disability?

The NSP assumptions create a system in which minimum two people are needed—one of those needs to have a disabilities card, and both want to make use of public goods, like: swimming pools, theaters, cinemas, etc. A person with disabilities is entitled to use in public transport for free and half precious for other public goods. In both cases, a person who does not have a disabilities card—named: *guardian of a disabled person* (GoDP)—can make use with a special price of public goods. The GoDP does not need to know a disabled person before the meeting and activity. The Polish law regulations are allowing for this, so the inclination for all European Union countries it will be the same. This will be good promotion for NSP project to the global.

The function of an assistant to a disabled person will be the main task for NSP project. The innovativeness of this project's idea is to personalize the assistant's function under the expectations of people with disabilities. Both, the person—one with disability and the guardian choose in terms of their own interests, passions or a momentary desire to spend time for less money. The idea of NSP project will enable easy scalability to the territory of Poland and is a business model to become implemented. It is addressed for people who want to use public goods for "less money fitness", it is and will be part of this communities and help in the convenience of their lives, and for "flowing more and more" with a wider stream of society. Moreover, it is possible to observe greater empathy and understanding for them in society, which

Fig. 3 The social entrepreneurship business model of the *Nshareplatform* (NSP) project

will allow for a long wave of implementation of the idea. Thanks to the innovative approach to the topic of disability, the target audience of recipients can be extended to people with various *disabilities*, physical but also societal ones, and the assistant can be limited only by a slight fear, which will be compensated by an outstanding mutual benefit.

The creation of the NSP model was based on interviews with people who have been disable since birth, people who in authors opinion know the best how to connect these two worlds—disabilities and not-disabilities persons. There was first step and basic rule to create innovative the *Nshareplatform* application, where one of the management function: motivating people, is basic rule to create impact—act in own interests (Fig. 3), (see also [12]).

One of such ideas had has occur in 2009-http://www.nieprzecietni.put.poznan.pl/index.php?zm=kimJestesmy data for 29.12.2019, as being originator and co-founder

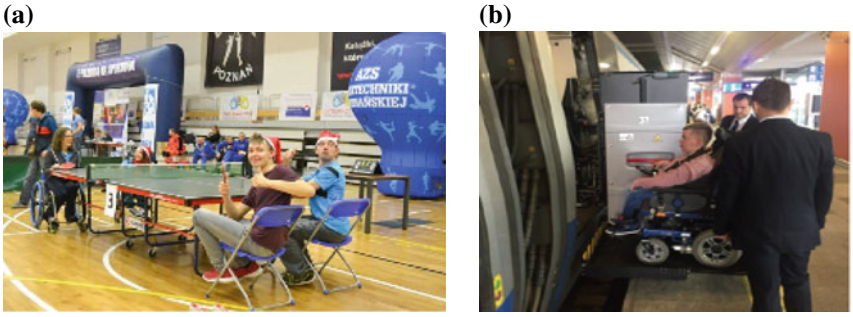


Fig. 4 **a** Picture 1. First Interactional Polish Champion of table tennis: PUT team gold medal. Poznan, 2th December 2016. **b** Picture 2. Journey organization for a person with motor disability. *Source* performed by authors

of the *Association of Students with Disabilities from Poznan University of Technology* "*Outstanding*" (polish:"*Nieprzecięnti*")—cyclical event encouraging to sports life by fun page, regular meeting with the target audience where the NSP project ideas have been tested with the expected, decisive effect (picture 1, 2) (Fig. 4).

In this way the tree determinates were establish—tree different, selected field of activities which have a strong impact for "return to society" for physical handicapped people: (1) regular, (2) reduced and (3) special. The regular level has been created for one person (for person without disabilities), similar the reduced one (for person with disabilities). The special level has been created for two persons (for person without disabilities + for person with disabilities). Those areas have been distinguish in three other aspects of activity: (1) public sport area, (2) short distance, local transport and (3) long distance train public transport, where started as a separate tables of costs contests (Table 2).

The greatest interest between respondents raises a "special packet"—almost 5 EUR for two people. Especially in the field of transport, where able-bodied people could use free journeys, in exactly the direction in which they are interested, to help people with physical disabilities. This is the best way for people who normally must buy a normal ticket, therefore, they pay less, and for a people with disabilities, hence, they receive have a help. *Nshareplatform* (NSP) can easily connect this group of people, who wants use of public goods in the same time. In public transport or cinemas, etc.,—the same dependence of cost calculation could be founded.

Important fact and challenge were created during research analysis. Special ticket is "the secret" knowledge only in the regulations of objects. Yet it is not shared like other information in public tables, even in the web. People mostly do not have any knowledge about it. This lack of awareness means a challenge to be overcome. This is one of the most important aspects of the creation of the mission, vision and value of enterprises, supported by a communicative message that resonates in society. The mission of *Nshareplatform* (NSP) is :
"Elicitation of disable people".

Table 2 Using idea of Nshareplatform (NSP) project in special sectors [in EUR****]

Type of ticket	Using the idea of Nshareplatform in				Public long-distance transport**	Public short-distance, transport***
	Public sport area*					
	1 h	2 h	3 h	All day	One tour	One tour
Regular	5.87	8.92	11.26	13.84	9.62	0.70
Reduced	4.69	7.74	8.92	11.50	4.69	0.35
Special	4.69	7.74	8.92	11.50	5.16	0.00

*Table of cost in the most popular Swim & Aqua & SPA (Sauna) Resort in Poznan, https://www.termymaltanskie.com.pl/en/aquapark/water-park-sport-pools/ (access of the day: 20.06.2019)

**Table of cost in the long-distance train public transport in Poland., https://www.pkp.pl/ (access of the day: 20.06.2019)

***Table of cost in the short-distance, local train public transport in Poznan, https://www.mpk.poznan.pl/ access of the day: (20.06.2019)

****The exchange rate from PLN to EUR = 4.2624; source: National Bank of Poland: www.nbp.pl (access of the day: 26.12.2019)

The vision is:
"Universality of social economy activities",
and the values are: activity, comprehension, unity and community.

The idea of both-side help with benefit (NSP) has a strong social impact. It shows differences between a non-disabled person and a person with a disability in human live. It tries to better communicate two different worlds: ability and disability. It also can help to personalize an assistant—a special person, who helps people with disabilities (GoDP). By analyzing individual stages of travel and related difficulties for people with disabilities, the many available solutions could be found [13]. Each of the designers and manufacturers offers the use of other devices, but the final decision belongs to the infrastructure manager and carrier. Unfortunately, there is no standardization in this respect. This results in the need for individual, individual training of the staff and the conductor team. Therefore, the simplest tools are preferred. *Nshareplatform* (NSP) will create a friendly public space for people with disabilities, understanding they needs.

There is so many types of disabilities [1, 9, 21, 31]. The need for help is a general and global necessity. Many people still have no access to that help. They stay at home [40], because the help is not dedicated for them. Big-Data Analytics tools can be useful and helpful in this case [20].

4.2 MIGPORT as a Liberated Social Entrpreneur Business Plan

Based on mentoring 2000 entrepreneurs, won over 20 awards and participated in top accelerator and incubation programs, the Migport as a LiBerated Business Model

was created—which includes business gap/problem, total targeted market value (in US dollars), competitors, network (existing clients and access to clients and most importantly the entrepreneur in the business).

An entrepreneur is not only the person who innovates but who is also the one who runs a start-up—starts a business and creates a start-up [14]. The entrepreneur is considered having no problem with profit earning, so with the high turn-over of profits, could even establish a monopolistic company which leads to more profits. Social entrepreneurs on the other hand are businesses that generate profits, but the main core of the business is returns to the society. Liberated Social Entrepreneurs are social entrepreneurs who is using business metrics such as profit, revenue and stock exchanges, to sustain social impact in order to change the systems. As a result of Liberated Social Entrepreneur Business Plan the *Migport Program* started. The Migport helps refugees exchange knowledge with each other, locals and organizations regarding their daily problems [25, 26]. Migport's Q&A application collects anonymous real-time data regarding daily problems of refugees in Turkey. Having profiles' categorization, skillsets and inventory, Migport's database function as anonymous digital ID connecting refugees with employment opportunities in Turkey. This program helps organizations prepare better social cohesion programs for refugees with evidence-data-based reports and reach refugees online easily, secured and anonymously. In addition, there exist also started the digital residence permit appointment system "e-residency" in Turkey that over 10 million foreigners used so far. This fact serves as another indicator—are going towards "*big-data*".

After considering this methods whether to initiate The LiBerated Social Entrepreneurship as a start-up business model in eleven, connected steps. The first step addresses the problem (target group and needs), second step deals with the solution, in other words, on how the problem would be solved. The third step focuses on what the product would be, having considered human centered design face-to-face interaction with targeted group and most possible product which they can produce. The fourth step the has social impact evaluation, development and fairness which the start-up can bring. The fifth step has the customers' segment; in this step—to consider who is going to purchase the product whether it would be for the use of targeted group such as the coat product of empowerment plan and micro-loans or any kind product that the target group would produce for the use of all people. The sixth step considers on who the partners are and what the channels are—to reflect on the channels that would distribute the product, reaching the targeted group and possible partners' start-up would operate such as international organizations or regional organizations do. The seventh step bases on the elaboration of the start-up summary, the financial structure of the start-up, whether the entrepreneur would get an investment, a return, sells assets, shares of a stocks, etc. It also summarizes whether the start-up would get donations and funding both from private and public aspects. For some countries, there exist social impact bonds and their margins go to social entrepreneurship firms' possible investments from social impact bond. This seventh step—a very essential one—helps the entrepreneur to decide and combine which business metrics on commercial and social entrepreneurship tools to use. The eighth step is designed to check competition that the liberated social entrepreneur would

face, for example, by similar products or the existence of similar start-ups. The ninth step focus on the cost segment of the start-up, what is needed as a minimum amount of money out of all expenses that constitute the total cost. The tenth step consider the revenue segment—the entrepreneur would examine possible sales of products and the revenue operated within the company. Last but not the least, the eleventh step is the profit step—to subtract cost from revenue. It does not matter whether the amounts are known—yet, if the idea of framework to examine the start up on how much profit would be generated in (1) very short run, (2) in short run, (3) in medium run and in the long run.

After considering profit at each step, the entrepreneur also prepares possible goals in each step to expand the company and lead to "*creative destruction*". As a result, this discussed methodology would provide the answer on whether to start a social business in order to introduce sustainable impacts in developing countries that would provide sustainability, healthcare, mobility, education, in terms of development in the developing countries (see also: [19, 43]). This case study provides guidance and it is a business model for social entrepreneurs—enables them to behave like commercial entrepreneurs, acting as the Migport—LiBerated Social Entrepreneur (Fig. 5).

Neutrally zone between two activities of entrepreneurs has been drawn to as mid-way between commercial entrepreneurship and social entrepreneurship. Social entrepreneur path is not essentially different from the commercial path. Social entrepreneur needs to be aware of market strength and develop a business model for social welfare accordingly [28, 42, 44]. That means—social entrepreneur should create own "creative destruction" and monopolistic start-up in order to survive and keep implementing a social impact. In order that a start-up would survive in the long run, its path passes directly through profits, whatever the aim is. The motto: "*you can never beat market*" means operating in liberal economies. environment. Social entrepreneurs by not combining all metrics are becoming prohibited from the market. Consequently, having all the metrics and combining from both aspects bring about more successful start-up and profits. This neutrally zone name is social entrepreneur, using metrics: profit revenue, increasing stock market price operation in business having outcome as products combined with the aim to improve social impact through employee profile, operation, mission and vision and, most importantly, who bring about social improvement in the society as *LiBerated Social Entrepreneur* [35, 36]. For this type of entrepreneur, consumers would know that a product are not lack of quality since profits are not divided and cut from the production process for social improvement as grant. Social improvement process would be impacting through company's tools such as employee profile, mission and vision. They would have an awareness that through consuming liberated social entrepreneurial products leading by the companies to social improvement, but having the same quality.

The Migport program is also providing entrepreneurship and innovation trainings to help locals and refugees be successful in their businesses—trainings are designed to base on Migport entrepreneurship journey as well as a history of mentoring 2000 entrepreneurs with the programs Migport run such as migrapreneurs, MAV Volunteers, several Migport Entrepreneurship Trainings in European Commission programs, Red Crescent, GIZ, ASAM-SGDD, Mudem and Boston University. In the

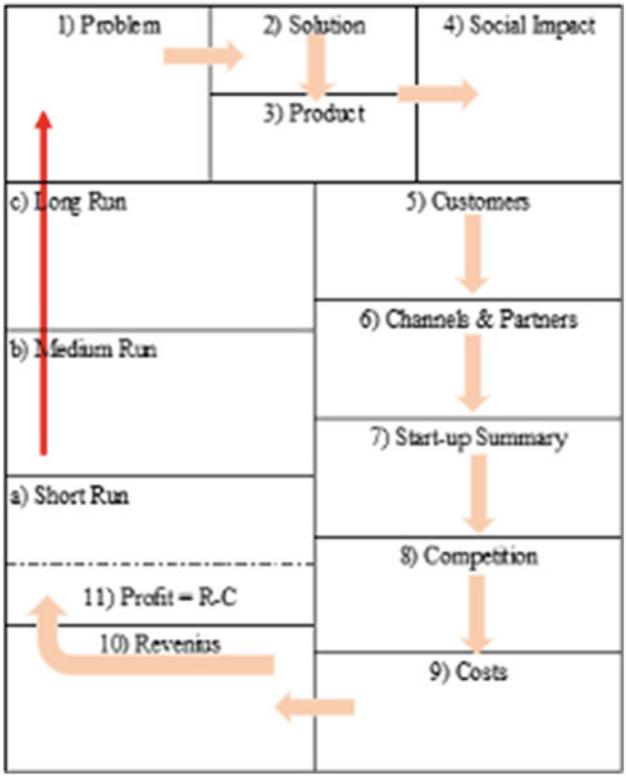


Fig. 5 The *Migport*—LiBerated Social Entrepreneurship project

relevant literature exist Guy Kawasaki 10-page presentation, to explain a business is one of the best pitch templates worldwide that can help entrepreneurs pitch/tell their businesses in simple 10 pages—Guy Kawasaki investor pitch template as an example as well as Simon Sinek Golden Circle model were used. On the other hand, Canvas Business Model is used to write down an idea and see the business how it works and shows value proposition the solution part very well, it doesn't show the gap in the market, market value, problem behind the scenes, competition and most importantly whether an entrepreneur exist in the business which is the key factor for the good businesses that have a big market to be successful [50]. Canvas business model indicates that solution always change—by written down the on the canvas with posted cards, the changing of the value proposition has change, clients and whenever needed. However, what should exist as not change is the total market value, competition, problem gap, entrepreneur and the network—existing base of business customers.

Liberate Business Model analyze startups on five criteria which are less likely to change after the establishment of the business. Those criteria are existing gap in the market that is problem entrepreneur is willing to change, market value in U.S.

Table 3 Evaluation of Migport

Evaluation of Migport—LiBerate Business Model	
71% < X	Most likely to be successful
60–70%	Uncertain
60% > X	Most likely to be unsuccessful

Table 4 Startups already raised Seed

Startups already raised Seed A, B, C or Exit	
90% < X	Fairly evaluated
70–85%	Over-evaluated
85% > X	Under-evaluated

dollars that entrepreneur aims to get or access 1–10% of the market, competitors that try to get the money in the market and which can be possible exit[2] routes for startups, network of entrepreneur such as existing clients and access to reach and talk to existing clients and the entrepreneur who leads the team dedicated, convincing, finds a way, most key factor, and who never gives up.

Firstly, entrepreneurs should define the of solutions, by searching the gaps in all process—real gap which entrepreneurs can start working and people will be able to pay for it. Best gaps in the markets, businesses require knowledge of the field. Entrepreneurs should known to find gaps in two areas: at common market and also at new markets. The 20 start-ups from Turkey were analyzed—successful brends, growing rapidly and which gaps defined as a 60% in distribution (rest gaps—10% in others areas of leading business). During investigation 5 startups failed. The Liberate Business Model anticipated 57% of them that they will not be successful due to Business Gap as the best anticipate criteria for startups are likely to fail. In addition, LiBerate Business Model also anticipated that 60% of uncertain startups but already did SEED A, B, C or Exit that they would be uncertain due to business gap. The combination of market value and business gaps, combination of entrepreneur and the network, combination of business gaps and the gaps of entrepreneur competence (Tables 3, 4 and 5).

Secondly, market value of total U.S. dollars also shows whether there exist possible cash flows for the product and solution. As creating a new market takes sometimes decades, if there does not exist potential and existing market value, success rate of startups depends on market value.

Thirdly, the biggest competitor could exist at the market not always as a one company. This case is strongly difficult because needs more assumption and distributed action. In addition, success rate of startups can be defined as exit. As there are potential and existing competitors, they may acquire the startup shares in the

[2] Exit is the term used to show selling the total or partial shares/stock options to startups to non-share holders.

Table 5 Prediction rates as the best predicted model in 20 startups

Successful and still surviving startups	33%	17%	0%	8%	100%	42%	33%	17%	8%
Uncertain but already did SEED A, B, C or Exit	0%	0%	0%	29%	0%	20%	14%	60%	0%
Closed	0%	0%	0%	20%	0%	14%	0%	57%	0%
Uncertain and closed ones	0%	0%	0%	25%	0%	17%	8%	58%	0%
Shere	Ent 50 & Network 20, rest 10 % each	Gap 50 & Ent 20, rest 10 % each	Each 20 %	60 % = Market $, rest 10 % each	60% = entrepreneur, rest 10 % each	60% = competition, rest 10 % each	60 % = network, rest 10 % each	60 % = Gap, rest 10 % each	Market 40 % & gap 30 %, rest 10 % each

future. Competitors can be seen as potential future exit possibilities. Results: from 20 startups evaluation that when the 60% was given to density to competitors and 10% other four each, the prediction model with competition was the second-best model to predict the startup success rates of successful and still surviving startups with roughly 42%.

Fourthly, existing clients and route to clients which is network also plays very critical role for startups to be successful and make sales. When startups receive seed fund, they are usually being part of angel investors or venture capitalists. Those venture capitalists and angel investor do not only invest money but also their network to invested business to make sales and sell the shares/exit of the business which will be successful. In addition, also network of the entrepreneur matters in this case to analyze whether the entrepreneur can meet with top clients and make sales to them. If the startup is already having clients or top clients in their market that means that startup have a strong network and portfolio. When the 60% exists as a density to network and 10% to rest four criteria each, the network prediction model predicts as third best model whether a startup can be successful with approximately 33% as the best prediction among other models.

Finally, the entrepreneur matter since without an entrepreneur nor the team neither none of other 4 criteria can be accomplished. An entrepreneur should be existing in a startup who leads the team, finds ways and pivots. The 10-star entrepreneur were establish as "dedicated, convincing, finds a way, most key factor, never give up". Although this cannot be mathematical, understanding an entrepreneur can be analyzed with whether its entrepreneurs first startup, whether entrepreneur have an established successful or unsuccessful startup, whether entrepreneur have deep knowledge on the field and the gap, market and competitors. As startup journey is defined as death-valley, surviving is one of the most critical aspects which makes a startup sustainable. This is one way to create a real economic environment which create favor assumption to change micro companies to small one, small to middle and then to the big one. The LiBerate Business Model could be also presented as a Canvas Business Model for starts-up. It is designed as one of the ways to analyze a startup and understand whether the problem startup is trying to solve whether its valid and make money. To having more insights from investors and startups for evaluation and grading on five criteria in the future, as well as more data on other startups, to predict success rates.

5 Conclusions

Entrepreneurship is an interdisciplinary issue. Therefore, it requires multilateral competence, and synergistic action based on the cooperation of specialized units in all fields involved becomes ideal. It takes into consideration the achievements of many sciences, of which the leading should be social sciences: psychology and sociology, next to economics and law, taking into account local institutional and legislative conditions as well as the ability to manage and conduct finances, which should be

considered in the aspect of competence of the managing person—entrepreneur. Social entrepreneurship, on the other hand, combines economic and social aims. Its main goals should be social, and the profits received by assumption should be reinvested or invested in expanding the business, and not serve the enrichment of owners. However, according to the authors, they are always a path supporting the primary goal of undertaking economic activity—profit. The presented programs: The Nshareplatform (NSP) and The Migport LiBerated Social Entrepreneurs fit in with both the goal and mission of action in the assumptions of social entrepreneurship. The Migport—LiBerated Social Entrepreneurs Program allow for sustainable impacts in developing countries for long-term solutions that would provide sustainability, healthcare, mobility, education in the developing countries though their liberated social entrepreneurship start-ups, having a direct effect on long-term solutions and though their profits impacting development in their liberal developing economies. The idea of both-sided help with benefit The Nshareplatform (NSP) ought to exercise a strong social impact in few opinions. It shows differences between a non-disabled person and a person with a disability in live. NSP tries to better communicate between two different worlds, called as ability and disability, respectively. It also can help to personalize an assistant (GoDP)—a particular people, who serves persons with disabilities. The development of NSP extend research scopes of Migport towards further groups of humans in need, while further benefitting from the economics, operational research and analytics methods and studies.

As a conclusion, a list of recommended units has been created that can benefit from the application of the Nshareplarform program in Poland and Liberated Social Entrepreneur and Migport in Turkey:

(1) SDG Impact Pre-Accelerator, Ministry of Foreign Affairs of Turkey, UNDP, Bill and Melinda Gates Foundation, Eczacibasi, Limak Holding and WFP
(2) PARP—Polish Agency for Enterprise Development
(3) Young Change Maker, Middle East Mediterranean Summit, USI, Lugano, Switzerland
(4) PEFRON—State Found for Rehabilitation of Disable People
(5) Samsun Region Social Entrepreneur Winner, TÜBİTAK
(6) Local Governments
(7) International Federation of OR Societies, its newsletter and DC online resources
(8) SOW of PEFRON—Support System for financing cooperating with PFRON
(9) International Visitor Leadership Program Fellowship
(10) Office of the Government Plenipotentiary for Disabled People
(11) Transatlantic Young Innovation Leaders Initiative Turkey Fellowship, Marshal Fund
(12) Polish Association of the Blind, Deaf and Disabled Sports “START”
(13) The Collective Global Acceleration London, top 11 of 5000 techno-social applications
(14) NFZ—National Health Found
(15) Social Impact Award, EO—Global Student Entrepreneur Awards, Toronto, Canada

(16) Polish Private Equity & Venture Capital Association
(17) MIT EF Pan Arab final Innovate for Refugees, semi-finalist, Amman, Jordan
(18) Foundations supporting disabled people
(19) Top 100 Projects & METU Technopolis Office Award; TechAnkara Project Market, Ankara Development Agency
(20) Eurostat
(21) Top 50, Startup Istanbul 2017, 140+ countries, 21 thousand applicants
(22) Start-up Poland
(23) Winner, Young Entrepreneurship Development Program, Ankara Development Agency & The Ministry of Labor and Social Security
(24) OECD—Organisation for Economic Co-operation and Development.

Internet sources:
www.termymaltanskie.com.pl/en/aquapark/water-park-sport-pools/ access of the day: 20.06.2019.
www.pkp.pl access of the day: 20.06.2019.
www.mpk.poznan.pl/ access of the day: 20.06.2019.
www.bankier.pl/consciousosc/Turecka-gielda-tylko-dla-inwestorow-o-mocnych-nerwach-7441545.html access od the day:22.07.2016.
www.nbp.pl access of the day: 26.12.2019
www.gpw.pl/statystyka-gpw access of the day 27.12.2019.
www.nieprzecietni.put.poznan.pl/index.php?zm=kimJestesmy access of the day: 29.12.2019.

References

1. Alvord, S.H., Brown, L.D.: Social entrepreneurship Leadership that facielitates societal transformation an exploratory study. Ceter for Public Lidership (2003)
2. Alvord, S.H., Zimmer, C.: Entrepreneurship though social network. In: Sexton, D., Smilor, R. (ed.) The Art and Science of Entrepreneurship. Ballinger Publishing Company, Cambridge MA.s.3–23 (1986)
3. American Foundation for the Blind. (2006). Blindness statistics. Retrieved January 28, 2006
4. Babelyuk, V.Y., Gozhenko, A.I., Dubkova, G.I., Korolyshyn, T.A., Kikhtan, V.V., Babelyuk, N., Popovych, I.L.: Causal relationships between the parameters of gas discharge visualization and principal neuroendocrine factors of adaptation. J. Phys. Educ. Sport **17**(2), 624–637 (2017). https://doi.org/10.7752/jpes.2017.02094
5. Beattie, G.: Visible Thought. Routledge Taylor & Francis Group, Londyn, The New Psychology of Body Language (2003)
6. Bohman, P.: Introduction to Web accessibility (2003). Retrieved November 25, 2005, from: https://webaim.org/intro/?templatetype=3
7. Brooks, A.C.: Social Entrepreneurship: A Modern Approach to Social Value Creation, 1st edn. Pearson Education, New Jersey (2009)
8. Bullock, M.: Tennis for the Blind and Partially Sighted, [w:] Coaching & Sport Science Review, International Tennis Federation, 2007 [online], http://www.tennisexplorer.narod.ru/English_Articles/ITF2007.PDF#page=10. Accessed 7 June 2019
9. Business Dictionary, Inc (2017). Definition of entrepreneurship Web site: http://www.businessdictionary.com/definition/entrepreneurship.html

10. Byzdran, K., Skrzypczynska, A., Piątek, M., Stepniak, R.: Physical activity, and the development of physical fitness in boys aged 13–15. J. Health Sci. **3**(10), 261–274 (2003)
11. Cantillon, R.: Essai Sur la Nature du Commerce en General. [in] translation H.Hoggins 1931 (1755). Macmillan, London
12. Crow, K.L.: Accommodating on-line postsecondary students who have disabilities. Dissertation submitted for publication, Northern Illinois University (2006)
13. Di Cagno, A., Iuliano, E., Aquino, G., Fiorilli, G., Battaglia, C., Giombini, A., Calcagno, G.: Psychological well-being and social participation assessment in visually impaired subjects playing Torball: a controlled study. Res. Dev. Disab. **34**, 1204–1209 (2013)
14. Egirisim, A.: Startup Istambul 2017 de ikinci aşamaya geçen 50 girisim (2017). Retrieved from: https://egirisim.com/2017/10/21/startup-istanbul-2017de-ikinci-asamaya-gecen-50-girisim/
15. Entrepreneurs' Organization (2018). Qzenobia to run for finals in MIT Enterprise Forum before GSEA| Retrieved from: https://www.eonetwork.org/turkey/chapterpressreleases/qzenobia-to-run-for-finals-in-mit-enterprise-forum-before-gsea
16. Friedman, M., Friedman, R.: Free to Choose: A Personal Statement. New York, N.Y: Harcourt Brace Jovanovich. Inc (1980). http://www.proglocode.unam.mx/sites/proglocode.unam.mx/files/docencia/Milton%20y%20Rose%20Friedman%20-%20Free%20to%20Choose.pdf
17. Galbraith, J.K.: Market Structure and Stabilization Policy. Rev. Econ. Stat. **39**(2), 124–133 (1957). https://doi.org/10.2307/1928529
18. Galbraith, J.K.: The bimodal image of the modern economy: remarks upon receipt of the Veblen-commons award. J. Econ. Issues **11**(2), 189–199 (1977)
19. Harris, C.B., Welsby, P.D.: Health advice and the traveller. Scottish Med. J. **45**(1), 14–16 (2000)
20. Kaul, I., Grunberg, I., Stern, M.A.: Global Public Goods: International. Cooperation in the 21st Century, United Nations Development Program, New York (1999)
21. Keynes, J.M.: Ogólna teoria zatrudnienia, procentu, pieniądza. PWN, Warszawa (1956)
22. Keynes, J.M.: The Economic Consequences of the Peace. Harcourt Brace & Howe, New York (1920)
23. Kirzner, I.: The entrepreneurial market process—an exposition. Southern Econ. J. **83**(4), 855–868 (2017). https://doi.org/10.1002/soej.12212. WOS: 000402904100002
24. Kirzner, I.: Information—knowledge and action—knowledge. Econ. J. Watch **2**(1), 75–81. WOS: 000203013400009 (2005)
25. Kjamili, B., Weber, G.W.: The role of LiBerated Social Entrepreneur in Developing Countries: A mid-way, in Societal Complexity, Data Mining and Gaming; State-of-the-Art 2017, Greenhill & Waterfront, Europe: Amsterdam, The Netherlands; Guilford, UK North-America: Montreal, Canada, 2017. ISBN /EAN (2017). 978-90-77171-54-7 (2005)
26. Kjamili, B., Weber, G.W.: (2014) IFORS: Opening Doors to International Students Retrieved from: http://ifors.org/newsletter/ifors-dec-2014.pdf
27. Kotarbiński, T.: Traktat o dobrej robocie. Ossolineum, Warszawa (1973)
28. Kramer, N.R.: Measuring innovation: Evaluation in the filedof sicial entrepreneurship. The skoll Foundation (2005)
29. Lange, O.: Ekonomia polityczna. PWN, Warszawa (1974)
30. Leibenstein, H.: Entrepreneurship, entrepreneurial training and X-efficiency theory. J. Econ. Behav. Organ. **8**(2), 191–205. Harvard Univ: Cambridge (1987). https://doi.org/10.1016/0167-2681(87)90003-5. WOS: A1987J422100003
31. Mackelprang, R.W., Salsgiver, R.O.: Disability: A Diversity Model Approach in Human. Brooks/Cole Publishing Company, Toronto, Service Practice (1999)
32. Małecka, J.: Comparison of entrepreneurial attitudes—a polish and Ukrainian case study. 12th ICEBE 2019, 12th International Conference on Engineering and Business Education (2019)
33. Małecka, J.: Knowledge management in SMEs— in search of a paradigm. Proceedings of the 19th European Conference of Knowledge Management. Published by Academic Conferences and Publishing International Limited Reading, UK. E-Book: ISBN: 978-1-911218-95-1. E-BOOKISSN: 2048-8971. Book version ISBN: 978-1-911218-94-4; Book Version ISSN: 2048-8963. pp. 485–493 (2018)

34. Małecka, J.: The perception of quality in Qualitology—aspects. The Proceedings of the 17th European Conference on Research Methodology for Business and Management Studies. Published by Academic Conferences and Publishing International Limited Reading, UK (2018). WOS:000461833200032
35. Małecka, J.: Venture capital as a source of financing small and medium-sized enterprises in Poland: selected aspects. Proceedings Paper Advancing research in Entrepreneurship in the Global Context p. 669–684. Krakow: Fundation Cracowv Univ Economics. WOS:000400596500044 (2016)
36. Małecka, J.: Revenues, expenses, profitability and investments of potential contenders for the status of a listed company in Poland. Oeconomia Copernicana **6**(4), 91–122. https://doi.org/10.12775/OeC.2015.031; WOS:000216511300006 (2015)
37. [MIT] Enterprise Forum (2018). Innovate for Refugees semifinalists. Retrieved from: https://innovateforrefugees.mitefarab.org/en/site/semifinalists
38. Mulgan, G., Tucker, S., Ali, R., Sanders, B.: Social innovation: what it is, why it metters and how it can be accelerated. Skool Center for Socila (2008). Enterpreneurship
39. Moore, J.: Extreme-do-gooderswhat maker them tick. Christien Science Monitor (2009)
40. [NYT] Liviing Whith Disability (2016). The New York Times. Retrieved from https://www.nytimes.com/2016/08/28/opinion/living-with-disability.html
41. Parsons, T., Smelser, N.J.: Economy and society. A Study of Integration of Economics and Social Theory, New York (1957)
42. Praszkier, R., Nowak, A.: Przedsibiorczość społeczna. Teoria i Praktyka. Wolters Kulwer Polska Sp. z o.o, Warszawa (2012)
43. Raghupathi, W.: Data Mining in Health Care. Healthcare Informatics: Improving Efficiency and Productivity. Edited by: Kudyba S. 2010, Taylor & Francis, 211–223 (2010)
44. Rogers, E.M.: Generalized expectancies for internal versus external control of reinforcement. Psychol. Monographs. **80**(1), 1–28 (2003)
45. Schnickus, C.: The valuation of social impact bonds: an introductory perspective with the Peterborough SIB. Res. Int. Bus. Finance, Elsevier **35**(C), 104–110 (2015)
46. Schumpeter, J.: Business Cycles. A Theoretical, Historical and Statistical Analysis of the Capital Process: Philadelphia: Porcupine Press (1989)
47. Seeman, L.: Inclusion of cognitive disabilities in the web accessibility movement. Paper presented at the 11th International World Wide Web Conference, Honolulu, Hawaii (2002)
48. Smith, A.: The Wealth of Nations. Random House Lcc Us (2000)
49. Smith, A.: An Inquiry Into the Nature and Causes of the Wealth of Nations, T.1& 2. FB & C Ltd (2016)
50. Strategyzer, Inc (2017). The Business Model Canvas Website: https://strategyzer.com/canvas/business-model-canvas
51. Supiński, J.: Szkoła polska gospodarstwa społecznego. [in:] Dzieła t. II i III, Lwów (1872)
52. Szymański, Z.: Józefa Supińskiego teoria rozwoju społeczno-gospodarczego. Wydawnictwo Uniwersytetu Marii Curie - Skłodowskiej, Lublin (1999)
53. Takei, M.: Lecture to 5th grade students in Takorozawa, Saitama 2007 [online], http://www.hanno.jp/matsui/Takei%20lecture.pdf. Accessed 20 June 2018
54. Weber, M.: Die protestanishe Ethik unde der Gheist des Kapitalismum. [in]: Gesamnte Aufsaetze zur religionssoziologie, t.1.Tuebingen (1920)
55. Weber, M.: Economy and society: an outline of interpretive sociology (an excerpt). Sotsiologiya **19**)(3), 68–78 (2018). https://doi.org/10.17323/1726-3247-2018-3-67-78. WOS: 000434049500005

An Iterative Process for Approximating Subactions

Hermes H. Ferreira, Artur O. Lopes, and Elismar R. Oliveira

Abstract We describe a procedure based on the iteration of an initial function by an appropriated operator, acting on continuous functions, in order to get a fixed point. This fixed point will be a calibrated subaction for the doubling map on the circle and a fixed Lipschitz potential. We study analytical and generic properties of this process and we provide some computational evaluations of subactions using a discretization of the circle. The fixed point is unique if the maximizing probability is unique. We proceed a careful analysis of the dynamics of this operator close by the fixed point in order to explain the difficulty in estimating its asymptotic behavior. We will show that the convergence rate can be in some moments like $1/2$ and sometimes arbitrarily close to 1.

Keywords Maximizing probability · Subaction · Iterative process · Fixed point · Convergence rate

1 Introduction

Here we analyze some properties of an iterative process (applied to an initial function) designed for approximating subactions. Properties for a general form of such kind of algorithm were considered in [12, 17, 22, 28] (see also [1] for more recent results). We analyze here the performance of a specific version of the algorithm which is useful in Ergodic Optimization.

In a companion paper [13] we will consider several examples. The sharp numerical evidence obtained from the algorithm permits to guess explicit expressions for the subaction.

H. H. Ferreira · A. O. Lopes (✉) · E. R. Oliveira
Instituto de Matemática e Estatística - UFRGS, Porto Alegre, Brazil

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_8

We identify $\mathbb{R}/\mathbb{Z}$ as S^1 and $T(x) = 2x$ the doubling map. We denote by $\tau_i(x) = \frac{1}{2}(x + i - 1),\ i = 1, 2$ the two inverse branches of T.

Definition 1 Given a continuous function $A : S^1 \to \mathbb{R}$ (or, $A : [0, 1] \to \mathbb{R}$) we denote by

$$m(A) = \sup_{\rho \text{ invariant for } T} \int A \, d\rho.$$

Any invariant probability μ attaining such supremum is called a **maximizing probability**.

The properties of the maximizing probabilities μ are the main interest of Ergodic Optimization (see [2, 7, 16, 18–21])

In Statistical Mechanics the limits of equilibrium probabilities when temperature goes to zero (see [2]) are called ground states (they are maximizing probabilities).

A interesting line of reasoning is the following: there is a theory, someone gives a particular example which leads to a problem to solve, then, use the theory to exhibit the solution. Is there a general procedure to find the solution of this kind of problem? Here we will address this kind of query on the present setting.

Definition 2 Given the Lipschitz continuous function $A : S^1 \to \mathbb{R}$ the union of the supports of all the maximizing probabilities is called the **Mather set** for A.

We will assume from now on that A is Lipschitz continuous and that the maximizing probability is unique.

It is known that for a generic Lipschitz potential A (in the Lipschitz norm) the maximizing probability is unique and has support on a T-periodic orbit (see [7, 9]). We do not have to assume here that the unique maximizing probability has support on a unique periodic orbit.

Definition 3 Given the Lipschitz continuous function $A : S^1 \to \mathbb{R}$, then a continuous function $u : S^1 \to \mathbb{R}$ is called a **calibrated subaction** for A, if, for any $x \in S^1$, we have

$$u(x) = \max_{T(y)=x} [A(y) + u(y) - m(A)]. \tag{1}$$

Note that if u is a calibrated subaction for A then u plus a constant is also a calibrated subaction for A.

For Lipschitz potentials A there exists Lipschitz calibrated subactions (see [5, 7]). If the **maximizing probability is unique** (our assumption) then the **calibrated subaction is unique** up to adding a constant (see [7] or [15]).

Calibrated subactions play an important role in Ergodic Optimization (see [2, 16, 27]). From an explicit calibrated subaction one can guess where is the support of the maximizing probability. Indeed, given u we have that for all $x \in S^1$.

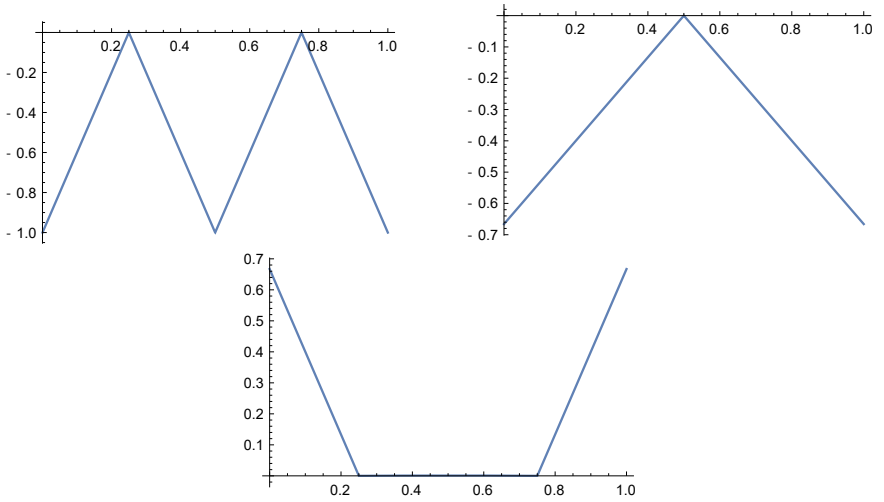


Fig. 1 From left to right: the graph of the potential A, the graph of the calibrated subaction u and the graph of R

$$R(x) := u(T(x)) - u(x) - A(x) + m(A) \geq 0, \tag{2}$$

and, for any point x in the Mather set $R(x) = 0$. Moreover, if an invariant probability has support inside the set of points where $R = 0$, then, this probability is maximizing (see [7]).

In [3] it is presented explicit expressions for the subaction in some nontrivial cases.

Example 1 We show in Fig. 1 the graph of a potential A, the graph of the calibrated subaction u and the graph of R. The potential A is zero at the points $1/4, 3/4$ and it is equal to -1 in the points $0, 1/2, 1$. The set $\{1/3, 2/3\}$ is contained on the Mather set (then, it is the support of a maximizing probability) and $m(A) = -1/3$. The calibrated subaction is 0 at the point $1/2$ and equal to $2/3$ at the points $0, 1$. The function R is equal to $2/3$ at the points $0, 1$ and it is equal to zero on the interval $[1/4, 3/4]$. We point out that we easily guessed the explicit expression for the subaction u from the picture obtained from the application of the algorithm on the initial condition $f_0 = 0$.

Given x, then, $u(x) = A(\tau_j(x)) + u(\tau_j(x)) - m(A)$, for some $j = 1, 2$. We say that $\tau_j(x)$ is a **realizer** for x. There are some points x that eventually get at the same time two realizers.

We are interested in an iteration procedure for getting a good approximation of the subaction in the case the maximizing probability is unique. As a byproduct we will also get the value $m(A)$. This will help to get R (as above) and eventually to find the support of the maximizing probability.

We will consider a map $\mathscr{G}$ acting on functions such that the subaction u is the unique fixed point (we will have to consider the action on continuous functions up to an additive constant). Unfortunately, $\mathscr{G}$ is not a strong contraction but we know that $\lim_{n\to\infty} \mathscr{G}^n(f_0) = u$ (for any given f_0). The performance of the iteration procedure is quite good and one can get easily nice approximations.

We explore here in Sect. 3 the generic point of view on the set of continuous functions. Given a fixed Lipschitz potential A we will show generic properties for the iterative process acting on continuous functions. In this direction expression (15) (and (16)) in Theorems 4, 3, Corollary 1 and also expression (14) in Remark 3 will provide this, and, therefore justify the excellent performance one can observe for the iterative process which we will describe here.

A natural question: when the calibrated subaction is unique is there an uniform exponential speed of approximation (or, something numerically good) of the iteration $\mathscr{G}^n(f_0)$ to the subaction? At least close by the subaction? In Sect. 4 we present a very detailed analysis of the action of the map $\mathscr{G}$ close by the fixed point u and we will show that this is not the case. We will consider in Example 5 a case where where $|\mathscr{G}(f_\varepsilon) - \mathscr{G}(u)| = |f_\varepsilon - u|$, $\varepsilon > 0$, for f_ε as close as you want to the calibrated subaction u. In the positive direction one can also show that close by u there are other g_ε, $\varepsilon > 0$, such that, $|\mathscr{G}(g_\varepsilon) - \mathscr{G}(u)| = 1/2\,|g_\varepsilon - u|$ (see Corollary 2).

Remark 1 We emphasize the fact that in our computational evaluations we are not going to consider numerical aspects of this iteration process as rate of convergence, complexity or comparative efficiency with respect to other numerical schemes. First because it is not our goal and more important because, as we are going to prove, there exist a generic obstruction to get an analytical precise estimate for the convergence nearby the fixed point. We will show (see Sect. 4) that the convergence rate can be in some moments like $1/2$ (at each iteration) and sometimes arbitrarily close to 1 (at each iteration).

For related numerical computations we refer the reader to [10, 11]. In these two papers the authors define a general rigorous approach to discretize points on an interval (considering a finite lattice of points) and also to discretize the action of some operators similar to the ones we will consider here. The aim is to find controlled approximations of a fixed point function for this discretized operator acting on a discrete lattice. One could employ the same ideas here with the appropriate adaptation but this is not the purpose of the present paper.

One final comment: there are two major settings that people analyze questions in Ergodic Optimization: (1) when it is assumed the potential is just continuous, and, (2) when it is assumed some regularity (as Lipschitz for instance) on the potential. The two cases are conceptually distinct: in the first case, generically, the maximizing probability has support on the all space (see [6, 18]) and in the second case, generically, the support has support on a periodic orbit (see [7, 9]). In the first case, generically, subactions are of no help. It is in the second case that subactions are of great help for identifying the support of the maximizing probability. In our work we introduce a nice tool for identifying, generically, the maximizing probability (see [13]).

2 The 1/2 Iterative Procedure

On the set of continuous functions $f : S^1 \to \mathbb{R}$ we consider the sup norm: $|f|_0 = \sup\{|f(x)|,\ x \in S^1\}$. This set is denoted by $C^0 = C^0(S^1, \mathbb{R})$.

Definition 4 In $C^0(S^1, \mathbb{R})$ we consider the equivalence relation $f \sim g$, if $f - g$ is a constant. The set of classes is denoted by $\mathscr{C} = C^0/\mathbb{R}$ and, by convention, we will consider in each class a representative which has supremum equal to zero.

In $\mathscr{C}$ we consider the quotient norm (see Sect. 7.2 in [25])

$$|f| = \inf_{\alpha \in \mathbb{R}} |f + \alpha|_0.$$

We can also consider this norm $|f|$ restricted to set of Lipschitz functions in $\mathscr{C}$. $(\mathscr{C}, |\cdot|)$ is a Banach space (see [25]) . As S^1 is compact we get that: for any given f there exists α, such that, $|f| = |f + \alpha|_0$.

We denote sometimes the constant α associated to f by $\alpha_f := -\frac{\max f + \min f}{2}$. We point out that when we write $|f(x)|$ this means the modulus of an element in $\mathbb{R}$ and $|f|$ means the norm defined above.

Definition 5 Given a Lipschitz continuous function $A : S^1 \to \mathbb{R}$ we consider the operator (map) $\hat{\mathscr{L}} = \hat{\mathscr{L}}_A$, such that, for $f : S^1 \to \mathbb{R}$, we have $\hat{\mathscr{L}}_A(f) = g$, if

$$\hat{\mathscr{L}}_A(f)(x) = g(x) = \max_{T(y)=x} [A(y) + f(y) - m(A)]. \tag{3}$$

for any $x \in S^1$.

For the given Lipschitz continuous function $A : S^1 \to \mathbb{R}$ the operator $\hat{\mathscr{L}}_A$ acts in $\mathscr{C}$ as well as in $\mathscr{C}_0$.

Note that u is a fixed point for such operator $f \to \hat{\mathscr{L}}_A(f)$, if and only if, u is a calibrated subaction. It is well known there exists calibrated subactions when A is of Lipschitz class (see for instance [2]).

One could hope that a high iterate $\hat{\mathscr{L}}_A^n(f_0)$ (n large) would give an approximation of the calibrated subaction. This operator will not be very helpful because we have to known in advance the value $m(A)$. Even if we know the value $m(A)$ the iterations $\hat{\mathscr{L}}_A^n(f_0)$ applied on an initial continuous function f_0 may not converge. This can happen even in the case the calibrated subaction is unique.

Definition 6 Given a Lipschitz continuous function $A : S^1 \to \mathbb{R}$ we consider the operator (map) $\mathscr{L} = \mathscr{L}_A : \mathscr{C} \to \mathscr{C}$, such that, for $f : S^1 \to \mathbb{R}$, we have $\mathscr{L}_A(f) = g$, if

$$\mathscr{L}_A(f)(x) = g(x) = \max_{T(y)=x} [A(y) + f(y)] - \sup_{s \in S^1}\{\max_{T(r)=s} [A(r) + f(r)]\}. \tag{4}$$

for any $x \in S^1$.

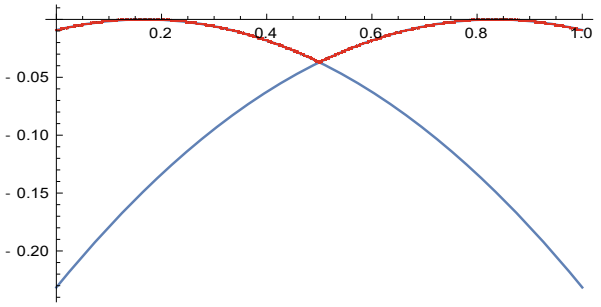


Fig. 2 Case $A(x) = -(x - 1/2)^2$ and $T(x) = -2x \pmod 1$—In this case $m(A) = -1/36$. The red graph describes the values of the approximation (via 1/2-algorithm) to the calibrated subaction u given by $\mathscr{G}^{10}(0)$ (using the language C^{++} and a mesh of points) and the two blue graphs describe, respectively, the graphs of $x \to -1/3x^2 + 1/9x$, and $x \to -1/3x^2 + 5/9x - 2/9$. The supremum of these two functions is the exact analytical expression for the graph of the calibrated subaction u. The red color obliterates the blue color

The advantage here is that we do not have to know the value $m(A)$. In the same way as before u is a fixed point for the operator $\mathscr{L}_A(f)$, if and only if, u is a calibrated subaction.

We call the iterative procedure (defined below and denoted by $\mathscr{G}$) the 1/2-iterative process. It is a particular case of the iteration procedure described on [12, 17]. From these two papers it follows that given any initial function $f_0 \in \mathscr{C}$ we have that $\lim_{n\to\infty} \mathscr{G}^n(f_0)$ exists and it is the subaction u (which belongs to $\mathscr{C}$).

Remark 2 The iterations $\mathscr{L}_A^n(f_0)$ applied on an initial continuous function f_0 may not converge. This can happen even in the case the calibrated subaction is unique as some examples can show. The bottom line is: we have to use $\mathscr{G}$ and not $\mathscr{L}_A$. $\diamond$

In order to show the power of the approximation scheme we consider an example where the subaction u was already known. The dynamics is $T(x) = -2x \pmod 1$ (not $T(x) = 2x \pmod 1$). The 1/2-algorithm works also fine in this case. According to Example 5 in pp. 366–367 in [24] the subaction u (see picture on p. 367 in [24]) for the potential $A(x) = -(x - 1/2)^2$ is

$$u(x) = \max\{-1/3x^2 + 1/9x,\ -1/3x^2 + 5/9x - 2/9\}.$$

More generally, in p. 391 in [24] is described a natural procedure to get the subaction u for potentials A which are quadratic polynomials. The maximizing probability μ in this case has support on the orbit of period two (according to [19–21]) and $m(A) = -1/36$. One can see from Fig. 2 a perfect match of the solution obtained from the algorithm described by $\mathscr{G}$ and the graph of the exact calibrated subaction u.

Definition 7 Given a Lipschitz continuous function $A : S^1 \to \mathbb{R}$ we consider the operator (map) $\mathscr{G} = \mathscr{G}_A : \mathscr{C} \to \mathscr{C}$, such that, for $f : S^1 \to \mathbb{R}$, we have $\mathscr{G}_A(f) = g$, if

$$\mathscr{G}_A(f)(x) = g(x) = \frac{\max_{T(y)=x}[A(y) + f(y)] + f(x)}{2} - c_f$$

for any $x \in S^1$, where

$$c_f := \sup_{s \in S^1} \frac{\max_{T(r)=s}[A(r) + f(r)] + f(r)}{2}. \tag{5}$$

We will show later in Theorem 2 that $|\mathscr{G}(f) - \mathscr{G}(g)| \leq |f - g|$, for any $f, g \in \mathscr{C}$. Therefore, $\mathscr{G}$ is Lipschitz continuous.

The operator $\mathscr{G}$ is not linear. As we already mentioned we called the procedure based on high iterations $\mathscr{G}^n(f_0)$ the $1/2$ iterative procedure.

The above Definition 7 was inspired by expressions (5.1) and (5.2) of [8]. This is a particular case of a more general kind of numerical iteration procedure known as the Mann iterative process (see [12, 17, 22, 26, 28]).

Assuming that the subaction u for the Lipschitz potential A is unique (up to adding constants) it follows (as particular case) from the general results of W. Dotson, H. Senter and S. Ishikawa (see Corollary 1 in [12, 17, 28]) that $\lim_{n\to\infty} \mathscr{G}^n(f_0) = u$, for any given $f_0 \in \mathscr{C}$.

The special $\mathscr{G}$ presented above was not previously consider in the literature (as far as we know).

Note that $\mathscr{G}_A(f + c) = \mathscr{G}_A(f)$ if c is a constant and also that for any f the supremum of $\mathscr{G}_A(f)$ is equal to 0.

When running the iteration procedure on a computer (using the language C^{++}) one fix a mesh of points in $[0, 1]$ and perform the operations on each site. The pictures we will show here are obtained in this way when we consider a large number of points equally spaced.

One important issue on the companion paper [13] with explicit examples is corroboration. By this we mean: we derive analytically some complicated expressions and we use the algorithm to compare and confirm that our reasoning was correct.

The next proposition is a direct consequence fo the definition of $\mathscr{G}_A$ but we will present a proof for the benefit of the reader.

Proposition 1 *If u is such that $\mathscr{G}_A(u) = u$, then, u is a calibrated subaction and*

$$m(A) = \sup_z \max_{T(y)=z} [A(y) + u(y)] + u(z). \tag{6}$$

Proof If

$$u(x) = \frac{\max_{T(y)=x}[A(y) + u(y)] + u(x)}{2} - c_u, \tag{7}$$

then, for all x, we obtain $u(x) = \frac{\max_{T(y)=x}[A(y)+u(y)]+u(x)}{2} - c$, where $c = c_u = \sup_z \frac{\max_{T(y)=z}[A(y)+u(y)]+u(z)}{2}$ is constant. This means that

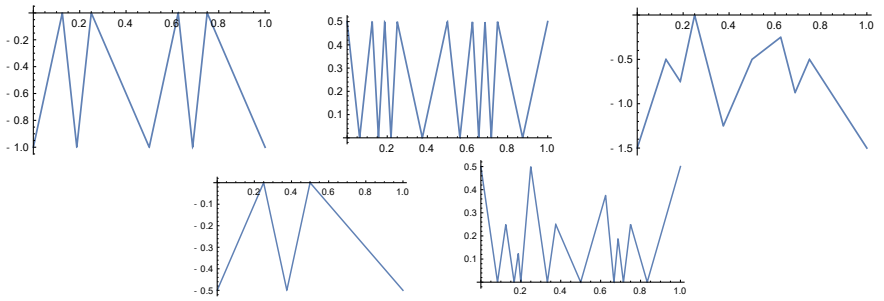


Fig. 3 On the top: from left to right the graph of $A = f_0$, the graph of $x \to |(f_0(x) - 0) + 0.5|$, the graph of $f_1 = \mathscr{G}(f_0)$. On the bottom: from left to right the graph of $g_1 = \mathscr{G}(0) = \mathscr{G}(g_0)$ and the graph of $x \to |f_1(x) - g_1(x) + 0.5|$. Therefore, $\mathscr{G}$ is not a strong contraction because $|f_0 - g_0| = 1/2 = |f_1 - g_1| = |\mathscr{G}(f_0) - \mathscr{G}(g_0)|$

$$2u(x) = \max_{T(y)=x} [A(y) + u(y)] + u(x) - 2c,$$

and, finally, we get $u(x) = \max_{T(y)=x}[A(y) + u(y)] - 2c$, for any x.

In the end of the proof of Theorem 11 in [4] it is shown that this implies that $m(A) = 2c$ and it follows that u is a calibrated subaction. □

Counter Example 1: $\mathscr{G}$ may not be a strong contraction (by a factor smaller than 1). We will present an example where $f_0, g_0 \in \mathscr{C}$ but $|\mathscr{G}(f_0) - \mathscr{G}(g_0)| = 1/2 = |f_0 - g_0|$.

Consider the potential A with the graph given by Fig. 3. This potential is linear by parts and has the value 0 on the points $1/8, 1/4, 3/4, 7/8$. The value -1 is attained at the points $0, 3/16, 1/2, 13/16, 1$.

Denote $g_0 = 0$ and $f_0 = A$. Then, $|f_0 - g_0| = |f_0 - g_0 + 1/2|_0 = 1/2$. We deno te $f_1 = \mathscr{G}(f_0)$ and $g_1 = \mathscr{G}(g_0)$. The graph of the function $x \to |f_1(x) - g_1(x) + 0.5|$ is described by the bottom rigth picture on Fig. 3. One can show that $|f_1 - g_1| = |f_1 - g_1 + 1/2|_0 = 1/2$. Therefore, for such potential A the transformation $\mathscr{G}$ is not a strong contraction. Theorem 2 shows that $\mathscr{G}$ is a weak contraction.

◇

For a fixed $K > 0$ we denote by $\mathscr{C}_K$, the set of Lipschitz functions $f : S^1 \to \mathbb{R}$ in $\mathscr{C}$, with Lipschitz constant smaller or equal to K. By Arzela-Ascoli Theorem $\mathscr{C}_K$ is a compact space in $\mathscr{C}$.

Theorem 1 *Suppose A has Lipschitz constant equal to K. Then, $\mathscr{G}(\mathscr{C}_K) \subset \mathscr{C}_K$. Therefore, the image of $\mathscr{C}_K$ by $\mathscr{G}$ is compact for the quotient norm in $\mathscr{C}$.*

Proof Denote $f_1 = \mathscr{G}(f_0)$.

Given a point y assume without loss of generality that $f_1(x) - f_1(y) \geq 0$.

Then,

$$f_1(x) - f_1(y) \leq \left[\frac{A(\tau_{a_0^{x,f_0}}(x))}{2} + \frac{1}{2}(f_0(\tau_{a_0^{x,f_0}}(x)) + f_0(x)) \right] -$$

$$\left[\frac{A(\tau_{a_0^{x,f_0}}(y))}{2} + \frac{1}{2}(f_0(\tau_{a_0^{x,f_0}}(y)) + f_0(y)) \right] =$$

$$\frac{1}{2}[A(\tau_{a_0^{x,f_0}}(x)) - A(\tau_{a_0^{x,f_0}}(y))] + \frac{1}{2}[f_0(\tau_{a_0^{x,f_0}}(x)) - f_0(\tau_{a_0^{x,f_0}}(y))] + \frac{1}{2}[f_0(x) - f_0(y)] \leq$$

$$K\,\frac{1}{2}\,|\tau_{a_0^{x,f_0}}(x)) - \tau_{a_0^{x,f_0}}(y)| + K\,\frac{1}{2}\,|\tau_{a_0^{x,f_0}}(x)) - \tau_{a_0^{x,f_0}}(y)| + \frac{1}{2}K\,|x-y| =$$

$$K\,\frac{1}{2}\,\frac{1}{2}|x-y| + K\,\frac{1}{2}\,\frac{1}{2}|x-y| + \frac{1}{2}K\,|x-y| = K\,|x-y|.$$

□

The next theorem is a direct consequence of the nonexpansiveness of $\mathscr{L}_A$ but we will present a proof for the benefit of the reader.

Theorem 2 *Given the functions $f, g \in \mathscr{C}$ we have*

$$|\mathscr{G}(f) - \mathscr{G}(g)| \leq |f - g|.$$

Proof Let $[f], [g] \in \mathscr{C}$ and $d = \alpha_{f-g} \in \mathbb{R}$ such that

$$|[f] - [g]| = |f - g + d|_0.$$

We denote $k = \alpha_{\mathscr{G}(f)-\mathscr{G}(g)}$ the value such that $|\mathscr{G}([f]) - \mathscr{G}([g])| = |\mathscr{G}([f]) - \mathscr{G}([g]) + k|_0$.

In order to estimate $|\mathscr{G}([f]) - \mathscr{G}([g])|$ consider $\mathscr{G}(f)(x) - \mathscr{G}(g)(x) =$

$$-c_f + \frac{1}{2}f(x) + \frac{1}{2}\max_{i\in\{1,2\}}[(A+f)(\tau_i(x))] + c_g - \frac{1}{2}g(x) - \frac{1}{2}\max_{i\in\{1,2\}}[(A+g)(\tau_i(x))],$$

which means $2\left(\mathscr{G}(f)(x) - \mathscr{G}(g)(x) + c_f - c_g\right) =$

$$f(x) - g(x) + \max_{i\in\{1,2\}}[(A+f)(\tau_i(x))] - \max_{i\in\{1,2\}}[(A+g)(\tau_i(x))].$$

We add d to both sides obtaining

$$2\left(\mathscr{G}(f)(x) - \mathscr{G}(g)(x) + c_f - c_g + d\right) =$$

$$f(x) - g(x) + d + \max_{i\in\{1,2\}}[(A+f+d)(\tau_i(x))] - \max_{i\in\{1,2\}}[(A+g)(\tau_i(x))],$$

which can be rewritten as $2\left(\mathscr{G}(f)(x) - \mathscr{G}(g)(x) + c_f - c_g + d\right) =$

$$(f(x) - g(x) + d) + \max_{i\in\{1,2\}}[(A+g+f-g+d)(\tau_i(x))] - \max_{i\in\{1,2\}}[(A+g)(\tau_i(x))].$$

We notice that $-|[f]-[g]| \leq f(y)-g(y)+d \leq |[f]-[g]|$ for any $y \in X$. By monotonicity of the supremum we get

$$-|[f]-[g]| + \max_{i\in\{1,2\}} [(A+g)(\tau_i(x))] \leq$$

$$\max_{i\in\{1,2\}} [(A+g+f-g+d)(\tau_i(x))] \leq |[f]-[g]| + \max_{i\in\{1,2\}} [(A+g)(\tau_i(x))],$$

which is equivalent to $-|[f]-[g]| \leq$

$$\max_{i\in\{1,2\}} [(A+g+f-g+d)(\tau_i(x))] - \max_{i\in\{1,2\}} [(A+g)(\tau_i(x))] \leq |[f]-[g]|,$$

thus

$$|\max_{i\in\{1,2\}} [(A+g+f-g+d)(\tau_i(x))] - \max_{i\in\{1,2\}} [(A+g)(\tau_i(x))]\,|_0 \leq |[f]-[g]|.$$

We assumed that $|f-g+d|_0 = |[f]-[g]|$. Therefore, using the two last inequalities we get $|2\left(\mathscr{G}(f)-\mathscr{G}(g)+c_f-c_g+d\right)|_0 \leq |[f]-[g]|+|[f]-[g]|$, which is equivalent to

$$|\mathscr{G}(f)-\mathscr{G}(g)+(c_f-c_g+d)|_0 \leq |[f]-[g]|. \tag{8}$$

We recall that $|\mathscr{G}([f])-\mathscr{G}([g])| =$

$$\min_{k\in\mathbb{R}} |\mathscr{G}(f)-\mathscr{G}(g)+k|_0 \leq |\mathscr{G}(f)-\mathscr{G}(g)+(c_f-c_g+d)|_0 \leq |[f]-[g]|,$$

and this finish the proof. □

3 Generic Properties

We will show a generic property for the iterative process acting on continuous functions for a given fixed Lipschitz potential A.

Definition 8 Consider the set $\mathfrak{A} \subset \mathscr{C} \times \mathscr{C}$ of pairs of functions (f_0, g_0), such that, if $|f_0-g_0| = |(f_0-g_0)+\alpha_{f_0-g_0)}|_0 = (f_0-g_0)(r)+\alpha_{f_0-g_0}$, for some r, then,

$$(f_0-g_0)(r) \neq (f_0-g_0)(\tau_1(r)) \text{ and } (f_0-g_0)(r) \neq (f_0-g_0)(\tau_2(r)).$$

Note that the above condition does not depends on the potential A. In the case $f_0(x)-g_0(x)+\alpha_{(f_0-g_0)}$ attains the supremum in a unique point then $(f_0, g_0) \in \mathfrak{A}$. Obviously, we could choose $\mathfrak{A} \subset \mathscr{C}$ the set of $h \in \mathscr{C}$ such that $h(r) \neq h(\tau_1(r))$ and $h(r) \neq h(\tau_2(r))$, but our choice $h = f_0 - g_0$ avoid this relabeling in the future.

We will show in Corollary 1 that the condition $(f, g) \in \mathfrak{A}$ is generic in $\mathscr{C} \times \mathscr{C}$.

Theorem 3 *Given the functions* $f_0, g_0 \in \mathscr{C}$, *assume* $(f_0, g_0) \in \mathfrak{A}$. *In this case, if* $|\mathscr{G}(f_0) - \mathscr{G}(g_0)| = |f_0 - g_0|$, *then,* $f_0 = g_0$.

Proof We denote by $d = \alpha_{f_0 - g_0}$ the value such that $|(f_0 - g_0) + d|_0 = |f_0 - g_0|$.

We denote by z_0 the point such that $|f_0 - g_0| = |f_0(z_0) - g_0(z_0) + d|$. Without loss of generality we assume that $f_0(z_0) - g_0(z_0) + d > 0$.

Note that $|(f_0 - g_0 + d)(z_0)|$ also maximizes

$$x \to |(f_0 - g_0 + d)(x)|. \tag{9}$$

Note that d was determined by the choice $(f_0 - g_0)$ (and, not $(g_0 - f_0)$).

We denote by $k = \alpha_{\mathscr{G}(f_0) - \mathscr{G}(g_0)}$ the value $|\mathscr{G}(f_0) - \mathscr{G}(g_0)| + k|_0 = |\mathscr{G}(f_0) - \mathscr{G}(g_0)|$.

Assuming $|\mathscr{G}(f_0) - \mathscr{G}(g_0)| = |f_0 - g_0|$, then, from (8) we get

$$|\mathscr{G}(f_0) - \mathscr{G}(g_0)| = |\mathscr{G}(f_0) - \mathscr{G}(g_0) + k|_0 \leq |\mathscr{G}(f_0) - \mathscr{G}(g_0) + (c_{f_0} - c_{g_0} + d)|_0 \leq |[f_0] - [g_0]|. \tag{10}$$

Therefore, k can be taken as $k = c_{f_0} - c_{g_0} + d$. Note that k was determined by d and the choice $(f_0 - g_0)$ (and, not $(g_0 - f_0)$).

We denote by z_1 a point such that $|\mathscr{G}(f_0) - \mathscr{G}(g_0)| = |\mathscr{G}(f_0)(z_1) - \mathscr{G}(g_0)(z_1) + k| = |f_1(z_1) - g_1(z_1) + k|$.

In the case $(f_1 - g_1)(z_1) + k \leq 0$ we know that there exists another point $\tilde{z_1}$, such that, $0 \leq (f_1 - g_1)(\tilde{z_1}) + k = |\mathscr{G}(f_0) - \mathscr{G}(g_0) + k|_0$.

Therefore, without loss of generality, we can always assume that it is true $(f_1 - g_1)(z_1) + k \geq 0$.

Assume that $(f_0, g_0) \in \mathfrak{A}$.

Under the above conditions in f_0, g_0, there exists z_0, z_1, $\bar{z} = \tau_{a_0^{z_1, f_0}}(z_1)$ and $\bar{w} = \tau_{a_0^{z_1, g_0}}(z_1)$ such that

$$(f_0 - g_0)(z_0) + d = |f_0 - g_0| = |\mathscr{G}(f_0) - \mathscr{G}(g_0)| = (f_1 - g_1)(z_1) + k =$$

$$[\frac{A(\bar{z})}{2} + \frac{1}{2}(f_0(\bar{z}) + f_0(z_1))] - [\frac{A(\bar{w})}{2} + \frac{1}{2}(g_0(\bar{w}) + g_0(z_1))] + k - c_f + c_g \leq$$

$$[\frac{A(\bar{z})}{2} + \frac{1}{2}(f_0(\bar{z}) + f_0(z_1))] - [\frac{A(\bar{z})}{2} + \frac{1}{2}(g_0(\bar{z}) + g_0(z_1))] + k - c_f + c_g =$$

$$[\frac{1}{2}(f_0(\bar{z}) + f_0(z_1))] - [\frac{1}{2}(g_0(\bar{z}) + g_0(z_1))] + k - c_f + c_g =$$

$$\frac{1}{2}(f_0(z_1) - g_0(z_1)) + \frac{1}{2}(f_0(\bar{z}) - g_0(\bar{z}) + k - c_f + c_g =$$

$$\frac{1}{2}(f_0(z_1) - g_0(z_1)) + \frac{1}{2}(f_0(\bar{z}) - g_0(\bar{z}) + d. \tag{11}$$

As $(f_0 - g_0 + d)(z_0) > 0$ is a supremum, it follows from the above that

$$(f_0 - g_0)(z_0) + d \leq \frac{1}{2}[(f_0 - g_0)(z_1) + d] + \frac{1}{2}[(f_0 - g_0)(\bar{z}) + d] \leq$$

$$\frac{1}{2}[(f_0 - g_0)(z_0) + d] + \frac{1}{2}[(f_0 - g_0)(z_0) + d] = (f_0 - g_0)(z_0) + d. \tag{12}$$

$(f_0 - g_0)(z_1) + d$ and $(f_0 - g_0)(\bar{z}) + d$ can not be both negative (because $(f_0 - g_0)(z_0) + d > 0$).

Both $(f_0 - g_0)(z_1) + d$ and $(f_0 - g_0)(\bar{z}) + d$ are positive. Otherwise, from (12) we get $(f_0 - g_0)(z_0) + d < \frac{1}{2}[\,(f_0 - g_0)(z_0) + d\,]$. This implies that $\frac{1}{2}[(f_0 - g_0)(z_1) + d] + \frac{1}{2}[(f_0 - g_0)(\bar{z}) + d] = (f_0 - g_0)(z_0) + d$.

Remember that $d = \alpha_{f_0 - g_0} = -\frac{\max(f_0 - g_0) + \min(f_0 - g_0)}{2}$. From 12 we get $(f_0 - g_0)(z_1) + d = (f_0 - g_0)(\bar{z}) + d = (f_0 - g_0)(z_0) + d$.

As $(f_0, g_0) \in \mathfrak{A}$ we get by Corollary 1 a contradiction. □

Remark 3 Given the point z_1 above (supremum of $x \to (f_1(x) - g_1(x)) + k$) we get from (11) that

$$(f_1 - g_1)(z_1) + k \leq \frac{1}{2}(f_0(z_1) - g_0(z_1) + d) + \frac{1}{2}(f_0(\tau_{a_0^{z_1, f_0}}(z_1)) - g_0(\tau_{a_0^{z_1, f_0}}(z_1) + d). \tag{13}$$

Note that if $f_0(z_1) - g_0(z_1) + d$ and $f_0(\tau_{a_0^{z_1, f_0}}(z_1)) - g_0(\tau_{a_0^{z_1, f_0}}(z_1) + d)$ have opposite signals, then we get a better rate

$$|\mathscr{G}(f_0) - \mathscr{G}(g_0)| \;=\; (f_1 - g_1)(z_1) + k \leq \frac{1}{2}|f_0 - g_0|. \tag{14}$$

During the iteration procedure this will happen from time to time for $f_n = \mathscr{G}^n(f_0)$ and $g_n = \mathscr{G}^n(u) = u$. This is a good explanation for the outstanding performance of the algorithm.

◇

Definition 9 Given a Lipschitz potential A with a unique subaction $u \in \mathscr{C}$ consider the set $\mathfrak{B} \subset \mathscr{C}$ of functions f_0, such that, if $|f_0 - u| = |(f_0 - u) + \alpha_{f_0 - u)}|_0 = (f_0 - u)(r) + \alpha_{f_0 - u}$, for some r, then,

$$(f_0 - u)(r) \neq (f_0 - u)(\tau_1(r)) \text{ and } (f_0 - u)(r) \neq (f_0 - u)(\tau_2(r)).$$

The set $\mathfrak{B}$ is dense in $\mathscr{C}$. The proof of this fact is basically the same as the proof that $\mathfrak{A}$ is dense on $\mathscr{C} \times \mathscr{C}$ and will be not presented.

In the same way as before one can show that:

Theorem 4 *Given the function $f_0 \in \mathscr{C}$, assume $f_0 \in \mathfrak{B}$. In this case, if $|\mathscr{G}(f_0) - u| = |f_0 - u|$, then, $f_0 = u$. This implies that if $f_0 \neq u$, then*

$$|\mathscr{G}(f_0) - u| < |f_0 - u|. \tag{15}$$

Therefore, if $\mathscr{G}^n(f_0) \in \mathfrak{B}$ and $\mathscr{G}^n(f_0) \neq u$, then

$$|\mathscr{G}^{n+1}(f_0) - u| < |\mathscr{G}^n(f_0) - u|. \tag{16}$$

Given an initial f_0 from time to time $\mathscr{G}^n(f_0) \in \mathfrak{B}$ for some n, and then the next iterate will experience a better approximation to the calibrated subaction u.

Now we will prove that $\mathfrak{A}$ is dense. We will need first to state some preliminary properties which will be used later. We recall that the norm in $\mathscr{C}$ is given by $|f| = \inf_{d \in R} |f + d|_0$ and the distance in $\mathscr{C} \times \mathscr{C}$ is the max distance $d((f, g), (f', g')) := \max(|f - f'|, |g - g'|)$ which is equivalent to the product topology. We will show now that the set $\mathfrak{A}$ is dense in $\mathscr{C} \times \mathscr{C}$ with respect to this topology.

Consider $X = [0, 1]$ and the maps $\tau_2(x) = \frac{1}{2}x$ and $\tau_2(x) = \frac{1}{2}(x + 1)$. Let $\mathscr{F} = \{(f, g) |\ f, g\ \in \mathscr{C}\} \subset \mathscr{C} \times \mathscr{C}$. Denote by β the map $\beta : X \times \mathscr{F} \to \mathbb{R}$ given by

$$\beta(x, f, g) = |f - g| - |f(x) - g(x)| + \min_{i \in \{0,1\}} \{|f - g| - |f(\tau_i(x)) - g(\tau_i(x))|\}.$$

We notice that $\beta(x, f, g) \geq 0$, and, moreover

- $\beta(x, f, g) = 0$, if and only if, $|f - g| = |f(x) - g(x)|$, and, $|f - g| = |f(\tau_1(x)) - g(\tau_1(x))|$ or $|f - g| = |f(\tau_2(x)) - g(\tau_2(x))|$;
- $\beta(x, f, g) > 0$, if and only if, one of the two conditions is true $|f - g| > |f(x) - g(x)|$, or, $|f - g| > |f(\tau_1(x)) - g(\tau_1(x))|$ and $|f - g| > |f(\tau_2(x)) - g(\tau_2(x))|$.

We define the set $\mathscr{O} \subset \mathscr{F}$ as being

$$\mathscr{O}_{\mathscr{F},\delta} = \{(f, g) \in \mathscr{F} | \beta(x, f, g) > 0,\ \forall x \in [\delta, 1 - \delta]\}.$$

If $d = -\frac{\max(f-g)+\min(f-g)}{2}$, then $|f - g| = |f - g + d|_0 = \frac{\max(f-g)-\min(f-g)}{2}$.

From the previous observation we conclude that for all $(f, g) \in \mathscr{O}_{\mathscr{F},\delta}$, if, x is such that $|f - g + d| = |f(x) - g(x) + d|$, then, $|f - g + d| \neq |f(\tau_1(x)) - g(\tau_1(x)) + d|$ and $|f - g + d| \neq |f(\tau_2(x)) - g(\tau_2(x)) + d|$.

To motivate our proof we are going to consider an explicit example where we made a perturbation of a pair $(f, g) \in \mathscr{C}$, but $\beta(x, f, g) = 0$, for some x.

Fig. 4 $f(x)$ of Example 2

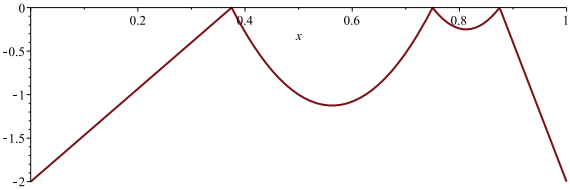
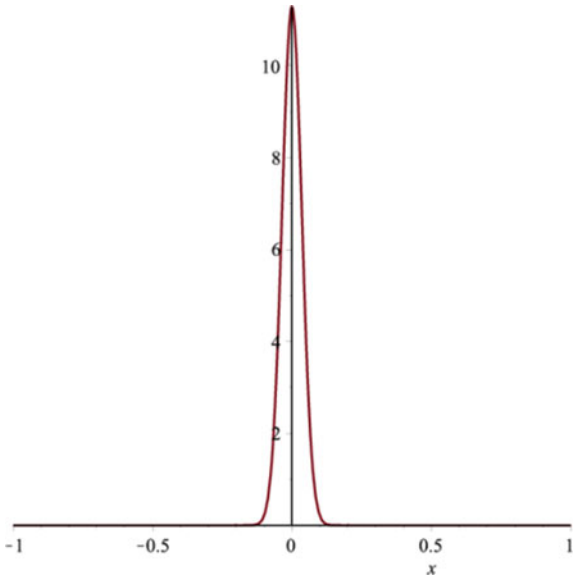


Fig. 5 u_ε



Example 2 Consider $(f, g) \in \mathscr{C}$ where

$$f(x) = \begin{cases} \frac{16}{3}x - 2 & 0 \leq x \text{ and } x < 3/8 \\ 32x^2 - 36x + 9 & 3/8 \leq x \text{ and } x < 3/4 \\ 64x^2 - 104x + 42 & 3/4 \leq x \text{ and } x \leq 7/8 \\ -16x + 14 & 7/8 \leq x \text{ and } x \leq 1. \end{cases}$$

and $g(x) = 0$.

It is easy to see that for $x = 3/4$ we have $|f - 0| = |f - 0 + 1| = f(3/4) + 1 = 1$, $f(\tau_2(3/4)) + 1 = 1$ and $f(\tau_1(3/4)) + 1 = 1$, (see Fig. 4) thus, $\beta(3/4, f, 0) = 0$, meaning that $(f, 0) \notin \mathscr{O}_{\mathscr{F}, \frac{1}{4}}$. The same is true for $x = 0$.

In order to obtain the perturbation $(f_\varepsilon, g_\varepsilon)$ we consider an ε-concentrated approximation via Dirac function $u_\varepsilon(x) := \frac{1}{\varepsilon\sqrt{\pi}} e^{-\frac{x^2}{\varepsilon^2}}$ (see Fig. 5) and we also we define for $\varepsilon = 0.005$ the modifications (see Fig. 6).

$$Q_\varepsilon(x) := \frac{1}{500} u_\varepsilon(x - (3/4 - 0.015)) \text{ and } W_\varepsilon(x) := -\frac{1}{1000} u_\varepsilon(x - (0 + 0.015)) :$$

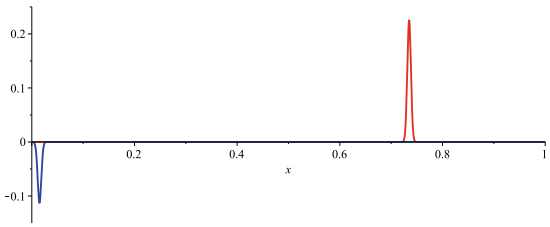


Fig. 6 Q_ε (red) and W_ε(blue)

We set $f_\varepsilon(x) = f(x) + Q_\varepsilon(x) + W_\varepsilon(x)$ and $g_\varepsilon(x) = g(x)$. In this case $|f_\varepsilon - f| = |-Q_\varepsilon - W_\varepsilon| = (0.113 - (-0.226))/2 = 0.1695$ as we can see by the picture (see Fig. 7).

As we can see, after the perturbation the maximum value is attained only for $x_0 = \frac{3}{4} - 0.015$ and for $x_1 = 0 + 0.015$ and neither of them are pre-image one of each other. Therefore, $(f_\varepsilon, g_\varepsilon) \in \mathscr{O}_{\mathscr{F},0}$ (see Fig. 14).

Theorem 5 *Let $\Lambda \subset \mathscr{F}$ a compact subset. Then the set $\mathscr{O}_{\Lambda,\delta}$ is an open and dense set. In particular, $\mathscr{O}_\Lambda := \bigcap_{n>2} \mathscr{O}_{\Lambda,\frac{1}{n}}$ is a dense set.*

As a consequence, taking $\mathfrak{A} = \mathscr{O}_\Lambda$, it will follow:

Corollary 1 *The set $\mathfrak{A}$ is dense. More precisely, if $|f - g| = |f - g + d| = f(x_0) - g(x_0) + d$, then $f(\tau_i(x_0)) - g(\tau_i(x_0)) + d \neq f(x_0) - g(x_0) + d$, for $i = 1, 2$.*

Proof The first step in the proof of Theorem 5 is the openness of $\mathscr{O}_{\Lambda,\frac{1}{n}}$.

In this direction we observe that β is continuous because the min operation and the sup-norm are continuous. Taking $(f_0, g_0) \in \mathscr{O}_{\Lambda,\frac{1}{n}}$ we obtain $\beta(x, f_0, g_0) > 0,\ \forall x \in \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$, as we can see in the Fig. 9.

Using the compactness and the continuity we can take $\alpha > 0$, such that, $\beta(x, f_0, g_0) > \alpha,\ \forall x \in [\frac{1}{n}, 1 - \frac{1}{n}]$. Therefore, if $(f, g) \in \mathscr{U}$, where $\mathscr{U}$ is an open neighborhood of (f_0, g_0), we get

$$\beta(x, f, g) - \frac{\alpha}{2} = \beta(x, f, g) - \beta(x, f_0, g_0) + \beta(x, f_0, g_0) - \alpha + \alpha - \frac{\alpha}{2} \geq$$

$$\leq \beta(x, f_0, g_0) - \beta(x, f, g) + \beta(x, f, g) - \alpha + \alpha - \frac{\alpha}{2} > -\varepsilon_x + 0 + \frac{\alpha}{2} > 0,$$

if we choose $\varepsilon_x < \frac{\alpha}{2}$, where ε_x is the continuity constant for the map $(f, g) \to \beta(x, f, g)$, for a fixed $x \in \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$.

Since the interval $\left[\frac{1}{n}, 1 - \frac{1}{n}\right]$ is compact we can take $0 < \varepsilon \leq \varepsilon_x,\ \forall x \in \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$ (Fig. 8).

This proves that the set

$$\mathscr{U}_\delta :=$$

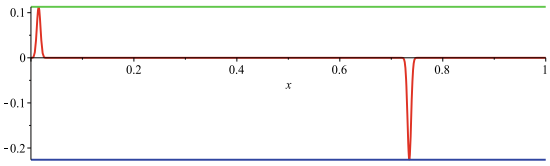


Fig. 7 *Calculating* $|-Q_\varepsilon - W_\varepsilon|$

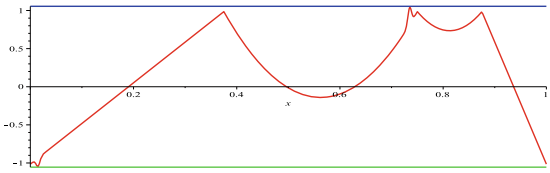


Fig. 8 $|f_\varepsilon - g_\varepsilon| = |f_\varepsilon - g_\varepsilon - (-0.985)|_0 = (0.07 - (-2.04))/2 = 1.055$

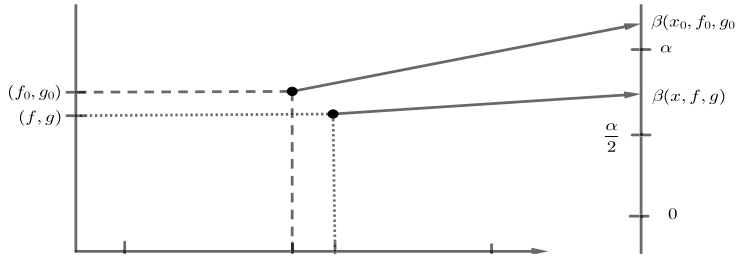


Fig. 9 Approximating (x_0, f_0, g_0)

$$\{(f,g) \mid \text{ if } d((f,g),(f_0,g_0))<\delta, \text{ then } |\beta(x,f,g)-\beta(x,f_0,g_0)|<\varepsilon, \forall x\in\left[\tfrac{1}{n},1-\tfrac{1}{n}\right]\}$$

is an open neighborhood of (f_0, g_0) in $\mathscr{O}_{\Lambda,\frac{1}{n}}$.

In order to prove the density of $\mathscr{O}_{\Lambda,\frac{1}{n}}$ we observe that if $x_0 \in \left[\frac{1}{n}, 1-\frac{1}{n}\right]$, then $\frac{1}{2n} + \frac{i}{2} - x_0 \leq \tau_i(x_0) - x_0 \leq \frac{i}{2} + \frac{1}{2} - \frac{1}{2n} - x_0$. Thus $|\tau_i(x_0) - x_0| \geq \frac{1}{2n}$ for all $x_0 \in [\frac{1}{n}, 1-\frac{1}{n}]$.

Using this estimate we can apply an ε-concentrated perturbations with $\varepsilon < \frac{1}{2n}$ (see Example 2 for a constructive approach) obtaining a pair $(f_\varepsilon, g_\varepsilon)$, in such way that, $g_\varepsilon = g, x_0$ and x_1 are the only points where $|f_\varepsilon - g_\varepsilon| = |f_\varepsilon(x_0) - g(x_0) + d| = |f_\varepsilon(x_1) - g(x_1) + d|$ and $x_0 \neq \tau_0(x_1), \tau_2(x_1), x_1 \neq \tau_0(x_0), \tau_2(x_0)$.

In particular $\beta(x, f_\varepsilon, g) > 0$, for any $x \in \left[\frac{1}{n}, 1-\frac{1}{n}\right]$, which means that $(f_\varepsilon, g_\varepsilon) \in \mathscr{O}_{\Lambda,\frac{1}{n}}$.

4 Perturbation Theory: Close by the Fixed Point

In this section we analyze the question: when the calibrated subaction is unique is there an uniform exponential speed of approximation of the iteration $\mathscr{G}^n(f_0)$ to the subaction? The question makes sense close by the subaction u. The answer is no. We will proceed a careful analysis of the action of $\mathscr{G}$ close by the fixed point $u \in \mathscr{C}$.

Section 4 is about the possibility of change a given point, in the neighborhood of a subaction, by a close one having different properties, with respect to the convergence rate of the operator $\mathscr{G}$. It can't be used for genericity, as far as we know because we say nothing close to other points in the space.

In some examples we may consider a different dynamical system on $X = [0, 1]$ given by the maps $\tau_i(x) = \frac{1}{2}(i + 1 - x)$, for $i = 0, 1$, which are the inverse branches of $T(x) = -2x \mod 1$.

Our main task is to evaluate the effect of a perturbation on the nonlinear operator ψ defined by

$$\psi(f)(x) = \max_{T(y)=x} (A + f)(y) = \max_{i=0,1}(A + f)(\tau_i(x))$$

for a fixed potential $A \in \mathscr{C}_k$ (Fig. 10).

The operator $H := H_A$ given by

$$H(f)(x) := \frac{1}{2} f(x) + \frac{1}{2}\psi(f)(x),$$

Note that $\mathscr{G} := \mathscr{G}_A$ is a normalized version of H

$$\mathscr{G}(f)(x) := H(f)(x) - \sup_{x\in X} H(f)(x).$$

It is usual to denote $c_f := \sup_{x\in X} H(f)(x)$ then $H(f)(x) = \mathscr{G}(f) + c_f$ (normalization means that $\sup_{x\in X} \mathscr{G}(f)(x) = 0$). Sometimes it is useful to look at the operator $H - Id$ given by $(H - Id)(f) := \frac{1}{2}\psi(f)(x) - \frac{1}{2} f(x)$.

We assume that there exists **a unique function** $u \in \mathscr{C}_K$ such that $\mathscr{G}(u) = u$ (this is true if the maximizing probability is unique). Thus, $H(u)(x) = \mathscr{G}(u) + c_u = u(x) + c_u$ where $c_u := \sup_{x\in X} H(u)(x)$.

The above equation is equivalent to $u(x) + c_u = \frac{1}{2}u(x) + \frac{1}{2}\psi(f)(x)$ which is equivalent to the sub-action equation

$$u(x) = \max_{i=0,1}(A - 2c_u + u)(\tau_i(x)).$$

We can assume that $m_A = 2c_u = 0$ (by adding a constant to A) and then, $H(u) = u$. It is useful to observe that under this assumption we also get $\psi(u) = u$.

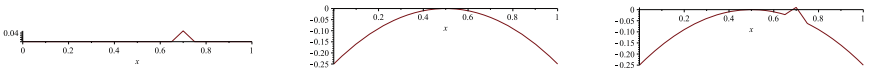


Fig. 10 The graph of the function $\alpha_{0.05,0.7}$ in the left side, $f(x) = -(x - 1/2)^2$ in the center and $f_{0.05} = f(x) + \alpha_{0.05,0.7}(x)$ in the right side

We start with a local perturbation lemma.

Let $\alpha_{\varepsilon,a} : X \to \mathbb{R}$ be a piecewise linear bump function defined by

$$\alpha_{\varepsilon,a}(x) = \begin{cases} 0, & 0 \le x \le a - \varepsilon \\ kx - k(a - \varepsilon), & a - \varepsilon \le x \le a \\ -kx + k(a + \varepsilon), & a \le x \le a + \varepsilon \\ 0, & a + \varepsilon \le x \le 1, \end{cases}$$

where $a \in (0, 1)$ and $\varepsilon > 0$ is arbitrary small.

Lemma 1 *If $f \in \mathscr{C}_K$, then $f_\varepsilon = f(x) + \alpha_{\varepsilon,a}(x) \in \mathscr{C}_K$. Moreover, $f_\varepsilon(x) \ge f(x)$ and $f_\varepsilon(x) = f(x)$ outside of the interval $[a - \varepsilon,\ a + \varepsilon]$. Finally, $|f_\varepsilon - f| = \frac{k\varepsilon}{2}$.*

Proof The proof is straightforward because $|f_\varepsilon - f| = |\alpha_{\varepsilon,a}|$ and $0 \le \alpha_{\varepsilon,a}(x) \le k\varepsilon$.

We will make the perturbations by choosing a fixed point $x_0 \ne 0, 1, 1/2$ in X and $\varepsilon > 0$, such that, the intervals $I = [x_0 - \varepsilon,\ x_0 + \varepsilon]$ and $T(I) = [T(x_0) - 2\varepsilon,\ T(x_0) + 2\varepsilon]$ are disjoint. Then, we take $f_\varepsilon = f(x) + \alpha_{\varepsilon,a}(x)$ and we will try to estimate $\psi(f_\varepsilon)$.

Lemma 2 *$\psi(f_\varepsilon) = \psi(f)$ outside of $T(I)$.*

Proof We notice that A remains unchanged and $[T(x_0) - 2\varepsilon,\ T(x_0) + 2\varepsilon] = T([x_0 - \varepsilon,\ x_0 + \varepsilon])$. Therefore, for any y such that $T(y) = x$ we can not have $y \in [x_0 - \varepsilon,\ x_0 + \varepsilon]$. Thus, $f_\varepsilon(y) = f(y)$, proving that $\psi(f_\varepsilon) = \psi(f)$.

Another question is about what happens in $T(I)$. For any x in this interval one of its pre-images y belongs to I therefore $f_\varepsilon(y) \ge f(y)$. Thus, ψ may change.

We recall that **a turning point** x (see also [23, 24]) is a point where $(A + f)(\tau_1(x)) = (A + f)(\tau_2(x))$. If x is not a turning point then there exists **a dominant realizer**, that is, $(A + f)(\tau_1(x)) > (A + f)(\tau_2(x))$, or, $(A + f)(\tau_1(x)) < (A + f)(\tau_2(x))$ (Fig. 11).

Lemma 3 *Suppose that x_0 is such that $T(x_0)$ is not a turning point and j is the dominant symbol. Let $i \in \{0,\ 1\}$ be such that $\tau_i(T(x_0)) = x_0$. We have two possible cases:*

- *If $j \ne i$, then $\psi(f_\varepsilon)(x) = \psi(f)(x)$, for any $x \in T(I)$.*
- *If $j = i$, then $\psi(f_\varepsilon)(x) = \psi(f)(x) + \alpha_{\varepsilon,x_0}(\tau_j(x)) \ge \psi(f)(x)$, for any $x \in T(I)$ and $|\psi(f_\varepsilon)(x) - \psi(f)(x)| = \frac{k\varepsilon}{2}$.*

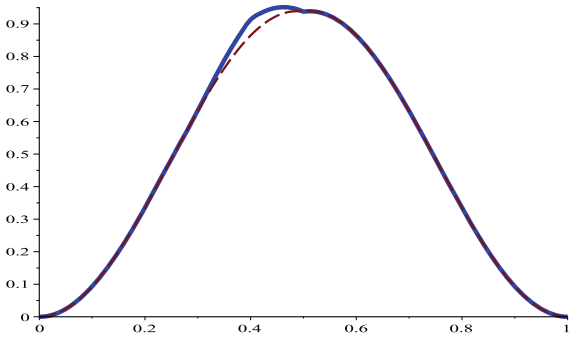


Fig. 11 The graph of the functions $\psi(f_\varepsilon)$ (blue line) and $\psi(f)$ (traced line) where, $A(x) = \sin^2(2\pi x)$, $f(x) = -(x - 1/2)^2$ and $f_{0.1} = f(x) + \alpha_{0.1,0.7}(x)$. The difference occurs only in the interval $T(I) = [0.2,\ 0.6]$ because $T(0.7) = 0.4$ and $I = [0.6,\ 0.8]$

Proof In the first case, in order to fix ideas we suppose, without lost of generality, $j = 0$ and $i = 1$, then $\tau_2(T(x_0)) = x_0$ and $(A + f)(\tau_1(T(x_0))) > (A + f)(\tau_2(T(x_0)))$. By the continuity of $A + f$ we can choose $\varepsilon > 0$ small enough in order to have $(A + f_\varepsilon)(\tau_1(x)) > (A + f_\varepsilon)(\tau_2(x))$, for all $x \in T(I)$. Therefore, $\psi(f_\varepsilon)(x) = (A + f_\varepsilon)(\tau_1(x)) = (A + f)(\tau_1(x)) = \psi(f)(x)$, for any $x \in T(I)$.
In the second case, $\tau_j(T(x_0)) = x_0$ and $(A + f)(\tau_j(T(x_0))) > (A + f)(\tau_i(T(x_0)))$. Once more we use the continuity of $A + f$ to choose $\varepsilon > 0$ small enough in order to have $(A + f_\varepsilon)(\tau_j(x)) > (A + f_\varepsilon)(\tau_i(x))$, for all $x \in T(I)$. Therefore, $\psi(f_\varepsilon)(x) = (A + f_\varepsilon)(\tau_j(x)) = (A + f)(\tau_j(x)) + \alpha_{\varepsilon,x_0}(\tau_j(x)) = \psi(f)(x) + \alpha_{\varepsilon,x_0}(\tau_j(x))$, for any $x \in T(I)$.

Our first task is to compare $H(f)$ and $H(f_\varepsilon)$. We can always assume that $T(I)$ and I are disjoint so the perturbation $f \to f_\varepsilon$ acts separately in each one as described by the previous lemmas.

Lemma 4 *Let f_ε a perturbation of f and x_0 such that is not a pre-image of a turning point (with respect to f). Then, $H(f)(x) \leq H(f_\varepsilon)(x)$, with equality only outside of $[T(x_0) - 2\varepsilon,\ T(x_0) + 2\varepsilon] \cup [x_0 - \varepsilon,\ x_0 + \varepsilon]$. Moreover, $H(f_\varepsilon)(x) - H(f)(x) \leq \frac{k\varepsilon}{2}$. (We can prove similar results for $(H - Id)$.)*

The proof is a direct consequence of the previous lemmas.

We want to study the relation between $|\mathscr{G}(f) - u|$ and $|f - u|$. We also want to see what happens when we make a perturbation $f \to f_\varepsilon$.

We start by choosing $d = \alpha_{f-u}$ such that $\delta = |f - u| = |f - u + d|_0$, then

$$-\delta \leq f(y) - u(y) + d \leq \delta,$$

for all $y \in X$. Multiplying the above by $1/2$ we conclude that (Fig. 12)

$$-\frac{\delta}{2} \leq \frac{1}{2}(f(y) - u(y)) + \frac{d}{2} \leq \frac{\delta}{2}.$$

Adding $A(y)$ we obtain the inequalities

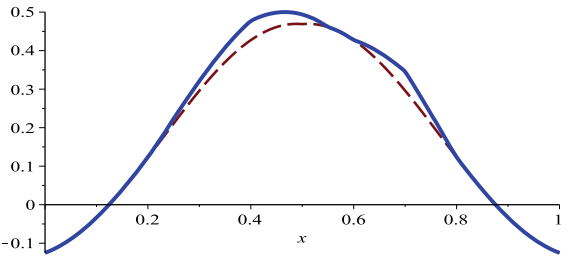


Fig. 12 The graph of the functions $H(f_\varepsilon)$ (blue line) and $H(f)$ (traced line) where, $A(x) = \sin^2(2\pi x)$, $f(x) = -(x - 1/2)^2$ and $f_{0.1} = f(x) + \alpha_{0.1,0.7}(x)$. The difference occurs only in the interval $[0.2,\ 0.6] \cup [0.6,\ 0.8]$ because $T(0.7) = 0.4$

$$-\delta \leq A(y) + f(y) - (A(y) + u(y)) + d \leq \delta,$$

and

$$-\delta + (A(y) + u(y)) \leq A(y) + f(y) + d \leq \delta + (A(y) + u(y)).$$

Taking the supremum in y, such that, $T(y) = x$, we get $-\delta + \psi(u)(x) \leq \psi(f)(x) + d \leq \delta + \psi(u)(x)$. Multiplying by $1/2$ we conclude that

$$-\frac{\delta}{2} \leq \frac{1}{2}(\psi(f)(x) - \psi(u)(x)) + \frac{d}{2} \leq \frac{\delta}{2}.$$

Note that

$$\mathscr{G}(f)(x) - u(x) + d = \mathscr{G}(f)(x) - \mathscr{G}(u)(x) + d =$$

$$= \frac{1}{2}(f(x) - u(x)) + \frac{1}{2}(\psi(f)(x) - \psi(u)(x)) - c_f + c_u + d =$$

$$\frac{1}{2}(f(x) - u(x) + d) + \frac{1}{2}(\psi(f)(x) - \psi(u)(x) + d) - c_f.$$

Using the inequalities

$$-\frac{\delta}{2} \leq \frac{1}{2}(\psi(f)(x) - \psi(u)(x)) + \frac{d}{2} \leq \frac{\delta}{2},$$

$$-\frac{\delta}{2} \leq \frac{1}{2}(f(y) - u(y)) + \frac{d}{2} \leq \frac{\delta}{2},$$

and, the fact that $c_u = 0$, we finally obtain

$$-\frac{\delta}{2} - \frac{\delta}{2} - c_f \leq \mathscr{G}(f)(x) - u(x) + d \leq \frac{\delta}{2} + \frac{\delta}{2} - c_f,$$

and,

$$-\delta \leq \mathscr{G}(f)(x) - u(x) + (d + c_f) \leq \delta.$$

Therefore,

$$|\mathscr{G}(f)(x) - u(x) + (d + c_f)| \leq \delta = |f - u|,$$

for all $x \in X$.

From this fundamental inequality we get a very important result about the operator $\mathscr{G}$.

We recall that $|\mathscr{G}(f)(x) - u(x)| = \min_\gamma |\mathscr{G}(f) - u + \gamma|_0 \leq |\mathscr{G}(f) - u + (d + c_f)|_0 = \sup_{x\in X} |\mathscr{G}(f)(x) - u(x) + (d + c_f)| \leq |f - u|$.

Theorem 6 *Let $\mathscr{G}$ be the operator associated to A and u the fixed point ($\mathscr{G}(u)(x) = u(x)$), then,*

(a) The contraction rate is controlled by $H - Id$;
(b) $|H(f) - f|_0 \leq 2|f - u|$;
(c) If $|H(f) - f|_0 = \beta$, then $|\mathscr{G}(f)(x) - u(x) + (d + c_f)|_0 \geq |f - u| - \beta$.

Proof (a) We recall that $\mathscr{G}(f)(x) + c_f = H(f)$, thus,

$$|\mathscr{G}(f)(x) - u(x) + (d + c_f)| \leq |f - u|$$

$$|\mathscr{G}(f)(x) + c_f - f(x) + f(x) - u(x) + d| \leq \sup_{x\in X} |f(x) - u(x) + d|$$

$$|[H(f) - f(x)] + f(x) - u(x) + d| \leq \sup_{x\in X} |f(x) - u(x) + d|$$

$$\sup_{x\in X} |[H(f) - f(x)] + f(x) - u(x) + d| \leq \sup_{x\in X} |f(x) - u(x) + d|.$$

(b) Here we use the triangular inequality

$$|H(f) - f(x)| \leq |[H(f) - f(x)] + f(x) - u(x) + d| + |f(x) - u(x) + d| \leq 2|f - u|.$$

(c) Using the triangular inequality we obtain

$$|f - u| = |f - u + d|_0 \leq |f - u + d + \mathscr{G}(f)(x) + c_f - f(x) - (\mathscr{G}(f)(x) + c_f - f(x))|_0 \leq$$

$$\leq |\mathscr{G}(f)(x) + c_f - f(x) + f - u + d|_0 + |\mathscr{G}(f)(x) + c_f - f(x)|_0 =$$

$$= |\mathscr{G}(f)(x) - u + (d + c_f)|_0 + |H(f)(x) - f(x)|_0 = |\mathscr{G}(f)(x) - u + (d + c_f)|_0 + \beta,$$

or, equivalently,

$$|\mathscr{G}(f)(x) - u + (d + c_f)|_0 \geq |f - u| - \beta.$$

We are dealing with a kind of technical problem: $|p(x) + q(x)| \leq |q|_0$, $\forall x \in X$, where $\max q = -\min q$. In our case, $p(x) = H(f) - f(x)$ and $q(x) = f(x) -$

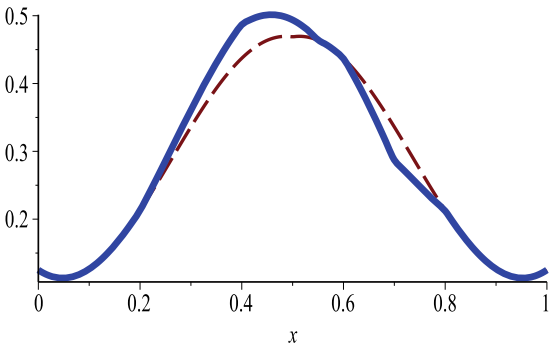


Fig. 13 Functions $(H - Id)(f_\varepsilon)$ (blue line) and $(H - Id)(f)$ (traced line) where, $A(x) = \sin^2(2\pi x)$, $f(x) = -(x - 1/2)^2$ and $f_{0.1} = f(x) + \alpha_{0.1,0.7}(x)$. The difference occurs only in the interval $[0.2,\ 0.6]$, where the perturbation is bigger, and, the interval $[0.6,\ 0.8]$, where the perturbation is smaller, because $T(0.7) = 0.4$

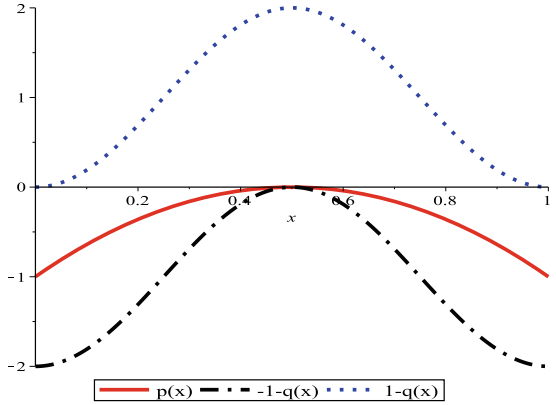


Fig. 14 Functions $-1 - q(x)$ and $1 - q(x)$

$u(x) + d$ are continuous functions. The first observation is that $|p(x) + q(x)| \leq |q|_0,\ \forall x \in X$, is equivalent to $-|q|_0 - q(x) \leq p(x) \leq |q|_0 - q(x)$. From this we can get interesting examples (Fig. 13).

Example 3 Consider $p(x) = -4\ (x - 1/2)^2$ and $q(x) = \cos(2\,\pi\,x)$. It is easy to see that $|q|_0 = \max q = -\min q = 1$ and the inequality $-1 - q(x) \leq p(x) \leq 1 - q(x)$ is described in the Fig. 14.

A simple calculation shows that $|p + q|_0 = 1 = |q|_0$, but $|p + q| = |p + q + 0.414|_0 = 0.586$.

The property $\max q = -\min q$ means that $|q| = |q + 0|_0$, therefore, $|p + q| = 0.586 < 1 = |q|$.

Lemma 5 *Consider* $|p(x) + q(x)| \leq |q|_0,\ \forall x \in X$, *with* $\max q = -\min q$. *Then, there exists* $z \in X$, *such that,* $p(z) = 0$. *In particular, taking* $p(x) = H(f) - f(x)$ *and* $q(x) = f(x) - u(x) + d$, *we have*

$$f(z) = \max_{T(y)=z} A(y) + f(y).$$

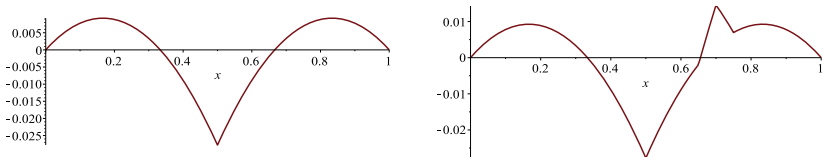


Fig. 15 In the left side the graph of u and in the right side the graph of f_ε

Proof We already know that there exists x_0 such that $|q|_0 = q(x_0)$, therefore, $p(x_0) + q(x_0) \leq |q|_0 = q(x_0)$, or, equivalently, $p(x_0) \leq 0$. Analogously, there exists x_1 such that $|q|_0 = -q(x_1)$ and $p(x_1) \geq 0$. Unless $q = cte$ we can always suppose that $x_0 \neq x_1$. If $p(x_0) = 0$ or $p(x_1) = 0$ the problem is solved. Otherwise, if $p(x_0) < 0$ and $p(x_1) > 0$ the intermediate value theorem for continuous functions claims that there exists $z \in [x_0,\ x_1]$, such that, $p(z) = 0$.

Note that for $p(x) = H(f) - f(x)$, the equation $p(z) = 0$ is equivalent to $f(z) = \max_{T(y)=z} A(y) + f(y)$.

The behaviour of $|\mathscr{G}(f)(x) - u + (d + c_f)|_0$ may be very different from $|\mathscr{G}(f) - u|$. On the one hand $|\mathscr{G}(f) - u| \leq |\mathscr{G}(f) - u + (d + c_f)|_0 \leq |f - u|$ and on the other hand we can find f arbitrarily close to u, such that, $|\mathscr{G}(f) - u| = \frac{1}{4} \leq |f - u|$.

Lemma 6 *Let u be the only sub-action of A ($m_A = 0$). Let $f_\varepsilon = u + \alpha_{\varepsilon,x_0}$ a perturbation of f and take x_0 not a pre-image of a turning point (with respect to f). Then, $|\mathscr{G}(f_\varepsilon) - u| = \frac{1}{2}|f_\varepsilon - u|$ and $|f_\varepsilon - u| = \frac{k\varepsilon}{2}$.*

Proof First, we observe that $|f_\varepsilon - u| = |\alpha_{\varepsilon,x_0}| = \frac{\max \alpha_{\varepsilon,x_0} - \min \alpha_{\varepsilon,x_0}}{2} = \frac{k\varepsilon - 0}{2} = \frac{k\varepsilon}{2}$.

Rewriting $|\mathscr{G}(f_\varepsilon) - u|$ we obtain

$$|\mathscr{G}(f_\varepsilon) - u| = |H(f_\varepsilon) - c_{f_\varepsilon} - u| = |H(f_\varepsilon) - u| = |\frac{1}{2} f_\varepsilon + \frac{1}{2}\psi(f_\varepsilon) - u| =$$

$$= |\frac{1}{2}(u + \alpha_{\varepsilon,x_0}) + \frac{1}{2}\psi(f_\varepsilon) - \psi(u)| = |\frac{1}{2}\alpha_{\varepsilon,x_0} + \frac{1}{2}(\psi(f_\varepsilon) - \psi(u))|.$$

The function α_{ε,x_0} is zero outside of the set $[x_0 - \varepsilon,\ x_0 + \varepsilon]$, and, $\psi(f_\varepsilon) - \psi(u) = 0$ outside of the set $[T(x_0) - 2\varepsilon,\ T(x_0) + 2\varepsilon]$ by Lemma 2.

Therefore, the $\min \frac{1}{2}\alpha_{\varepsilon,x_0} + \frac{1}{2}(\psi(f_\varepsilon) - \psi(u)) = 0$, and, $\max \frac{1}{2}\alpha_{\varepsilon,x_0} + \frac{1}{2}(\psi(f_\varepsilon) - \psi(u)) = \frac{k\varepsilon}{2}$. By definition $|\mathscr{G}(f_\varepsilon)(x) - u| = \frac{k\varepsilon}{4}$.

Example 4 Consider the dynamics $T(x) = -2x \mod 1$.

Let $A(x) = -(x - \frac{1}{2})^2 + \frac{1}{36}$ be the potential and u the subaction (see Figs. 15, 16 and 17)

$$u(x) = \begin{cases} -1/3\,x^2 + x/9, & 0 \leq x \leq 1/2 \\ -1/3\,x^2 + 5/9\,x - 2/9, & 1/2 \leq x \leq 1. \end{cases}$$

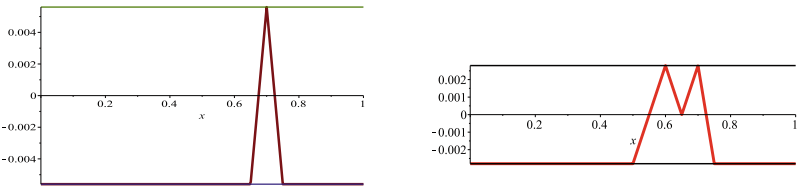


Fig. 16 In the left the graph of $f_\varepsilon(x) - u(x) - \frac{k\varepsilon}{2}$ and in the right the one for $1/2\, f_\varepsilon(x) + 1/2\, \psi(f_\varepsilon)(x) - u(x) - \frac{k\varepsilon}{4}$

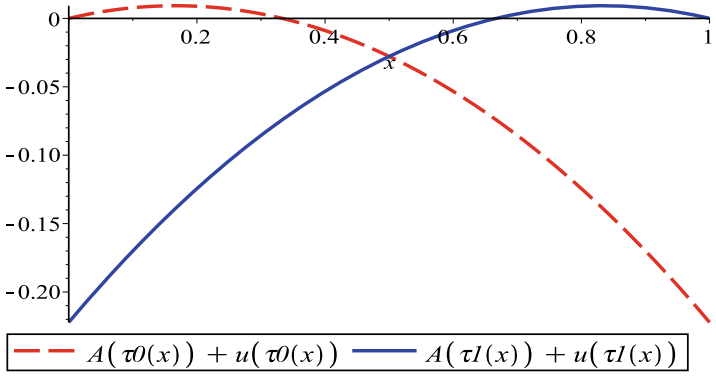


Fig. 17 The graph of the functions $(A + u)(\tau_1(x))$ and $(A + u)(\tau_2(x))$

From the graph of u we see that $x = \frac{1}{2}$ is the only turning point. Therefore, we can take $x_0 = 0.7$, $\varepsilon = 0.05$ and $f_\varepsilon = u + \alpha_{0.05,0.7}$. We also know that $Lip(A) = 1$ and $Lip(u) = \frac{2}{9}$, thus, we can take $k = \frac{2}{9}$.

As predicted $|\mathscr{G}(f_\varepsilon)(x) - u| = \frac{k\varepsilon}{4} = 0.0028$ and $|f_\varepsilon - u| = \frac{k\varepsilon}{2} = 0.0056$.

From Lemma 6 we get

Corollary 2 *For any $\varepsilon > 0$ there exists a function f which is ε-close to u, such that, $\mathscr{G}$ contracts by $1/2$ in f, that is, $|\mathscr{G}(f) - u| = \frac{1}{2}|f - u|$.*

We may ask if there exists some neighborhood of u where $|\mathscr{G}(f_\varepsilon)(x) - u| \leq (1 - \delta)|f_\varepsilon - u|$. The answer is no. Actually, it is the opposite of that. We can exhibit a sequence $f_\varepsilon \to u$, and, $|\mathscr{G}(f_\varepsilon)(x) - u| = |f_\varepsilon - u|$.

Example 5 We will show an example where $|\mathscr{G}(f_\varepsilon) - \mathscr{G}(u)| = |f_\varepsilon - u|$, $\epsilon > 0$, for f_ε as close as you want to the calibrated subaction u.

Consider again the dynamics $T(x) = -2x \pmod 1$. Let $A(x) = -(x - \frac{1}{2})^2 + \frac{1}{36}$ be the potential and u the subaction

$$u(x) = \begin{cases} -1/3\, x^2 + x/9, & 0 \leq x \leq 1/2 \\ -1/3\, x^2 + 5/9\, x - 2/9, & 1/2 \leq x \leq 1. \end{cases}$$

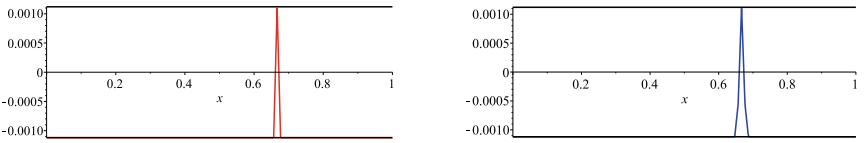


Fig. 18 In the left side the graph of $f_\varepsilon(x) - u(x) - \frac{k\varepsilon}{2}$ and in the right side the graph of $1/2\, f_\varepsilon(x) + 1/2\, \psi(f_\varepsilon)(x) - u(x) - \frac{k\varepsilon}{2}$

We fix $x_0 = \frac{2}{3}$. The function α_{ε,x_0} is zero outside of $I = [\frac{2}{3} - \varepsilon,\ \frac{2}{3} + \varepsilon]$ and $\psi(f_\varepsilon) - \psi(u) = 0$ outside of $T(I)$ by Lemma 2.

We know that $T(\frac{1}{3}) = \frac{1}{3}$ and $T(\frac{2}{3}) = \frac{2}{3}$. As we can see in the Fig. 17, $\{0,\ \frac{1}{2},\ 1\}$, are the only turning points and the dominant symbol in $x_0 = 2/3$ is $j = 1$. Also, $\tau_2(T(\frac{2}{3})) = \frac{2}{3}$, and thus $i = 1 = j$.

Once more

$$|\mathscr{G}(f_\varepsilon) - u| = \left|\frac{1}{2}\alpha_{\varepsilon,x_0} + \frac{1}{2}(\psi(f_\varepsilon) - \psi(u))\right|.$$

Since $I \subset T(I)$, we get, by Lemma 3, that α_{ε,x_0} attains the value $k\varepsilon$ and $\psi(f_\varepsilon)(x) - \psi(u)(x) = \alpha_{\varepsilon,x_0}(\tau_2(x))$ attains the value $k\varepsilon$ at least in x_0. Thus, $\frac{1}{2}\alpha_{\varepsilon,x_0} + \frac{1}{2}(\psi(f_\varepsilon) - \psi(u))$ attains the value $\frac{k\varepsilon}{2} = |f_\varepsilon - u|$ (see Fig. 18 for $\varepsilon = 0.01$ and $x_0 = \frac{2}{3}$).

Therefore, $|\mathscr{G}(f_\varepsilon) - u| \geq |f_\varepsilon - u|$. This work was part of the MSc dissertation of H. H. Ferreira [14].

References

1. M. Bachar and M. Khamsi, Recent contributions to fixed point theory of monotone mappings. J. Fixed Point Theory Appl. 19, no. 3, 1953–1976 (2017)
2. Baraviera, A., Leplaideur, R., Lopes, A.O.: Ergodic Optimization, zero temperature and the Max-Plus algebra, 29º Coloquio Brasileiro de Matematica. IMPA, Rio de Janeiro (2013)
3. Baraviera, A.T., Lopes, A.O., Mengue, J.: On the selection of subaction and measure for a subclass of potentials defined by P. Walters, Erg. Theo. Dyn. Syst. 3**3**(5), 1338–1362 (2013)
4. Baraviera, A.T., Cioletti, L.M., Lopes, A.O., Mohr, J., Souza, R.R.: On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. Rev. Math. Phys. **23**(10), 1063–1113 (2011)
5. Bousch, T.: Le poisson n'a pas d'aretes, Ann. Inst. Henri Poincare, Proba. et Stat., **36**, 489–508 (2000)
6. T. Bousch and O. Jenkinson, Cohomology classes of dynamically nonnegative Ck functions, Inventiones mathematicae 148 (2002), 207–217
7. Contreras, G., Lopes, A.O., Thieullen, P.: Lyapunov minimizing measures for expanding maps of the circle. Ergodic Theory and Dynamical Systems **21**, 1379–1409 (2001)
8. Chou, W., Griffiths, R.: Ground states of one-dimensional systems using effective potentials. Physical Review B **34**(9), 6219–6234 (1986)

9. G. Contreras, Ground states are generically a periodic orbit, Invent. Math. 205, no. 2, 383–412. (2016)
10. da Cunha, R.D., Oliveira, E.R., Strobin, F.: A multiresolution algorithm to generate images of generalized fuzzy fractal attractors. Numer. Algor. (2020). https://doi.org/10.1007/s11075-020-00886-w
11. da Cunha, R.D., Oliveira, E.R., Strobin, F.: A multiresolution algorithm to approximate the Hutchinson measure for IFS and GIFS, arXiv https://arxiv.org/abs/1909.03052 (2020)
12. Dotson, W.G.: On the Mann iterative process. Trans. Amer. Math. Soc. **149**, 65-73. 65–73 (1970)
13. Ferreira, H.H., Lopes, A.O., Oliveira, E.R.: Explicit examples in Ergodic Optimization. Sao Paulo J. Math. Sci **14**, 443–489 (2020)
14. Ferreira, H.H.: Um processo iterativo para aproximar sub-acoes calibradas e exemplos explicitos em Otimizacao Ergodica, MSc Dissertation, UFRGS - Porto Alegre (2021)
15. Garibaldi, E., Lopes, A.O.: On the Aubry-Mather Theory for Symbolic Dynamics. Erg. Theo. and Dyn Systems **28**(3), 791–815 (2008)
16. Garibaldi, E.: Ergodic Optimization in the Expanding Case. Springer (2017)
17. Ishikawa, S.: Fixed Points and Iteration of a Nonexpansive Mapping in a Banach Space. Proceedings of the American Mathematical Society **59**(1), 65–71 (1976)
18. Jenkinson, O.: Ergodic optimization in dynamical systems. Ergodic Theory Dynam. Syst. 39(10), 2593–2618 (2019)
19. O. Jenkinson, A partial order on x 2 -invariant measures, Math. Res. Lett. 15, no. 5, 893–900 (2008)
20. Jenkinson, O., Steel, J.: Majorization of invariant measures for orientation-reversing maps. Ergodic Theory Dynam. Systems **30**(5), 1471–1483 (2010)
21. O. Jenkinson, Optimization and majorization of invariant measures, Electron. Res. Announc. Amer. Math. Soc. 13, 1–12 (2007)
22. Krasnoselski, M.A.: Two remarks on the method of successive approximations. Uspehi Mat. Nauk **10**(1) (63), 123-127. (Russian) MR 16, 833 (1955)
23. Lopes, A.O., Oliveira, E.R., Smania, D.: Ergodic transport theory and piecewise analytic Subactions for analytic dynamics. Bull. Braz. Math Soc. **43**(3), 467–512 (2012)
24. Lopes, A., Oliveira, E., Thieullen, Ph.: The dual potential, the involution kernel and transport in Ergodic optimization, dynamics, games and science. In: Bourguignon, J.-P., Jelstch, R., Pinto, A., Viana, M. (eds.) International Conference and Advanced School Planet Earth DGS II, Portugal (2013), Edit. Springer, pp. 357–398 (2015)
25. Ma, T.-W.: Classical Analysis on Normed Spaces, W. Scie (1995)
26. Robert Mann, W.: Mean value methods in iteration. Proc. Amer. Math. Soc. **4**, 506–510 (1953)
27. S.V. Savchenko, Homological inequalities for finite topological Markov chains, Funct. Anal. Appl. 33 (1999), 236–238
28. H. Senter and W. Dotson, Approximating fixed points of nonexpansive mappings, Proc. Amer. Math. Soc. 44, 375–380 (1974)

"Beat the Gun": The Phenomenon of Liquidity

Alfredo D. Garcia and Martin A. Szybisz

Abstract We describe the formation of liquidity condensations, in a random matrix that represents a complex financial market (high network interconnectivity) exposed to systemic risk. In a percolation model we simulate runs for different strategies of economic agents to study diverse types of fluctuations and the limits for an eventual phase transition (characterized by the existence of booms or market crash). The liquidity of financial assets arises as a result of agent's interaction and not as intrinsic properties of the assets. We focus on the formation process of the phase transition probability. Small differences in the strategic rules adopted by the agents lead to divergent paths of market liquidity. Our simulation also supports the idea that the higher the maximum local allowed fluctuation the higher the path divergence.

Keywords Liquidity · Finance · Percolation · Stochastic models · Simulations

1 Introduction

Physical investments (machinery, equipment, real estate) require time (generally a long period) not only to be decided, but also to be implemented before achieving the expected results. Once the investment decision is made, it is irreversible [1]. The same applies to loans for expenses, which require not only time; the repayment sources of must be originated as an added value of real economic activity. Once the debt is instrumented, it is irreversible unless early redemption. In both cases (physical investment and loans) there are risks that remain over time, linked to the difference between the recovering expectations of the invested or borrowed amount and the actual ability to achieve it. Traditionally physical investments are illiquid (given time maturation and the uncertainty that they entail) when compared to financial investments. However, the latter represent the first in eventually tradable instruments (in the form of shares which express the economy of a company as notes or bonds)

A. D. Garcia (✉) · M. A. Szybisz
Faculty of Economic Science, Buenos Aires University, Av. Cordoba 2122, C1120 AAQ Buenos Aires, Argentina

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_9

promising a future flow of payments to be obtained from the productive activity. Physical and financial investments are mirrored, as two sides of the same coin, or, in accounting terms, are computed on one side or the other on the same balance sheets. The illusion that a machine operating in a factory is illiquid but the stock that represents it is liquid; produces a division, purely virtual, between those who maintain long-term physical investment as part of their wealth and those who buy and sell, in shorter terms, the financial papers that represent it.

In this paper we evaluate the consequences of this virtual division. Small differences in the strategic rules adopted by the agents lead to divergent paths of market liquidity. Our simulation also supports the idea that the higher the maximum local allowed fluctuation the higher the path divergence. Section 2 discusses the main characteristics of liquidity, in Sect. 3 the implications of the percolation model for the study of economics are analysed, Sect. 4 describes the model and simulations performed whereas Sect. 5 concludes.

2 Liquidity

Liquidity can be seen as the ability of an asset to be bought or sold at a value that remains stable at any future time.

There is no consensus in the literature on the definition of liquidity. Sometimes it focuses on specific assets (relative position in a pyramid created by the bank system, sovereign bonds, blue chip stock), some other times definitions refer to a characteristic of an agent (its ability to generate cash flow) and sometimes to a medium of exchange [2]. Some authors try to establish links between what they call macro liquidity (monetary flows) and micro liquidity (transactional liquidity) where monetary expansions increase liquidity in the stock market and interest rate movements impact the volatility of bonds and shares [3].

Another refinement points to the existence of different types of liquidity. A distinction is made between central bank,[1] fund and market liquidity [5]. From these aspects we can infer that conditions necessary to study liquidity include the analysis of various types of instruments (money, bonds, stocks and other financial positions) that are related in multiple aspects (for example, market and funding liquidity) with heterogeneous agents (central banks, banking institutions, sovereign wealth funds, hedge funds, high net worth individuals) that operate within the market.

In terms of dynamics, Nikolaou [6] considers liquidity as a flow. Due to the time-variability of liquidity Nikolaidi [7] proposes a definition of liquidity that varies over time.

With the background cited above and following these last two authors we can understand that liquidity is a relationship between different assets that emerges from complex dynamics[2] and not a intrinsic characteristic of any of the particular assets. In other words, the same asset (for instance, a blue chip stock) can be a perfect liquid

[1] Monetary base offer [4].

[2] Expanding Chordia's cited work in terms of formation of transmission channels.

asset during long time, entering in different portfolios, but in a stress market situation may become illiquid (have to be sold accepting a big price discount).

Categorical definitions often cause paradoxical situations. For example, money issued by a central bank is liquid in times of low inflation but may lose liquidity in the face of inflation expectations, eventually being replaced by another asset (a foreign reserve currency, inflation protected bond, or some alternative commodity, such as gold).

Liquidity "fluctuates", leading to situations that may be extreme, when for example it is said that the financial market is "dry" because only a few assets maintain the certainty of being recognized at their par value in the future and are retained in the portfolios; on the contrary it is said that the market has excess of liquidity since more assets (discounted checks, securitized debt, sovereign bonds) are considered close substitutes. Liquidity is therefore an outcome, an emerging phenomena, in the interaction of agents and their perceptions about the future and not an essential property of a particular asset that is fixed over time.

To measure liquidity, it is necessary to consider the set of relationships that link and value the entire spectrum of assets which are operated in the markets, and requires knowledge of interconnection rules, the network topology that form the transactions and the dynamics between the financial assets and the real economy.

The complexity of the aforementioned interactions is widely addressed in Keynes's General Theory, and the liquidity problem specifically mentioned in the following paragraph:

> Thus the professional investor is forced to concern himself with the anticipation of impending changes, in the news or in the atmosphere, of the kind by which experience shows that the mass psychology of the market is most influenced. This is the inevitable result of investment markets organised with a view to so-called "liquidity". Of the maxims of orthodox finance none, surely, is more anti-social than the fetish of liquidity, the doctrine that it is a positive virtue on the part of investment institutions to concentrate their resources upon the holding of "liquid" securities. It forgets that there is no such thing as liquidity of investment for the community as a whole. The social object of skilled investment should be to defeat the dark forces of time and ignorance which envelop our future. The actual, private object of the most skilled investment to-day is "to beat the gun", as the Americans so well express it, to outwit the crowd, and to pass the bad, or depreciating, half-crown to the other fellow. [8]

Holders of financial assets have in their hands a varied and growing number of instruments (since Keynes's time) to regulate liquidity.[3]

The hierarchical pyramid that structures liquidity [9] expands and contracts cyclically and generates debates as to what should be defined as money proper (M1, M2, M3).

The existence of liquidity arises from a continuous range of interchangeable assets on par or near par, that is, with little uncertainty on their value over extended periods of time. With a perfect forecast on the physical investment results, we would be living in the world thought by Keynes of "forever" investors, where there would be

[3] In all its variants, central bank, funding and market liquidity. From central bank deposits, bank deposits to repurchase agreements and financial derivatives.

no liquidity needs (and therefore no illiquidity). In this world, original owners will kept their physical investments until the end, collecting the expected return.

When forecasts includes uncertainty, liquidity is created as it becomes possible to transfer property rights to other agents. The previous "forever holders" may want to diversify their portfolios looking to diversify risks.

The emergent liquidity gives way to some agents to think and act as if the real economy can be anticipated and "beat" uncertainty (beat the gun), trading its future results.

In this framework where uncertainty, expectations, agents relationship forms, assets and financial positions are intertwined to produce the phenomenon of liquidity; we propose a model where liquidity is the result of a complex interaction that arises from different perceptions about the future of the real economic system.

3 Percolation Model

Percolation[4] studies the connectivity between defined units (sites or nodes) in an array where cluster formation occurs, defined as groups of connected neighbouring nodes. The properties that arise in such systems can be consulted for example in Stauffer [10]. The dynamic combines deterministic rules for agents and stochastic results that do not require normal or previously defined probability distributions.

Another innovative element of using percolation with respect to traditional models[5] is that it can identify the phenomenon of liquidity in the complex interaction between agents and not by the emission of instruments from an exogenous agent (in general a central bank). In our model, liquidity is constantly fluctuating and agents attempt to "beat the market" that is, anticipate others with respect to an expected result, can be successful or frustrated without the need to assume defined behaviours based on the characteristics of agents (rational or irrational) but of heuristic rules.

Our model does not need to establish a theoretical framework regarding agent's behaviour. Since they are part of pre-established institutional games and part of a culture, they have a narrow scope for decision making. Additionally, as they don't know the final prices of each period until the market is closed, optimistic or pessimistic approaches only will be taken at the end of each period. Portfolio diversification, degree of exposure, have to be re-evaluated in every period.

The chances of liquidity entering the system is represented by the appearance of coalitions in clusters of liquidity, the combination of which the model interprets as the general state of the system in each period. Agents' response to a specific business climate may or may not (probabilistically) result in speculative bubbles. When a cluster becomes large enough to cross the system from side to side, the overall situation of the system changes radically (a phase transition occurs).

[4] In our model we will refer to site percolation.

[5] Models such as Dynamic stochastic general equilibrium (DSGE).

The idea of a complex network emphasizes some important issues: (a) agents need to evaluate their portfolios looking at others portfolios (value is a relative concept, financial assets haven't a value "per se"), (b) since prices and stocks transactions are known at closing, all agents are only aware of their final achievement simultaneously at closing time, (c) agents operate with rules that modify the desired liquidity of their portfolios, which may be different compared to the value of the portfolio finally achieved in each period, (d) at closing they will know if the market is near the equilibrium point when assets prices movement show lower volatility.

The interactions are deterministic as to the rule that the agents follow, but the results for the set are stochastic, since there is no Walrasian auctioneer that identifies a general balance before all transactions are completed. Nor are there "representative agents" that synthesize aggregate behaviour in different markets, such as in DGSE models, but rather interactions at an intermediate level between micro and macro results.

The characterization of a "complex system" emphasizes that there is a network topology and that a single asset is not capable of constantly preserving a certain degree of liquidity defined a priori. Portfolios value is a result related to all the others portfolios, since value is a relative concept in relation to all other assets.

The percolation model also allows to reflect the interconnectivity between the "real" sector (physical investments represented by the matrix) and the financial sector (virtual investments that mirror portions of the matrix in the form of clusters). The propagation of shocks that connect one sector with the other, allows to highlight the cases where the existence of systemic risk can be verified.[6]

That a system percolated means that there is at least one cluster of liquidity large enough to pass through the matrix from one of its edges to the opposite. An example of percolation for a square matrix is shown in Figs. 1 and 2. The probability that this occurs is centred around a particular saturation value (probability of occupation or connection) and do not depend on the size of the matrix.

We can mention that for two dimensions the triangle structure percolate for an occupancy probability of $p = 0.5$ while the honeycomb structure does it at $p = 0.6970$ and the square one at $p = 0.5927$ [12]. In general, given the size, the more neighbours the percolation limit is lower. Which of them is the one that best represents

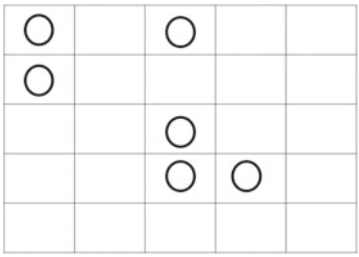
[6] This type of risk is highlighted by the Basel banking supervision committee report in the Basel III documents [11] that argues that:

> While procyclicality amplified shocks over the time dimension, excessive interconnectedness among systemically important banks also transmitted shocks across the financial system and economy.

Systemic risk is generated largely via:

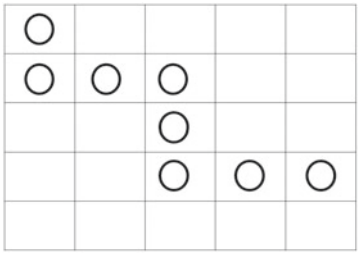
> an array of complex transactions.

.



(a) Clusters

Fig. 1 Example of three clusters: in a square of 5 by 5 sites each locker is occupied given a certain probability of occupation. A cautionary note with the definition of neighbour: in this case the diagonal boxes are not considered neighbours. A cluster is defined if at least one site (or more neighbour sites) is (are) occupied and the neighbour sites are not



(a) Percolation

Fig. 2 Example of cluster percolation: it is possible to follow a path from the occupied site to the left to the right through the entire square

the complex financial world and how many dimensions it has are open questions for future research.[7]

On the other hand, the range in which there is a greater than zero probability that percolation may occur does depend on the size of the matrix. The smaller the size, the greater the range of percolation probability. In a previous work we have studied this topic highlighting that the number of nodes of an economic system (in the range of hundreds or thousands) is very far from the size of the systems studied in physics (10^{23}) or the infinite size considered in the theory [13].

Thus, a larger matrix connects clusters for longer time before percolation occurs and when it does, it corresponds to a more massive cluster. In this type of system, a larger size does not make the system more robust but simply differs the time in which the phase transition eventually occurs.

[7] In this work we do not present any particular topology, since the choice of a specific structure would extend the work outside acceptable limits.

3.1 Percolation and Finance

Some authors have used this model in economics as a reaction to standard models that assume normal probability distributions which are not consistent with empirical evidence of kurtosis, fat-tails distributions and fluctuations outside the ranges the usual models predict.[8]

The network topology has been studied in detail in reference to the number of connections of each node, the degree of interconnection between them and the degree of asortativity (degree of homogeneity between nodes). In their work, Loepfe [16] state that the relevant properties of real-world banking systems include high grouping; size heterogeneity and contagion sensitivity even with low density connections. The authors also point out that the transition from "safe" to "risky" regimes can be very fast [17], which is another element the percolation model provides. The models we know so far use percolation to study fluctuations in short-term financial markets, our variation considers a regime that expands and contracts the liquidity of the system over longer terms.

With regard to assets prices, financial operators are usually divided by the adoption of two practical analysis: (a) those who use projections of future cash flows (dividends, bond yields etc.) discounted at an adequate interest rates (the so-called fundamental analysis) and those that analyse various curves of moving averages and trends (technical analysis). Both formulations allow us to emphasize that use of simple heuristic rules that consider a trade off between expected returns from the real side of the economy and portfolio liquidity, since what is relevant to decide is based always on the future expected price. Historic prices are relevant information, in the formalization here proposed, agents use at least immediate recent ranges of past prices.[9]

The efficient-market hypothesis (EMH) [18] states that asset prices reflect all available information, so it is impossible to "beat the market" consistently since market prices should only react to new information. Since relevant information is also related to expectations on what other agents are doing uncertainty is relevant to the idea of periodic portfolio review about the level of liquidity they have. This give room for two aspects to be considered: (a) all financial agents deal with uncertainty about the fulfilment of expected cash flows and choose a relative level of liquidity. (b) all financial assets belong to some portfolio and only part of the assets are traded, the rest is valued at the current price, so, the financial portfolios have permanent changes in value. This two aspects related to liquidity give us the chance to divide financial trading between conservative agents and short term speculators. Conservative agents (who have a strategy closer to the "forever" in Keynes view) and very short-term speculators who seek to anticipate the market (to beat the gun) giving up liquidity when they think the payment promises will be fulfilled and running to liquidity when doubts appear. Our model combines these type of simple rules that, when interacting, generate complex results.

[8] See for example [14, 15].

[9] The point of the significance of past prices was suggest by professor Hartwell.

The interaction between liquidity and real economy shows significant instability. The price/earning ratio (P/E) for the 500 shares of the most relevant US companies according to financial data [19] fluctuates between 10 in 1954 to maximum exceeding 30 in 2000; with lows near 7 in the late seventies and early eighties averaging 16.56 for the period 1954–2017. Volatility is also high, in 1992–2009 values reached 27, whereas in early 2008 (and in 1988) its value exceed 10 by little margin.

The liquidity of portfolios is a central indicator for agents. For conservatives because it is a bridge between the uncertain and long-term results that the real economy may provide. It is often said that in conservative portfolios "cash is king", as they tent to maintain a high proportion of liquid assets. For speculators, liquidity is the instrument to wait for the moment to enter the market, incorporate less liquid assets and then sell them to make a profit without waiting for the real investment maturity (beat the gun).

3.2 The Structure of the Model

We consider a topology of nodes (sites) and connections where each node represents a physical asset (real estate, productive companies, obligations of final debtors such as employees, companies and governments). If there were no possibility of issuing financial assets, each original investor would keep his real asset until the end of his useful existence. In the absence of the concept of liquidity, nodes would remain empty.

As investors (professionals and speculators, in the words of Keynes) produce (inject) financial assets, nodes are occupied and liquidity permeates in the matrix with a certain probability (see Fig. 3), which arises from the decision rules of the agents.

We distinguish two types of actors: financial investors (rentiers), who decide to increase or reduce the liquidity of their portfolios through placements (deposits) in specialized financial institutions (professional investors and speculators). These institutions buy and sell financial assets based on the deposits received.[10]

The model simulates the transition from an initial liquidity situation to another where liquidity clusters are formed. We do not distinguish in this model between the different instruments mentioned in Sect. 2.

Each investor may know what degree of liquidity he wants to inject into the system (buying rights to future income of the real sector), but the results of the aggregate of the decisions are only known at the end of each period and nobody can anticipate the aggregate result. The information used in each period takes into account what

[10] It is important to note that we differentiate the normal depositor from the rentier, the normal depositor uses their deposits to cover their normal flows of expenditure related to consumption or direct productive activities, the rentier on the contrary injects liquidity to the system, while its income is not used in the current expenditure circuit [20]. Examples of rentier are large family fortunes, sovereign wealth funds or pension funds. Examples of institutions are investment banks, mutual funds, brokers, hedge funds, etc.

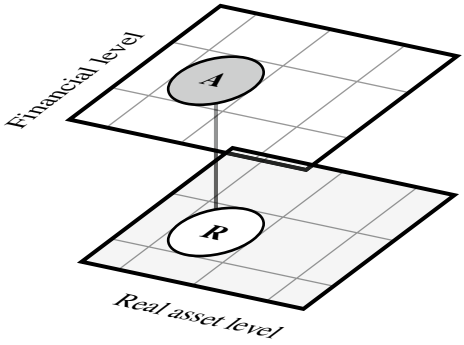


Fig. 3 Real and Financial levels

happened in previous periods. The rules followed by the agents reflect two aspects recognized by the operators: (a) what the market is doing (results review of several recent previous periods) and (b) agents are aware of the real business cycles and react to consecutive market results of the same sign.

The heuristic rules of Sect. 4 are defined as (1) for rentiers there is a w proportion that determines how much of their wealth is deposited in specialized financial entities and (2) for these institutions there is a b proportion of the deposits received, which are dedicated to the purchase of financial assets (always backed by real assets—mainly stocks and bonds).

The variable w also indicates the confidence of the rentiers in financial institutions, since the larger w, the more confidence in the financial system and the greater the proportion of the deposited wealth.

Proportion b represents the aggregate liquidity position that results from the decisions of all financial institutions. The higher b, the greater the liquidity that penetrates the matrix. For example, banks can expand or reduce the liquidity received in deposits, applying reserve requirements or taking advances from the Central Bank; other specialized institutions can maintain a conservative position or conversely get leverage (take debt) and buy more financial assets than they are allowed by the original deposits.

In this way the proportion of assets that gain liquidity p is at the same time the probability of percolation of the system.

$$p = w \times b \tag{1}$$

Variables w, b and p have a minimum value of 0 and a maximum of 1.

In Fig. 3 we represent the case for a physical and a financial good at square matrix levels.

4 Rules Definitions and Simulations

The combination of activities defines a matrix of n elements that are represented by a plot of m nodes. These nodes may or may not be interconnected, depending on whether it acts in isolation or as a coalition (cluster).

4.1 Rules Definitions

Each node represents the combination of participating rentiers (providing liquidity) and institutions willing to acquire (or sell) financial assets.

The Δw variation of the w proportion of their wealth that rentiers deposit in financial institutions is modelled as a random variation with a maximum threshold.[11]

The variation Δb of the proportion b representing the aggregate decisions of the financial institutions on their degree of aggregate liquidity is also modelled as a random variation.[12] The higher b, the greater the use of resources by financial institutions and the greater the liquidity that penetrates the matrix.

In this sense, we deviate from the classic Stauffer and Penna [21] models that do not consider any economic information and Tanaka [22] that only takes into account information of nearby neighbour interactions. But, as Stauffer and Penna cited by Tanaka admit, all agents share the same information within a few minutes. Therefore, our approach allows agents to act based on global (not only local) information while retaining the possibility of phase transitions.

The randomness of the variation represent the fact that new information is valuable [23], cannot be known in advance and produce revaluation of all portfolios at the same time. In this sense, this work departs from the hypothesis in which all the economically relevant information about an asset is contained in its price [24]. In the context of this work, each individual rentier or institution cannot control the confidence of the market as a whole or become totally independent of the strategy followed by the others.

Because of the system's reliance on payment promises mentioned in Sect. 3.2, two aspects are to be considered: (a) uncertainty of future results and (b) the level of liquidity is adjusted every time new information is available.

We assume scenarios where one every i periods (in random terms) is negative to reflect the fact that recessive periods are usually scarcer.

We define two types of rules. With γ agents take into account the last n variations Δw for rentiers and Δb for banks. If these are all negative then they multiply the last variation by j. The β rule causes agents to react by multiplying their variation by j,

[11] Distributed uniformly to calculate its absolute value, a proportion k of these are negative variations.

[12] Distributed uniformly to calculate its absolute value, from those values a proportion k are negative variations.

Table 1 Rules and scenarios of simulations

			Scenarios	
	If	Negative fluctuation (ng)	Stable Path every 10 periods MRF 1%	Unstable Path every 5 periods MRF 2%
			Action	
Rules	Rule β	One period	Multiplies by 10 ng	Multiplies by 5 ng
	Rule γ	Three periods	Multiplies by 10 ng	Multiplies by 5 ng

if a single Δw or Δb variation is negative ($n = 1$). In all cases j includes the current period.

In this way we have two types of agents, those who follow the γ rule wait more time having more confidence on their positions. Agents following the β rule have less confidence in their expectations and change them quickly and expose themselves less.

Rentiers only take information on variations of w and react with $j\Delta w$ if the rule applies. In this case, financial institutions apply the same rule[13] $(j(-)|\Delta b|)$[14] (n periods of Δw negative). Financial institutions additionally follow the rules β or γ also with respect to Δb, so they use additional information that can lead to $j\Delta b$.

4.2 Simulations

For each graph we have run 30 series of 100 periods each,[15] series of the value p are shown. We take an initial value of p of 0.2[16] for all series. In the case of combination of rules, the proportion of financial institutions and rentiers using the same rule are equal.

First, we can assume a financial region with a stable expansion path where the amount of negative periods is relatively small (1 every 10, $i = 10$) and the fluctuations of proportions are usually not wide (in our case we take 1% as the maximum of the range of fluctuation (MRF)) which we display in Fig. 4. When the β or γ rule is applied, the fluctuation becomes 10 ($j = 10$) times more negative. The percentages in Table 1 refer to the MRF band.[17] For all the graphs an average series is calculated (highlighted in black).

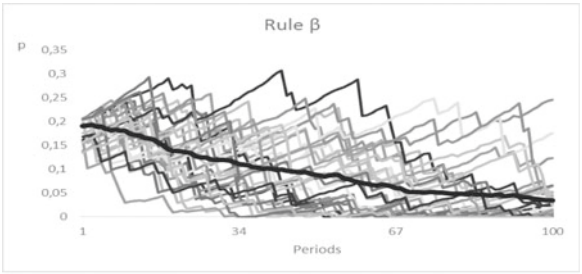
[13] In this case the variation of b is always taken as negative and the rule is applied.

[14] In this way a coordination phenomenon occurs.

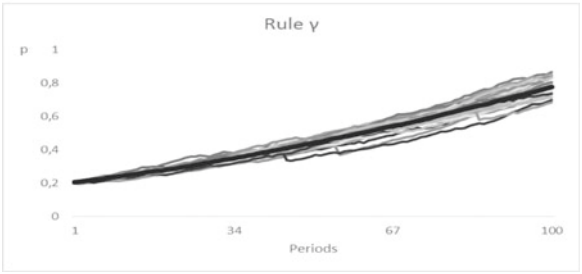
[15] For this we develop a program in C++, the obtained series are analysed with Excel.

[16] We use this initial value because it constitute a conservative combination of w and b that indicates modest financial deepening.

[17] To these variations rules β and γ can be applied if the conditions are met.

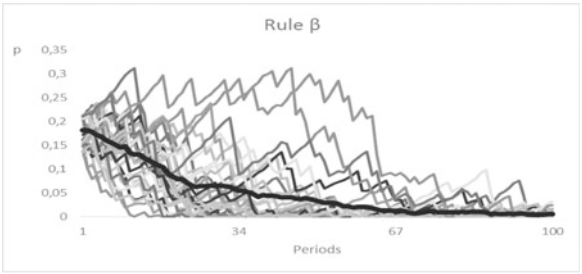


(a) Rule β

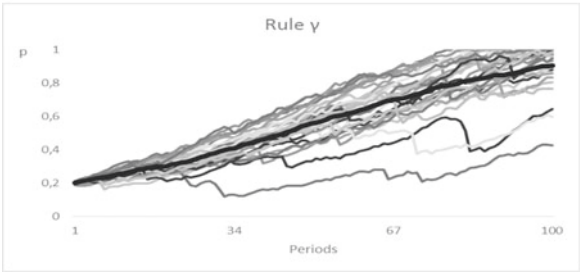


(b) Rule γ

Fig. 4 30 Series of rule evolution separately, 1 out of 10 negative periods, maximum normal fluctuation 1%



(a) Rule β

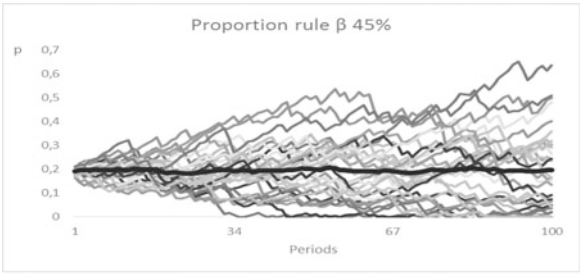


(b) Rule γ

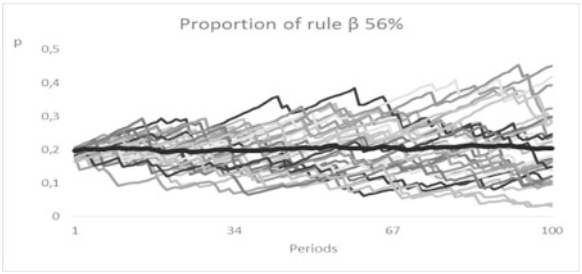
Fig. 5 30 Series of rules evolution separately, 1 out of 5 negative periods, maximum normal fluctuation 2%

Table 2 Average and Standard Deviation of final values of series

	Stable path		Unstable path	
	Average	Std dev	Average	Std dev
Rule β alone	0,03484	0,05606	0,005256	0,00801
Rule γ alone	0,77619	0,04918	0,90492	0,13788
Rule β	56%		45%	
	0,20444	0,10795	0,19701	0,17795



(a) Rule $\beta\gamma$ unstable path

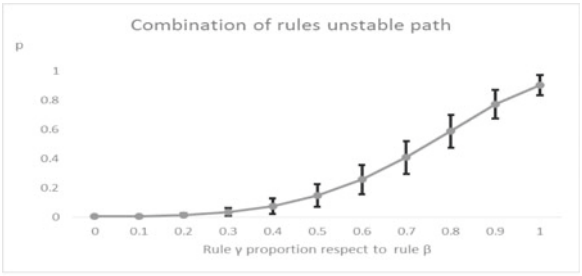


(b) Rule $\beta\gamma$ stable path

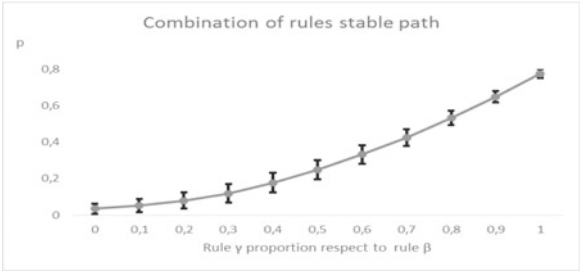
Fig. 6 30 Series of evolution of rules as a whole, stable and unstable paths

These parameters can be compared to a financial region with a more unstable expansion path (negative periods 1 every 5, $i = 5$) and wider fluctuations (in our case we take 2%) in Fig. 5. When either β or γ rules are applied, the fluctuation becomes 5 ($j = 5$) times more negative.

Table 2 shows average values of the series for the different hypotheses. The first part of the table shows the values obtained using the β and γ rules separately for each hypothesis. The second part refers to the results obtained by combining the rules, with the β rule used by 56% of agents in the stable case and by 45% in the unstable case. The proportion of the β rule used is shown, the complement with the γ rule adds up to 100%. In the case of a combination of rules, we show the combination that leaves the initial value of the series invariant.



(a) Unestable path



(b) Stable path

Fig. 7 Average of the final value of 30 series of joint Evolution with standard deviation for each point, 100 periods each series

Figure 6 shows the temporal evolution of the results displayed in Table 2. Finally, we present the graphs of the proportional relationships of the rules in cases of unstable and stable paths in Fig. 7.

In neither experiment the relationship is linear as seen in Fig. 7. In the case of a stable path, the first derivative and the second derivative of the relationship curve are positive, so p grows at increasing rates within the definition range. The standard deviation is proportionally greater in the middle range of the graph.

In the case of an unstable path, the curve suggests a positive first derivative for the entire range, but the second derivative takes positive first and negative values at values close to $\gamma = 1$. Standard deviations are markedly higher in the region with a high proportion of the γ rule.

5 Conclusions

We have presented a model that describes the formation of liquidity based on interactions between agents that increase or retract their financial exposure to the expected results of the real sector. To model these circumstances we implement two rules that describe possible strategies to follow when agents receive new information.

In this sense, we deviate from the classic Stauffer and Penna [21] model that do not consider any economic information and that of Tanaka [22] which only takes into

account nearby neighbour interactions. Stauffer and Penna cited by Tanaka recognize the importance of global information even if they do not use it. Our model aims to be an advance in this regard.

Unlike traditional models [24], we do not consider current prices as the relevant and almost exclusive variable for making decisions. Historical prices are a result of the complex interaction of portfolio strategies. Negative expectations contract liquidity (many assets that were previously liquid lost this characteristic) and positive expectations extend it (there is a greater propensity to get into and stay tied to the results of the real economy).

The randomness of the liquidity variation expose the fact that new information, produced by multiple agents at the same time, is valuable [23] but cannot be anticipated with certainty in its aggregate result. In the context of this work, each individual rentier or institution cannot anticipate how much liquidity (market confidence) will result from the strategies that other agents are following.

The creation of liquidity leads to the appearance of clusters, which in turn can percolate. The percolation of the system represents an excess of liquidity with respect to the underlying illiquidity of physical assets. Interconnection allows to diffuse in the system the effects of the fall of any of the nodes [25] due to non-compliance with the expectations on the underlying real asset.

This approach allows to consider a wider range of underlying economic phenomena that can be incorporated into these types of models, such as divergences of expected returns with respect to real returns on investments.

The key point of the study is that slightly different decision rules (such as those defined for β and γ) generate very different paths of liquidity expansion. The first one (β) diverge from the percolation limit while the second (γ) produces it almost inevitably.

More stable expansion paths tend to have higher percolation probabilities and therefore the system is more vulnerable in terms of interconnections.

The results suggest that high proportions of liquidity in the economic system, being connected in clusters, may be detrimental in a system with low volatility, in principle, every financial asset that deviates from calculated expectations has the ability to infect the system and turn it highly volatile.

On the other hand, the existence of a certain proportion of financial assets can encourage the transfer of resources to the best projects [26, 27].

From the point of view of active policies, considering these kinds of theories may not be seen as a sign of the inevitability of the disorder and the impossibility of analysis, but rather as a fertile ground where new types of interventions may be explored [28] to more effectively stabilize the system.

Acknowledgements The authors would like to thank the organizers and participants of the XX JOLATE Conference and professor Christopher Hartwell for constructive criticism of the manuscript, errors remain ours.

References

1. Bernanke B.S.: Irreversibility, uncertainty, and cyclical investment. Q. J. Econ. **98**(1), 85–106 (1983)
2. Kawamura, E., Antinolfi, G.: Some observations on the notion of liquidity. Technical report, Society for Economic Dynamics (2010)
3. Chordia, T., Sarkar, A., Subrahmanyam, A.: An empirical analysis of stock and bond market liquidity. Rev. Financ. Stud. **18**(1), 85–129 (2004)
4. ECB: The monetary policy of ECB. ECB Publications (2004)
5. Brunnermeier, M.K., Pedersen, L.H.: Market liquidity and funding liquidity. Rev. Financ. Stud. **22**(6), 2201–2238 (2008)
6. Nikolaou, K.: Liquidity (risk) concepts: definitions and interactions. ECB Wokring Paper Series, vol. 1008 (2009)
7. Nikolaidi, M.: Bank liquidity and macroeconomic fragility: empirical evidence for the EMU (2016)
8. Keynes, J.M.: The General Theory of Employment, Interest, and Money. Macmillan Co., London (1936)
9. Gabor, D., Vestergaard, J.: Towards a theory of shadow money. In: Institute for New Economic Thinking Working Paper (2016)
10. Stauffer, D., Aharony, A.: Introduction to percolation theory. Taylor & Francis (2014)
11. Basel Committee on Banking Supervision: A global regulatory framework for more resilent banks and banking systems (2010)
12. Tarasevich, Y.Y., Van Der Marck, S.C.: An investigation of site-bond percolation on many lattices. Int. J. Mod. Phys. C **10**(07), 1193–1204 (1999)
13. Garcia, A.D., Szybisz, M. :Dinámica de transiciones de fase de largo plazo en mercados complejos. Anales Reunion Anual AAEP (2016)
14. Cont, R., Bouchaud, J.-P.: Herd behavior and aggregate fluctuations in financial markets. Macroecon. Dyn. **4**(2), 170–196 (2000)
15. Stauffer, D., Sornette, D.: Self-organized percolation model for stock market fluctuations. Phys. A: Stat. Mech. Its Appl. **271**(3–4), 496–506 (1999)
16. Loepfe, L., Cabrales, A., Sánchez, A.: Towards a proper assignment of systemic risk: the combined roles of network topology and shock characteristics. PloS One **8**(10), e77526 (2013)
17. IMF: Global financial stability report navigating monetary policy challenges and managing risks (2015)
18. Malkiel, B.G., Fama, E.F.: Efficient capital markets: a review of theory and empirical work. J. Financ. **25**(2), 383–417 (1970)
19. bespokepremium.com. [P/e ratio] (2017)
20. Nicolini Llosa, J.L., Garcia, A.D., Guerra, E.: The financial rentier in the 21st century. Anales Reunion Anual AAEP (2018)
21. Stauffer, D., Penna, T.J.P.: Crossover in the cont-bouchaud percolation model for market fluctuations. Phys. A: Stat. Mech. Its Appl. **256**(1–2), 284–290 (1998)
22. Tanaka, H.: A percolation model of stock price fluctuations. Math. Econ. **1264**, 203–218 (2002)
23. Stiglitz, J.E., Weiss, A.: Credit rationing in markets with imperfect information. The Am. Econ. Rev. **71**(3), 393–410 (1981)
24. Fama, E.F.: The behavior of stock-market prices. J. Bus. **38**(1), 34–105 (1965)
25. Cont, R., Moussa, A. et al.: Network structure and systemic risk in banking systems. In: Edson Bastos e, Network Structure and Systemic Risk in Banking Systems (December 1, 2010) (2010)
26. John Stuart Mill: Principles of Political Economy: With Some of their Applications to Social Philosophy, vol. 1. Reader, and Dyer, Longmans, Green (1871)
27. Schumpeter, J.A.: Theory of Economic Development. Routledge (2017)
28. Bullard, J., Butler, A.: Nonlinearity and chaos in economic models: implications for policy decisions. Econ. J. **103**(419), 849–867 (1993)

Board Knowledge and Bank Risk-Taking. An International Analysis

E. Gómez-Escalonilla and L. Parte

Abstract Corporate governance mechanism and board knowledge are crucial variables to mitigate bank risk taking. This study focuses on two variables related to board members profile: the knowledge of board members by the accumulated experience in the same company (Board Tenure), and the expertise of board members derived from previous competences obtained in the industry or by specific training in the financial field (Board Skills). Using a sample of 156 banks from 37 countries for the period 2009–2017, the association between both corporate governance variables and bank risk is tested through the Generalized Method of Moments model. The results show that specific financial skills decrease the level of bank risk. In contrast, there are no conclusive results on board tenure and the risk measures analyzed.

Keywords Corporate governance · Board tenure · Board skills · Bank risk

1 Introduction

The role of boards and corporate governance prescriptions in the banking industry are particularly relevant for several reason. First, the banking industry is a crucial industry for the economy worldwide. Second, banks' failures not only affect the banking industry but also have implications to other industries and across countries. Third, the role of boards in the banking industry has been noted as one of the critical factors that failed in the financial crisis in 2007–2008. Hence, market supervisors and

E. Gómez-Escalonilla · L. Parte (✉)
Faculty of Economics and Business Administration, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain
e-mail: lparte@cee.uned.es

E. Gómez-Escalonilla
e-mail: egomezesc1@alumno.uned.es

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_10

regulatory bodies have published a set of guidelines and recommendations during the last decade, with a special focus on the board structure and composition (e.g. [5, 15, 19, 31, 32]). To date, these recommendations and regulations continue being updated on a regular basis: for example, according to a report published by OECD [32], half of all jurisdictions have revised their national corporate governance codes in the past two years.

Board directors exert a dual function in the company: "Management Board" and "Supervisory Board". On one hand, the Management Board deals with operational issues and provides information and recommendations for decision-making. On the other hand, the Supervisory Board is responsible for the decision-making and for monitoring the firm including risk control. Consequently, the board has not only the responsibility of adopting and implementing proper policies that result in effective internal control systems but also the responsibility of electing the board members to guarantee an efficiency control of business risk in line with the company's mid and long-term strategies. In this context, board knowledge and specialization are a matter of considerable interest to guarantee an effective monitoring mechanism and especially, in the area of business risk.

Although the inefficiencies in the bank boards and the relationship between bank governance and performance have been deeply analyzed before the financial crisis (see e.g. [11, 14]), little evidence exists about the effect of corporate governance characteristics on bank risk (total risk, systematic risk and idiosyncratic risk) in the post crisis period. This study contributes to understand the importance of corporate governance variables to mitigate business risk, following the prescriptions and recommendations of market supervisors and regulators.

Recently, European regulator bodies in the banking and financial industries (the European Banking Authority (EBA), and the European Securities and Markets Authority (ESMA)) have published guidelines on the assessment of the suitability of directors and other key functions, which have been the main germ of this research. The joint publication by EBA and ESMA published in 2017 establishes the minimum criteria to be considered: (i) honorability; (ii) sufficient knowledge, competence and experience; (iii) independence of mind; and (iv) time dedication. The notion of "independence of mind" is more than a formal position of independence on the board. In this context, the guidelines indicate that the ability to ask questions, to obtain more information or to have a greater knowledge of the business and its risks, as well as the ability to resist "group thinking", should be evaluated. This eligibility criterion is relevant for monitoring business risk in general, but is even more important in the banking system due to the characteristics of complexity and opacity of the business [20], thus implying a significant need for specialization and diversity of the directors and a high level of financial knowledge.

This study focuses on two main corporate governance variables of interests. The first is the board tenure, as a proxy variable to the knowledge of the board members by the accumulated experience in the same company, which supposedly brings a greater knowledge of the business model in the mid and long term and of the risks they are taking, allowing them to execute a better supervisory function. The second is the expertise of the board members derived from previous competences obtained in the

industry or by specific training in the financial field. In our opinion, the combination of previous expertise and experience (*Board Skills*) and knowledge acquired by board tenure (*Board Tenure*) can be useful when making decisions about the company, including risk business, and to constrain certain aspects as "group thinking" that could create inefficiencies on the board. To measure bank risk, we use two proxies: non-performing loans ratio and loan loss provision ratio.

The sample consists of 156 banks for the period 2009 to 2017 from 37 countries located in Europe, America, Asia and Africa. Hence, early corporate governance studies have focused mainly on the U.S. setting (e.g. [1, 28, 30, 33]), followed by European countries (e.g. [3, 8]) but in the last decade other economies and countries are attracting the attention of the academia (e.g. [29]). However, little research to date has focused on international samples around the world. Some exceptions are the studies by Laeven and Levine [25], Erkens et al. [14] but they do not include Asian countries. With respect to the period of time, we selected the years after the financial global crisis because most corporate governance guidelines and recommendations have been published after 2008 in order to overcome some corporate governance weaknesses detected before the financial crisis. Methodologically, we used a dynamic Generalized Method of Moments (GMM) proposed by Arellano and Bover [4], and also used in recent empirical papers that focus on the relationship between corporate governance variables and financial variables in the banking industry (e.g. [3, 11, 34]). In fact, recent empirical literature reveals that the mixed results in the corporate governance could be related to the use of inappropriate statistical techniques.

The originality and contribution of this research with respect to the previous literature can be summarized in the following ideas. First, to our knowledge, no research specifically examines the association between bank risk and board tenure and board expertise together in order to shed light on the role of the board and management risk in the banking industry. Second, the financial expertise variable has not been explored in detail in previous risk-corporate governance literature, because in most previous research this variable has been considered secondary. We consider that the complexity of the bank industry, and especially in the area of risk, makes the study of this variable interesting. Third, in the most recent literature, there is a gradually increasing number of banks and countries analyzed in the studies in order to obtain more generalizable conclusions. Consequently, this paper uses a large sample of banks worldwide.

The rest of the chapter is organized as follows: Sect. 2 presents a brief literature of corporate governance in banks and develops the hypotheses; Sect. 3 describes the sample, variables and methodology; Sect. 4 presents the descriptive statistics and the results; and Sect. 5 concludes.

2 Brief Literature Review and Hypotheses

There is a large body of literature that examines corporate governance in financial and non-financial firms. To illustrate the volume of publications dealing with this issue, a search of the words "corporate governance" in Google Scholar in 2010 returned approximately 287,000 hits, while ten years later the same search returns approximately 1,010,000 hits. The recommendations and guides published by regulatory bodies and supervisors have contributed to a growing development in the last decade. Most previous research has focused on the analysis of corporate governance indexes and corporate governance variables, with special emphasis on the board of directors. In particular, the relation between bank performance and variables, such as board size, percentage of independent directors, number of meetings and the duality of the CEO and Chairman in the same person, has been extensively analyzed (see e.g. [2, 11, 34]). More recently, academia has expanded the research objectives, including other board variables and indirect relationships, overcoming some methodological problems, and extending the setting for generalizability, among others.

Considerable effort was spent on understanding bank risk from different points of view. The literature shows that governance mechanisms influence bank risk taking (e.g. [1, 8, 13, 14, 33, 35]). Most of the previous research treats financial firms separately from non-financial firms because the banking industry has specific regulation and some corporate recommendations and guides do not hold for the rest of the industries, as for example, board independence and board size (see e.g. [1, 14, 20, 24, 27, 34]). In fact, Laeven [24] states that differences can be found in at least four main aspects of the business model because banks are "highly leveraged, they have diffuse debtholders, they are large creditors, and they are systemically important and therefore heavily regulated."

It is important to mention that bank business models are represented by funded commercial banks and capital-markets-oriented banks. The first business model involves the retail-funded commercial bank, which is the traditional banking model with the classical structure of assets and liabilities, accounts and loans, and also the wholesale-funded commercial bank, which is a model more diversified. The second business model is investment banking, which is the regular model among global banks with a focus on trading assets and market funding. Globalization has provoked banking business models to become more heterogeneous than ever before with continuous changes to operate in highly competitive markets. Furthermore, the complex nature of new products has caused high levels of growth and expansion and, apparently, a decrease in the bank risk level. The perception of a decrease in bank risk level could lead to a relaxation of different controls, such as the risk analysis responsible for conferring a loan or credit [16]. In fact, corporate governance studies point out a lack of perception about the risk of new complex financial products by boards of directors, which took risks out of the balance sheet of these institutions.

A brief summary of the most remarkable studies dealing with corporate governance and risk in banks is provided below, although the results are mixed due to differences in the variable measures, samples and settings. Pathan's study [33] finds that boards reflecting more the shareholders interest positively affect bank risk-taking while, in contrast, the partnership with CEO power is negative. Beltratti and Stulz [9] analyze this problem from a similar point of view and conclude that shareholder-friendly boards are positively associated with default risk, but the association is not robust with other risk measures (i.e., leverage risk and portfolio risk). Erkens et al. [14] analyze the influence of independent directors on risk, and found that increasing the number of these directors reduced leverage during the financial crisis. Adams [1] also found a positive relationship between the percentage of independent directors and the bank bailouts request during the financial crisis. A different perspective for studying the behavior of the board was given by the work of Elyasiani and Zhang [13], in which they examine the association between "busyness" of the board (i.e., serving on multiple boards) and bank risk. The results confirm that the bank risk (in different measures) is inversely related to this concept. Finally, Berger et al. [7] using a sample of German banks in a period from 1994 to 2010, found different results depending on the board variable (younger executives, female directors, directors with doctorate) analyzed. According to the joint publication by EBA and ESMA [12], board knowledge is crucial to understand the business model and business risk. Specifically, the publication points out "the degree to which an individual is dented to have good repute and to have, individually and collectively with other individuals, adequate knowledge, skills and experience to perform her/his/their duties". Rajgopal et al. [35] argued that the bank board did not have adequate knowledge and experience to understand the business model and thus, to take actions to control risk management at least before the financial crisis. Based on prior literature, our study focuses on two relevant variables of board knowledge that we develop in the following paragraphs.

Vafeas [38] explains that board tenure can affect the firm governance in two different ways. On one hand, the so-called "expertise hypothesis", which suggests that board tenure enables a better knowledge of the business (just as a consequence of permanence and attendance to the boards) and a greater commitment to firm strategy in the mid and long term. On the other hand, the "management friendliness hypothesis" suggests that this permanence implies a decrease in the level of control of the CEO and managers. That is, a long relationship would produce "group-think", without any questioning of the issues dealt with, and therefore, move the board away from the much desired "independence of mind" pointed out by financial supervisors such as EBA and ESMA [12]. Hence, corporate governance recommendations and rules consider that independent external directors are no longer referred to as "independent" if they exceed a certain period of permanence.

In this area, empirical literature shows results in both directions, although there prevails a favorable result regarding permanence on the board (e.g. [30, 36, 37]). For example, Schnake et al. [37], using a period of time before the crisis, obtained a negative and significant association between board tenure and the number of investigations and legal proceedings against the company. Therefore, board member permanence implies better control of the company with a lower level of risk in the decisions of

the board. From a different point of view, but with a similar result, Rutherford and Buchholtz [36] concluded that increasing the permanence of the outside directors is associated with a more frequent exchange of information within the boards. This increased flow of information is an indication of the board's proper functioning, and it is therefore expected to have an impact on better management and control of the company. An interesting study of O'Sullivan et al. [30] used a sample of 150 banks and showed that both CEO tenure and board tenure enhanced bank performance consistently with the "expertise hypothesis". On the other hand, and among the authors supporting the "management friendliness hypothesis" formulated by Vafeas [38], are Vafeas himself, and Bebchuk and Cohen [6].

In recent years, some research detected that board permanence is positive for a company until a point in which the relationship is maintained for a long time and starts to yield negative results. Therefore, and according to this argument, this relationship would have an inverted U-shaped form in respect to different firm value variables of the company (see e.g. [22, 26]). Both articles conclude that the effectiveness of the board directors starts to decrease after approximately nine years of tenure in the position.

EBA and ESMA [12] suggest that firms should continuously monitor the suitability of the board to identify—in light of any relevant new fact—those situations in which such suitability needs to be re-evaluated. One of the competencies that this publication points out is loyalty, defined as: "identifies with the undertaking and has a sense of involvement. Shows that he or she can devote sufficient time to the job and can discharge his or her duties properly, defends the interests of the undertaking and operates objectively and critically. Recognises and anticipates potential conflicts of personal and business interest." This loyalty is only possible after a minimum permanence that enables the identification with the goals set for the mid and long term. Therefore, according to this line of arguments, we enunciate the first hypothesis of the study as follows:

H1. There is a negative association between the board tenure and bank risk.

In general, prior studies that focused on board expertise shows a positive association with risk-taking levels (e.g. [7, 18, 21]). As market supervisors have pointed out that one of the main board weaknesses before the financial crisis was the lack of adequate board experts, this variable has been extensively reviewed by the academia. For example, Fernandes and Fich [18] focused on the financial experience of the banks' outside directors of 398 U.S. entities during 2006–2007, concluding that their financial experience was inversely related to the likelihood of bank failure. Hau and Thum [21] obtained similar conclusions using a sample of German banks in the same period. In particular, they studied the biographical background of 593 supervisory board members from the 29 largest banks and found that low levels of competence are related to bank losses during the financial crisis. Also, Berger et al. [7] focused on board composition in Germany banks through three variables: age, gender, and educational level. The results show that young directors are more likely to increase portfolio risk. Remarkably, the level of education evolved in the opposite direction; when the directors have high academic degrees, such as a PhD, the risk portfolio is

lowered. It is important to mention that these levels of experience and education can contribute to larger board diversity, as advocated by the regulators (e.g. [15, 19]) but, as Van Peteghem et al. [39] pointed out, if these differences are large among the directors, they could generate strong fault lines, which are associated with lower firm performance.

Minton et al. [28] analyzed the relationship between different risk measures and the presence of independent advisors considered as experts in the financial field. However, the results refute the argumentation of the different institutions and some of the literature stating that greater financial experience on bank boards would unequivocally reduce their risk profile. In particular, the presence of financial experts among independent directors is related to greater risk-taking in the period prior to the financial crisis. This positive association with risk-taking was not penalized by markets before the financial crisis, but just when the crisis broke out. Therefore, the conclusions of this work are different depending on the period analyzed, and a negative relationship is only shown during the crisis period. In the previous literature, we have not found precedents of similar results, except for a report from the International Monetary Fund [23] but is not the main objective of that work and the results are not significant, and also from an article devoted to the analysis of the experience of the CEO [17]. To complete this review of the literature on the expertise of directors, we consider it relevant to mention Chen's recent work [10], which concluded that the expertise of the boards in banks increases the likelihood of forced CEO turnover and the succession by an external. The results also show that CEO succession boosts performance and reduces the bank risk-taking.

Based on this literature and the recommendations of regulatory bodies, especially the publication of the EBA and ESMA [12] in which they state that board members should have up-to-date knowledge of the company business and its risks and should be aligned with their responsibilities, we enunciate the second hypothesis of the study as follows:

H2. There is a negative association between board skills and bank risk.

3 Data, Variables and Methodology

The sample consists of 156 banks for the period 2009 to 2017 from 37 countries located in Europe, America, Asia and Africa. The period analyzed in the study comprises the years after the financial crisis because most of the recommendations and guides to overcome the weaknesses of corporate governance have been published after 2008 year. The financial data and corporate governance variables are obtained from Eikon. In addition, we use World Development Indicators Database (World Bank website) to obtain the annual real Gross Domestic Product (GDP) growth rate (*GGDP*).

The objective of the paper is to test the association between both board tenure and board financial expertise on bank risk. Two proxies are used to measure the bank risk: the ratio of non-performing loans (*NPL*), which is the ratio of non-performing loans to total assets, and the loan loss provision ratio (*LLP*), which is the ratio of the bank's loan loss provisions to net interest income. To measure the variables of interest in our analysis, *Board Tenure* and *Board Skills*, we use the definition provided by Eikon database. Specifically, *Board Tenure* is the average number of years each board member has been on the board, and *Board Skills* is the percentage of board members who have either an industry specific background or a strong financial background.

The model also includes five control variables: the size of the bank (*Bank Size*), calculated as the natural logarithm of a bank's total assets (at book value), the total equity book to total assets ratio (*Equity Cap Ratio*), the bank's average growth in total revenues during the last 3 years (*Revenues 3Yr*), the liquidity ratio (*Liquidity Ratio*) measured by cash and trading assets to total assets, and the percentage of annual real GDP growth rate (*GGDP*) to control the influence on the country's risk. Previous empirical research that includes these variables are Laeven and Levine [25], Nguyen [29], Minton et al. [28], among others.

Methodologically, we use a dynamic model that addresses the problem of endogeneity between corporate governance variables and bank financial variables. According to recent empirical corporate governance studies, the most suitable model is the Generalized Method of Moments (GMM) in two steps [3, 11, 34], among others. Thus, this study uses the GMM model, including lagged differenced values as instruments for the equations in levels, and Hansen's test and Arellano and Bond's serial autocorrelation test.

The model specification is as follows:

$$Risk\ measure_{it} = \alpha + \beta_1 Board\ Tenure_{it} + \beta_2 Board\ Skills_{it} + \sum \beta_j X_{it}^{j} + \\ + \sum \beta_k Year_t + \varepsilon_{it}$$

As explained before, the model includes two proxies of bank risk: the ratio of non-performing loans (*NPL*) and the loan loss provision ratio (*LLP*). The independent variables are *Board Tenure* and *Board Skills* and the model also includes a set of control variables, X_{it}, such as *Bank Size*, *Equity Cap Ratio*, *Revenues 3Yr*, *Liquidity Ratio*, *GGDP*. The variable $Year_t$ is a time dummy variable; α denotes the constant, and ε_{it} is the error. Subscript i refers to each bank observed in the sample, while subscript t refers to each observed year.

Table 1 Descriptive statistics

Variables	Obs.	Mean	SD	Min.	Max.
Risk variables					
NPL	1341	1.9641	2.2489	0.0023	20.6757
LLP	1394	0.1973	0.2315	−2.0979	2.1321
Corporate governance variables					
Board tenure	1137	7.2048	3.6392	1.0000	19.6200
Board skills	1227	54.1423	21.6012	5.2600	100.0000
Control variables					
Bank size	1399	8.2212	0.5956	7.0573	9.5773
Equity cap ratio	1398	0.0831	0.0338	0.0082	0.2957
Revenues 3Yr	1404	5.8457	12.9912	−23.8000	101.0367
Liquidity ratio	1399	0.3323	0.1325	0.0774	0.9733
GGDP	1404	2.6480	3.2558	−7.0761	25.5573

The table shows the mean, median, standard deviation, minimum, and maximum values. The dependent variables are bank risk measured by two proxies: *NPL* (Non-performing loans) that is the ratio of non-performing loans to total assets, and *LLP* (Loan loss provision ratio) that is the ratio of the bank's loan loss provisions to net interest income. The Corporate governance variables are *Board Tenure* that is the average number of years each board member has been on the board and *Board Skills* that is the percentage of board members who have either an industry specific background or a strong financial background. The control variables includes: *Bank Size* is the natural logarithm of a bank's total assets; *Equity Cap Ratio* is the total equity book to total assets ratio; *Revenues 3Yr* is the bank's average growth in total revenues during the last 3 years; *Liquidity Ratio* is measured by the cash and trading assets to total assets ratio; and *GGDP* is the percentage of annual real GDP growth rate

4 Descriptive Statistic and Results

Table 1 shows the descriptive statistic of the sample. The average *Board Tenure* is 7.2 years, with a maximum of about twenty years. The majority of countries has developed recommendations and guides to limit the number of year that members can be on the boards in order to be considered as "independent". In fact, the maximum duration ranges from 5 and 7 years in Turkey, China and Russia, up to 12 to 15 years in many European Union countries, such as Belgium, France, Luxembourg, Poland, Spain, Portugal, etc. The average *Board Skills* in the sample is 54.14% but the standard deviation is high.

Table 2 exhibits the sample distribution per region. The majority of the banks are located in North America (22.44%), Europe and Central Asia (26.92%), and Rest of Asia and Pacific (39.10%). Asia is well represented in the sample due to the growing boom of the economy of these countries after financial crisis.

Table 3 shows the bivariate correlation matrix of the variables used in the model. The results indicate a negative and significant correlation between *Board Skills* and *NPL* ($p < 0.01$) and Loan loss provision ratio or *LLP* ($p < 0.01$). The correlation

Table 2 Sample distribution per region

Region	Banks	%
North America	35	22.44
Latin America and the Caribbean	5	3.21
Europe and Central Asia	42	26.92
Rest of Asia and Pacific	61	39.10
Africa and the Middle East	13	8.33
	156	100.00

Table 3 Pearson correlations

	NPL	LLP	Board tenure	Board skills	Bank size	Equity cap ratio	Revenues 3Yr	Liquidity ratio
LLP	0.4515***							
	(0.0000)							
Board tenure	−0.0919***	−0.0557*						
	(0.0023)	(0.0612)						
Board skills	−0.2118***	−0.1400***	−0.1429***					
	(0.0000)	(0.0000)	(0.0000)					
Bank size	−0.1057***	0.0681**	−0.2846***	−0.0085				
	(0.0001)	(0.0111)	(0.0000)	(0.7668)				
Equity cap ratio	0.1069***	0.0173	0.2646***	−0.1428***	−0.4149***			
	(0.0001)	(0.5185)	(0.0000)	(0.0000)	(0.0000)			
Revenues 3Yr	−0.1067***	0.0422	−0.0906***	−0.0059	−0.0814***	0.1019***		
	(0.0001)	(0.1154)	(0.0022)	(0.8370)	(0.0023)	(0.0001)		
Liquidity ratio	−0.2018***	−0.0983***	−0.2105***	−0.0107	0.4549***	−0.2118***	−0.0645**	
	(0.0000)	(0.0002)	(0.0000)	(0.7072)	(0.0000)	(0.0000)	(0.0159)	
GGDP	−0.1381***	−0.2257***	−0.1265***	0.0447	−0.0601***	0.1208***	0.2092***	−0.0497*
	(0.0000)	(0.0000)	(0.0000)	(0.1175)	(0.0245)	(0.0000)	(0.0000)	(0.0633)

The asterisks (***, **, and *) denote significance at the 0.01, 0.05, and 0.10 levels, respectively

between *Board Tenure* and *NPL* is negative and significant *($p < 0.01$)* while the correlation between *Board Skills* and *LLP* is negative but statistically marginal *($p < 0.1$)*.

Table 4 presents the results of GMM models. The result shows a negative and statistically significant association between *Board Skills* and both bank risks variables, *NPL ($p < 0.01$)* and *LLP ($p < 0.05$)*. In contrast, the association between *Board Tenure* and *LLP* is not statistically significant *($p > 0.05$)*. The results reveal that board expertise is important for risk management, at least in the post-crisis period (2009–2017). The board competencies confer quality in monitoring the bank. The results are in line with previous corporate governance studies that deal with board expertise (e.g. [7, 18, 21]). Also, the evidence is aligned with the recommendations of EBA and ESMA [12] that point out the bank board should be able to exert a management control that includes a reduction of bank risk. Consequently, if managers and the

Table 4 Explanatory model for risk measures: NPL and LLP

	Vardep: NPL						Vardep: LLP					
	Coef	P> |t|		Coef	P >|t|		Coef	P> |t|		Coef	P> |t|	
Board tenure	−0.1577	(0.034)	**	−0.1724	(0.013)	**	0.0032	(0.645)		−0.0068	(0.316)	
Board skills	−0.0329	(0.008)	***	−0.0405	(0.004)	***	−0.0027	(0.010)	**	−0.0037	(0.001)	***
Bank size	0.1530	(0.835)		0.2368	(0.760)		0.0437	(0.483)		0.0293	(0.638)	
Equity cap. ratio	0.3676	(0.963)		3.3696	(0.522)		0.2558	(0.755)		0.0447	(0.962)	
Revenues 3yr	−0.0140	(0.223)		−0.0102	(0.416)		−0.0009	(0.250)		0.0000	(0.993)	
Liquidity ratio	−5.2431	(0.019)	**	−5.5989	(0.031)	**	−0.4388	(0.082)	*	−0.4030	(0.079)	*
GGDP	−0.0648	(0.013)	**	−0.1394	(0.001)	***	−0.0142	(0.003)	***	−0.0122	(0.006)	***
const	5.2367	(0.402)		5.5658	(0.396)		0.1058	(0.848)		0.3172	(0.582)	
Year effect	No			Yes			No			Yes		
Hansen	88.19	(0.385)		89.04	(0.361)		103.07	(0.089)		100.31	(0.123)	
AR1 (p-value)	−0.55	(0.583)		−0.54	(0.591)		−0.70	(0.483)		−0.99	(0.323)	
AR2 (p-value)	−0.76	(0.449)		−1.34	(0.182)		−0.84	(0.399)		−0.93	(0.353)	
Number of obs	1073			1073			1100			1100		
Number of groups	154			154			156			156		
Number of instruments	93			101			93			101		

The table shows the GMM model (two steps) with robust fit. Dependent variables are measures of bank risk: *NPL* (Non-performing loans) is the ratio of non-performing loans to total assets, and *LLP* (Loan loss provision ratio) is the ratio of the bank's loan loss provisions to net interest income. Corporate governance variables: *Board Tenure* is the average number of years each counselor has been on the board; *Board Skills* is the percentage of board members who have either an industry specific background or a strong financial background; Control variables are: *Bank Size* is the natural logarithm of a bank's total assets; *Equity Cap Ratio* is the total equity book to total assets ratio; *Revenues 3Yr* is the bank's average growth in total revenues during the last 3 years; *Liquidity Ratio* is measured by the cash and trading assets to total assets ratio; and *GGDP* is the percentage of annual real GDP growth rate. The analysis that incorporates *Year Effect* is presented in columns II and IV. Constraints for endogenous variables are included as follows: their delays (*lags*) t-1 to t-2 as instruments, and instruments in levels. The *p-value* is shown in parentheses. The asterisks (***, **, and *) denote significance at the 0.01, 0.05, and 0.10 levels, respectively

board have adequate knowledge and expertise to understand the business model and bank risk, one of the main corporate governance weaknesses detected before the crisis can be overcome.

5 Conclusions

Corporate Governance mechanisms and board profiles are crucial to mitigate bank risk taking. In fact, previous authors of corporate governance research argue that some inefficiencies in the bank boards failed in the financial crisis in 2007–2008. During the last decade, market supervisors and regulatory bodies have published a set of guidelines and recommendations, with a special focus on board structure and composition, to revise and update corporate governance codes, including the requirements for compliance, monitoring and enforcement.

The objective of this study is to explore the effect of two corporate governance variables, *Board Tenure* and *Board Skills*, on bank risk. The first variable of interest *Board Tenure* is a proxy of knowledge because board permanence provides an understanding of the business model. Then the departure of a board member who has acquired this knowledge implies a loss of talent that can hardly be mitigated. The second variable of interest, *Board Skills,* is a proxy of the specific skills in the financial sector in terms of industry specific background or a strong financial background. Consequently, *Board Skills* is a crucial variable to understand the business model and business risk. Using a sample of 156 banks from 37 countries for the period 2009 to 2017, the association between both the corporate governance variables and the bank risk variable is tested through GMM model.

The results show that board members with extensive experience in finance and banking (*Board Skills*) contribute to decrease the bank credit risk. Consequently, board members with more experience in finance and banking can be more conservative with prudential risk control. The evidence reinforces the results of prior corporate governance studies and the position of several market supervisors and regulator setters. The results also show that the association between *Board Tenure* and bank risk is not statistically significant in all cases. Therefore, board permanence could not be enough to fully understand bank risk. Hence, some recent researchers argue that the association between board tenure and bank performance could be an inverted U-shape instead of linear. Furthermore, some countries and jurisdictions have adopted limits to board tenure members to guarantee their independence.

This study contributes to understand the importance of corporate governance variables to mitigate business risk, following the prescriptions and recommendations. Given the calls for more studies focusing on business risk, this research deals with a strategic industry that changes continuously according to the market and regulatory updates. In this context, bank governance systems and the factors that help to mitigate

risk levels are crucial in the industry and have implications for other industries and across countries. The health of this industry is the driving force behind global economic growth.

Future studies could continue exploring the association between board knowledge and bank performance. *Board Tenure* offers an interesting avenue considering national and local jurisdictions and the potential inverted U-shape. It could be also interesting to examine in detail the "independence of mind", as pointed out by EBA and ESMA [12].

References

1. Adams, R.B.: Governance and the Financial Crisis. International Review of Finance **12**(1), 7–38 (2012)
2. Adams, R.B., Mehran, H.: Bank board structure and performance: evidence for large bank holding companies. J. Financ. Intermed. **21**(2), 243–267 (2012)
3. Ahmad, S., Kodwani, D., Upton, M.: Non-compliance, board structures and the performance of financial firms during crisis: UK Evidence, In: International Finance and Banking Society (IFABS), Risk in Financial Markets and Institutions: New Challenges, New Solutions, 1–3 June (2016). http://oro.open.ac.uk/46837/
4. Arellano, M., Bover, O.: Another look at the instrumental variable estimation of error-components models. Journal of Econometrics **68**(1), 29–51 (1995)
5. Basel Committee on Banking Supervision: Principles for Enhancing Corporate Governance. October 2010. https://www.bis.org/publ/bcbs176.pdf
6. Bebchuk, A., Cohen, A.: The costs of entrenched boards. J. Financ. Econ. **78**(2), 409–433 (2005)
7. Berger, A.N., Kick, T., Schaeck, K.: Executive board composition and bank risk taking. Journal of Corporate Finance **28**(5), 48–65 (2014)
8. Berger, A.N., Imbierowicz, B., Rauch, C.: The roles of corporate governance in bank failures during the recent financial crisis. Journal of Money, Credit and Banking **48**(4), 729–770 (2016)
9. Beltratti, A., Stulz, R.M.: The credit crisis around the globe: Why did some banks perform better? Journal of Financial Economics **105**(1), 1–17 (2012)
10. Chen, Z.: Does Independent Industry Expertise Improve Board Effectiveness? Evidence from Bank CEO Turnovers. International Review of Finance (2020). https://doi.org/10.1111/irfi.12236
11. de Andrés, P., Vallelado, E.: Corporate governance in banking: The role of the board of directors. Journal of Banking and Finance **32**(12), 2570–2580 (2008)
12. EBA and ESMA: Joint ESMA and EBA Guidelines on the assessment of the suitability of members of the management body and key function holders under Directive 2013/36/EU and Directive 2014/65/EU. Final Report. EBA/GL/2017/12 (2017). https://eba.europa.eu/eba-and-esma-provide-guidance-to-assess-the-suitability-of-management-body-members-and-key-function-holders
13. Elyasiani, E., Zhang, L.: Bank holding company performance, risk, and "busy" board of directors. J. Bank. Financ. **60**(1), 239–251 (2015)
14. Erkens, D.H., Hung, M., Matos, P.: Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide. Journal of Corporate Finance **18**(2), 389–411 (2012)
15. European Commission. Corporate Governance in Financial Institutions: Lessons to be drawn from the current financial crisis, best practices. Commission Staff Working Document. SEC(2010) 669 (2010)

16. European Economic and Social Committee. De Larosière Report. ECO/259-EESC-2009-1476 (2009). https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/de-larosiere-report
17. Farag, H., Mallin, C.: The influence of CEO demographic characteristics on corporate risk-taking: evidence from Chinese IPOs. The European Journal of Finance **24**(16), 1528–1551 (2018)
18. Fernandes, N., Fich, E.: Does financial experience help banks during credit crises? Working paper, Drexel University (2009)
19. G30: Toward Effective Governance of Financial Institutions. Washington (2012). https://group30.org/images/uploads/publications/G30_TowardEffectiveGovernance.pdf
20. Haan, J., Vlahu, R.: Corporate governance of banks: A survey. Journal of Economic Surveys **30**(2), 228–277 (2016)
21. Hau, H., Thum, M.: Subprime crisis and board (In-)competence: private vs. Public banks in Germany. Econ. Policy **24**(60), 701–752 (2009)
22. Huang, S., Hilary, G.: Zombie board: Board tenure and firm performance. Journal of Accounting Research **56**(4), 1285–1329 (2018)
23. International Monetary Fund: Risk-taking by banks: The role of governance and executive pay. In: Global Financial Stability Report: Risk-Taking, Liquidity, and Shadow Banking: Curbing Excess While Promoting Growth, Washington (2014). https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Risk-Taking-Liquidity-and-Shadow-Banking-Curbing-Excess-While-Promoting-Growth
24. Laeven, L.: Corporate governance: what's special about banks? Annual Review of Financial Economics **5**(1), 63–92 (2013)
25. Laeven, L., Levine, R.: Bank governance, regulation and risk taking. Journal of Financial Economics **93**(2), 259–275 (2009)
26. Livnat, J., Smith, G., Suslava, K., Tarlie, M.: Do directors have a use-by date? Examining the impact of board tenure on firm performance. Am. J. Manag. **19**(2) (2019). https://doi.org/10.33423/ajm.v19i2.2073
27. Mehran, H., Morrison, A.D., Shapiro, J.D.: Corporate governance and banks: what have we learned from the financial crisis? FRB of New York Staff Report, vol. 502 (2011). https://www.newyorkfed.org/research/staff_reports/sr502.html
28. Minton, B.A., Taillard, J.P., Williamson, R.: Financial expertise of the board, risk taking, and performance: Evidence from bank holding companies. Journal of Financial and Quantitative Analysis **49**(2), 351–380 (2014)
29. Nguyen, P.: Corporate governance and risk-taking: Evidence from Japanese firms. Pacific-Basin Finance Journal **19**(3), 278–297 (2011)
30. O'Sullivan, J., Mamun, A., Hassan, M.K.: The relationship between board characteristics and performance of bank holding companies: before and during the financial crisis. Journal of Economics and Finance **40**(3), 438–471 (2016)
31. OECD: Corporate Governance and the Financial Crisis: Key Findings and Main Messages, (2009). https://www.oecd.org/corporate/ca/corporategovernanceprinciples/43056196.pdf
32. OECD: OECD Corporate Governance Factbook (2019). https://www.oecd.org/corporate/Corporate-Governance-Factbook.pdf
33. Pathan, S.: Strong boards, CEO power and bank risk-taking. Journal of Banking and Finance **33**(7), 1340–1350 (2009)
34. Pathan, S., Faff, R.: Does board structure in banks really affect their performance? Journal of Banking and Finance **37**(5), 1573–1589 (2013)
35. Rajgopal, S., Srinivasan, S., Wong, Y.T.F.: Bank boards: what has changed since the financial crisis? (2019). Available at SSRN: https://ssrn.com/abstract=2722175 or https://dx.doi.org/10.2139/ssrn.2722175
36. Rutherford, M.A., Buchholtz, A.K.: Investigating the relationship between board characteristics and board information. Corporate Governance: An International Review **15**(4), 576–584 (2007)
37. Schnake, M.E., Fredenberger, W.B., Williams, R.J.: The influence of board characteristics on the frequency of 10-K investigations of firms in the financial services sector. Journal of Business Strategies **22**(2), 101 (2005)

38. Vafeas, N.: Length of board tenure and outside director independence. Journal of Business Finance and Accounting **30**(78), 1043–1064 (2003)
39. Van Peteghem, M., Bruynseels, L., Gaeremynck, A.: Beyond Diversity: A Tale of Faultlines and Frictions in the Board of Directors. The Accounting Review **93**(2), 339–367 (2018)

The Shopping Experience in Virtual Sales: A Study of the Influence of Website Atmosphere on Purchase Intention

F. Jiménez-Delgado, M. D. Reina-Paz, I. J. Thuissard-Vasallo, and D. Sanz-Rosa

Abstract The conceptual framework for this research comes from the study of environmental psychology in response to a given environment that condition approach-avoidance behaviors. These three states were pleasure (satisfaction/happiness), arousal (stimulation), and dominance (feeling of control), or PAD. Subsequently, further research applied this theory to the context of a retail setting in order to better understand the influence of a physical store's atmosphere on consumers. The main purpose of this study is to analyze the effects of certain marketing tools and the atmosphere created at a Point-Of-Sales (PoS) on the online shopping experience and purchase intention. The methodology for the work is based on the well-known stimulus-organism-response (S-O-R) model, which has been the subject of numerous consumer behavior studies. The approach for this study is an experiment in which a sample of individuals takes part in a double online shopping experience involving two different scenarios with different dimensions and atmospheres. Despite other existing research has demonstrated in the past the impact of vividness' info and video materials on the purchase intention, the originality of this study is the usability of 3D-dimension as a virtual presence enhancer to demonstrate both, better experience and more realistic shopping behavior closer to the findings we can expect from a consumer on the physical traditional stores.

Keywords Consumer behavior · Multi-channel retailing · Online buying experience · Internet shopping · Online atmosphere · Virtual presence · E-commerce

F. Jiménez-Delgado · M. D. Reina-Paz (✉)
UNED, Madrid, Spain
e-mail: mreina@cee.uned.es

I. J. Thuissard-Vasallo · D. Sanz-Rosa
Universidad Europea, Madrid, Spain

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_11

1 Introduction

Today, there is growing interest in all aspects of how consumers experience brands [26], as well as in the evolution of the new channels consumers are being offered for their commercial transactions with them [29]. This paper focuses on these new online shopping channels and, specifically, the validity and effectiveness of the elements used to stimulate purchases at physical points of sale (merchandising material, in-store displays (ISDs)) when they are deployed in an online shopping environment [2, 13, 14]. Paradoxically, while the shopping experience in a physical environment always takes place in a three-dimensional space, nearly all online retailers, or "e-tailers," choose 2D navigation for their online stores, giving rise to starkly different brand images and experiences [20].

The product category chosen for this study is appliances. Within the broader family of electronics and technological equipment, appliances are the category with the lowest share of online sales (8% in Spain in 2015, according to data from the market research consultancy GfK-Emer). This is because, in the large majority of cases, prospective appliance consumers prefer to use online channels to research their purchase decision, but ultimately perform the purchase/transaction at a physical POS, preferably one where the product they have previously picked out is on display. This effect is known as "Ro-Po" (Research Online, Purchase Offline) or "webrooming" [27]. This stands in contrast to other product categories, such as consumer electronics (audio and video), computers (PCs and printers), and mobile devices (tablets and smartphones). With those products, consumers are more willing to perform all stages of the purchasing process (research, comparison, selection, transaction) without leaving the online channel and, in some cases, even choose to purchase the product online after inspecting it in person at a brick-and-mortar outlet, an effect known as "showrooming" [23].

This paper focuses on "webmosphere" and the impact of interactivity in a highly realistic virtual environment designed to enhance the sense of virtual presence. It also analyses how these features enhance the consumer experience in online stores [4, 18]. Although there is extensive literature in this field in Spain, including on how website design variables influence the consumer shopping experience [21], most of it either compares 2D static elements (images) with 2D dynamic images (videos) [11] or compares 2D static designs (images) with 3D static ones (images). The originality of the present study relies on the fact that it compares a 2D static environment with a highly realistic 3D immersive one [1] that allows subjects to interact freely with the environment, thereby fostering their sense of virtual presence or "being there" [15]. In short, it creates a virtual reality effect, which maximizes the sense of realism, emulating a physical shopping environment and the known effects such environments have on consumer behavior [34].

Consequently, the objective of this paper is to demonstrate the more realistic the on-line store is, the better shopping experience the consumer perceives. The methodology to test the hypotheses is an experiment with real consumers engaging with 2 different e-commerce set-ups. The results, as expected, show the consumer

feels better and more confident on a scenario closer to a real look and feel physical store.

The structure of the paper starts with the theoretical foundations based on the traditional off-line consumer behavior and how can be applied on the new on-line consumer when purchasing online. After the introduction, the paper highlights the key targets as well as the hypotheses to be validated throughout the research. The next sections contain the research method with the sample selection and the model variables and the results of the study. Finally, the conclusion section explains the main model contributions, key implications and the research limitations and future research venues.

2 Theoretical Foundations

The conceptual framework is grounded in the study of environmental psychology, first introduced by [22], who identified three emotional states in response to a given environment that condition approach-avoidance behaviors. These three states were pleasure (satisfaction/happiness), arousal (stimulation), and dominance (feeling of control), or PAD. Subsequently, [28] applied this theory to the context of a retail setting in order to gain insight into how the stimuli of the atmosphere of a physical store influence consumers [28].

Ever since, the model generally used in these types of studies has been the stimuli-organism-response (S-O-R) model, which has been widely developed in the literature on consumer behavior. There is extensive literature on the S-O-R model in a real atmosphere (physical store), based on the study of various stimuli variables in a store or POS (external, interior, design, decoration, and human) and how they affect buyers and sellers and their corresponding responses or behaviors [34]. Since the rise of e-commerce (beginning in 2001), numerous papers have likewise applied the S-O-R model in the context of online retail settings [12], where the store is no longer a physical space but nevertheless seeks to replicate the shopping experience in a virtual environment [6, 13].

The study of consumer behavior in online environments, in which some of the five senses are limited, is a very challenging field in neuromarketing, which has traditionally focused heavily on environments offering a multisensory experience [19]. Although an alternative model exists for addressing the issue of virtual presence and web atmosphere (i.e., the Technology Acceptance Model or TAM [5], in the authors' view it is more focused on the man-machine interaction from the perspective of learning, usefulness, and the ease of use of new technologies [35]. As the present study is more interested in the experience of users as they browse, the S-O-R model seemed more appropriate. Specifically, the chosen model was the S-O-R model as revised by authors such as [3] in keeping with their concept of "webmospherics," which are determined by three variables: Structural/design attributes, Media dimensions and Site layout/dimensions.

Several non-Spanish authors have researched this topic [6, 17]. A more recent study at Oregon State University [30] applied this model to the visual presentation of products at online retailers, using image size and the number of product views as variables.

3 Objectives and Hypotheses

The S-O-R model formerly introduced in the "Theoretical Foundations" section above, is the point of departure for the consumer behavior research we aim to run thought this paper, where we will analyze the impact of specific stimuli influence (WD), under certain control factors (BT and PP), on the consumer perceived shopping experience (EX and NPI) and the purchase intention patronage (PI).

As "Stimuli" variables, 2 different environment scenarios will be selected, Consumers sample participating on this research, will be guided throughout these 2 online shopping frameworks Online 2D and Online 3D (WD).

The "Organism" variables to be analyzed are the Need for Inspection (NPI) and shopping experience Satisfaction (EX) both resulting right after the participation on the research different environments (2D versus 3D)

As "Response" variable the Purchase intention (PI) after shopping experience has been selected and we have also included a tracking record of the 2 control variables: Shopping-Navigation Time spent (BT) and prior product purchase experience (PP) on the research design.

The preliminary hypothesis on the proposed model to be validated throughout the research are:

- Hypothesis 1: A 3D presentation offers an enhanced consumer shopping experience compared to 2D navigation.
- Hypothesis 2: A 3D presentation reduces the need to physically inspect the product.
- Hypothesis 3: A 3D presentation increases the time consumers spend browsing before leaving.
- Hypothesis 4: The more time a consumer spends browsing, the more likely he or she is to purchase the target product (appliance).
- Hypothesis 5: Reducing the need to physically inspect the product increases purchase intention.
- Hypothesis 6: Enhancing the shopping experience increases the intention to purchase the target product (appliance).
- Hypothesis 7: Prior purchase of an appliance on an online channel increases the intention to purchase the target product (washing machine).

Figure 1 shows the initially proposed model with the 7 hypotheses to be tested and how they are positively or negatively related to the analyzed variables (Table 1).

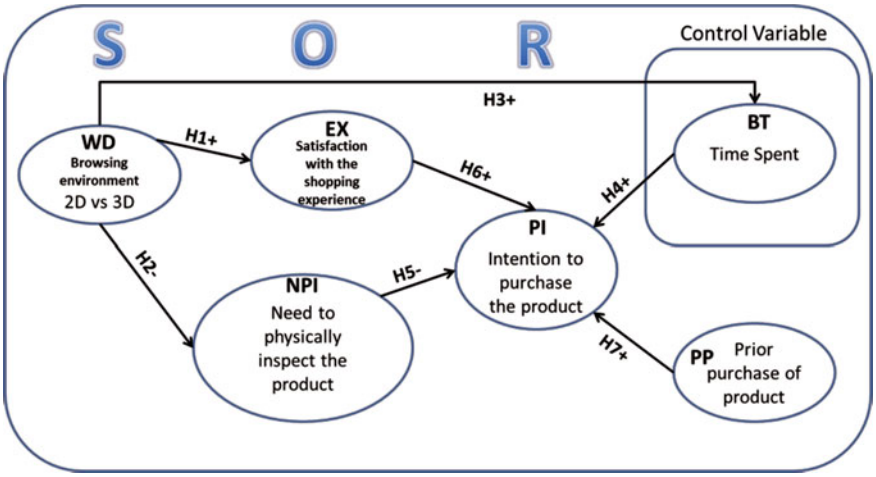


Fig. 1 Proposed model

Table 1 Initial model variables

		Dimensions	
Variables related to the object of study (independent variables)	(WD) Browsing environment dimensions	2D	3D
	(PP) Prior purchase of product	YES	NO
	(BT) Time spent Browsing	Minutes spent browsing	
Variables related to the outcomes (dependent variables)	(EX) Satisfaction with the shopping experience	On a Likert scale (1–5) from "strongly disagree" to "strongly agree"	
	(PI) Intention to purchase the product (washing machine)	On a Likert scale (1–5) from "strongly disagree" to "strongly agree"	
	(NPI) Need to physically inspect the product*	On a Likert scale (1–5) from "strongly disagree" to "strongly agree"	

4 Research Method

The experiment's design is largely based on the work of Suh, including his empirical research on the effects of virtual reality on consumer learning [33] and his research on how "telepresence," achieved through the user interface, affects consumer perceptions in online stores [32].

The research method is based on a remote laboratory experiment consisting of the observation of the behavior of the participating consumers, who connect to a terminal from their homes. The experimental technique is based on a user interface equipped with state-of-the-art Artilux VSO virtual reality technology developed by Quasar Labs, making it possible to record all the movements consumers make during their browsing experience in a simulated purchase process.

Two "simulated" online stores were developed, one with a traditional 2D static browsing interface, like most e-commerce websites today, and the other with a 3D virtual reality interface that allows users to move freely through the store, interacting with merchandising and product display elements, just as they would in a physical store. To ensure the symmetry of the information received by the consumer in each browsing experience (2D and 3D), only visual stimuli were included. The inclusion of other stimuli, such as sound in the case of the 3D virtual-reality browsing experience, was ruled out.

The individuals in the sample were subjected to an initial experience in the 2D environment (store). Data were gathered via an ad hoc survey that they were asked to complete upon finishing browsing. They then moved onto the second experience, entering in the 3D environment (store). Upon finishing browsing, they were asked to complete a second ad hoc survey in order to enable the collection of data on the second experience.

The product chosen for the experiment was a household appliance (specifically, washing machines), as appliances are a virtually-high-experiential (VHE) [33] product category, whose most prominent attributes are their external appearance and features, and which can thus be fully represented through visual stimuli ideally suited for an online store environment such as the one proposed here.

The Partial Least Squares (PLS) regression model was chosen to confirm the hypotheses. We used SmartPLS Software to perform this analysis and estimations.

4.1 Sample Design

A trial test was conducted with 9 users from academia (researchers) to validate the construction of the pilot stores [33] and verify the natural process of the experiment.

Once the initial trial test of the experiment was complete, a sample of 160 individuals linked to the world of marketing professionals (managers and marketing students) was assembled, and the participants were emailed a link they could use to access the experience with the two online stores (2D and 3D).

Although the experiment was designed to allow participants to browse both stores at will and leave whenever they wished, the data on browsing time were recorded with the control variable BT (time spent browsing) and reasonable minimum and maximum "experience" times were set in order to ensure the validity of the results. To this end, 15 participants were eliminated from the total sample as their browsing times fell outside the established time margin (14–673 s). Thus, the final sample on which the statistical analysis was performed consisted of 145 users.

4.2 Statistical Analysis

In order to obtain the final model, we performed the following analysis using a previously published model:

1. We checked the type of relationship (positive or negative) according to the hypothesis established in the original model using a linear or logistic regression analysis.
2. Assessment of the original scheme, measuring the quality by using the statistical model of Partial Least Squares (PLS) allowing the analysis of all the hypotheses as a method for estimating multiple simultaneous correlations between different study variables [9] with a view to validating the consistency of the resulting model. This provided the best fitting model of the parameters studied.

5 Results

The final model was obtained upon the study of the adjusted indexes in different situations of the model. As the ordinal variables presented ≤ 5 categories, a Diagonally Weighted Least Squares method was used to estimate the values of the model.

We analyzed the following statistical indexes:

- Browne's ADF Chi-Square: It is considered a goodness-fit test being acceptable when the p value $>5\%$.
- Comparative Fit Index (CFI): It is a goodness-fit test corrected by the degree of freedom and it's acceptable when the value is between 90% and 95% and good $> 95\%$.
- Root Mean Square Error of Approximation (RMSEA): It is a population level adjusted model in which a value < 0.08 confirms the model reliability.

On a second stage we also run an analysis of the reliability measurement index as it is shown below (Table 2).

Regarding the Reliability measurement index the Cronbach Alpha value is higher than >0.5 for every variable as it is shown on Table 2. On top of that, the composite reliability index are higher than > 0.5 in every case too.

Table 2 Reliability measurement index

	Cronbachs Alpha	R Square	Composite **Reliability**	AVE
Prior Purchase (PP)	0.541681	0.685365	0.700666	0.569239
Shopping Experience (EX)	0.623472	0.624269	0.942112	0.890587
Purchase Intention (PI)	0.532487	0.547939	0.911796	0.838383
Need for Inspection (NPI)	0.600000	0.725985	0.823175	0.701958
Web Navigation Time (BT)	0.535901	0.595474	0.805314	0.676047

Webmosphere—*(WD) "Browsing environment dimensión"*: is a dichotomous variable that identifies the web navigation environment (2D vs. 3D)

Table 3 Discriminant validity

	PP	EX	PI	NPI	BT
PP	**0.754**				
EX	0.043	**0.944**			
PI	0.055	0.740	**0.916**		
NPI	−0.144	0.032	0.044	**0,838**	
BT	0.059	−0.008	0.061	−0.041	**0.822**

*As per diagonal Reading the figures in bold show the AVE square root value

Regarding he Average Variance Extracted (AVE) results are close or even higher than >0.5, as recommended by [8].

The Discriminant Validity, [8] is also validated since for every variable their corresponding AVE is higher than the square of the estimated correlation among them as it is shown on Table 3.

Regarding the R2 coefficients linked to the latent variables regression are also solid value higher than >0.1 in every single case [7].

A deeper analysis across the factors, as it is shown on Table 4, highlights the existing dependence across variables and confirms all the preliminary hypothesis but hypotheses 3 and 4 , that are related to the navigation time (BT), in this case value is below <0.1

As a result of the statistical analysis we end with a validated resulting model, as it is shown below on Fig. 2. The model has been revisited and Navigation Time (BT) variable has been removed out of the original proposed model as hypotheses 3 and 4 could not be statistically validated

The results of the analysis showed that hypothesis 1, on the enhanced shopping experience provided to consumers by a 3D browsing environment as opposed to a 2D one, was supported. The PLS analysis yielded a very solid value of 1.016, thereby supporting this hypothesis, which suggests that the shopping experience in a virtual

Table 4 Discriminant validity

	PP	EX	PI	NPI	BT
PP			0.523		
EX			2.643		
PI					
NPI			−0.937		
BT			**0.002**		
WD		1.016		−0.557	**– 0.005**

*As per diagonal Reading the figures in bold show the AVE square root value

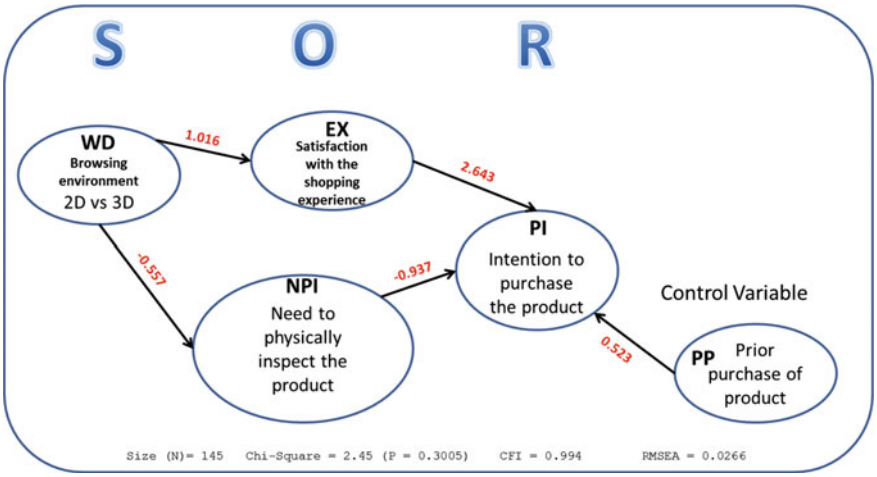


Fig. 2 Resulting structural model

store is enhanced when the level of realism of the browsing is increased through the use of a more dynamic, information-rich environment with higher-quality graphics, which "teleports" consumers to an environment more similar to those they are used to in the world of physical retail [24].

Regarding the second hypothesis, according to the PLS analysis, which yielded a score of −0.557, a 3D presentation does reduce the need to physically inspect the product. The need for physical inspection was found to be smaller in the 3D environment than in the 2D environment, and the difference was statistically significant. This supports the hypothesis that the more information users receive, the better the mental representation they are able to make of the studied object and, thus, the smaller their need to go to a physical retail outlet to inspect the product. Notwithstanding the above, while the enhanced virtual representation of a product was found to mitigate the need that many users still feel to physically see and "touch" a product before making a decision, it did not entirely eliminate it [33].

The hypothesis that a 3D browsing experience increases the time consumers spend browsing before leaving a store, initially formulated based on the expectation that

the more "pleasurable" experience would increase the customer's desire to continue browsing in the 3D environment, was not statistically supported. This could be due to the design of the experiment itself, which established a specific browsing order—first in 2D and then in 3D—that may have conditioned the results for individuals, who experienced the 2D view before the 3D version. It would be interesting to conduct the experiment in parallel with another statistical population that was exposed first to the 3D experience and then to the 2D one in order to test whether the survey results are conditioned by the order of the experiences with each scenario.

Another possible interpretation, based on the negative result of the PLS analysis (−0.005), is that the correlation is precisely the opposite, i.e., that the more realistic 3D browsing experience provides more detail about the product under study, thereby enabling users to get a better idea of the product more quickly and, based on this enhanced mental representation, to make their purchase decision sooner. However, this second interpretation cannot be categorically confirmed either, as the correlation result was, in any case, very weak.

Similarly, no strong correlation was found for fourth hypothesis, which stated that increasing the time customers spend browsing in the virtual store increases the likelihood that they will purchase the target product (appliance). The resulting value (0.002), while positively correlated, was limited. This behavior has already been reported in the world of physical stores, in which numerous studies on shopping duration have shown that simply having customers in the POS for a long time is not enough to increase sales receipts and/or customer satisfaction (especially if the increase in time spent at the store is due to lines or an overly long/forced consumer journey [31]).

The fifth hypothesis, which stated that reducing the need to physically inspect the product increased purchase intention, was supported (−0.937). This outcome was expected in light of the many studies that have measured this correlation in the past [25]. However, of particular interest was the fact that this need for "physical inspection" seemed to be even further reduced through the use of a realistic and highly dynamic representation, such as that provided in 3D, which generates an even stronger sensation of "being there" [32].

Very strong support (2.643) was also found in the 3D scenario for the sixth hypothesis, which proposed a positive correlation between an enhanced shopping experience and the intention to purchase the target product (appliance). This finding suggests that most companies with websites currently designed for classic browsing environments should consider migrating to a more evolved, fluid, and attractive environment [16], able to provide more pleasurable experiences to potential buyers.

The last hypothesis, which proposed that prior experience with online purchases (of any product) has a positive effect on subsequent purchases via the same online sales channel, was generically supported (0.523) in the specific case examined in this study (washing machines). As intuited, once consumers overcome their initial inhibitions regarding the virtual world, they are more likely to continue shopping in online stores in the future [10].

6 Conclusions

The most important technical contribution of this research is the creation of an ad-hoc website as a "test lab." In this regard, support was provided by Quasar Labs, a company specialized in the development of virtual platforms, which reproduced the environment designed for the study in 3D.

Today, some companies have begun to take virtual 3D representation into account as a tool not only for offering customers a reality closer to the product itself, but also for providing them with simulations they can use to see the product in scenarios they create themselves.

In this regard, leading furniture and home décor companies are beginning to successfully test applications that allow customers to create their own space design, tailored to their needs, in an augmented 3D virtual reality environment. This enables them to give customers an almost perfect idea of what the final result will be like once the chosen products (furniture, appliance, decorative objects, etc.) have been incorporated into their particular reality.

Virtual 3D representation entails not only access to a more realistic scenario, closer to the experience of shopping at traditional brick-and-mortar stores, than 2D representation does, with all the implications this entails with regard to enhanced customer experience and purchase intention outcomes, as this study has shown.

As per our research results, the variable that shows a higher correlation directly linked to the Purchase Intention, is the Shopping Experience. This is also very solid correlated to the different shopping environment where it takes place. Therefore, we can assume there is an existing indirectly relation between the shopping environment and purchase intention from the consumer patronage. This supports our hypothesis than the more realistic the virtual environment looks like, the better the shopping experience is, thus the higher purchase intention consumers will have.

The outcome from our research also supports this hypothesis since the results have validated the less need for consumer physical inspection exits that will result on higher consumer purchase intention. Since there is a solid negative correlation between 3D environment and Need for physical inspection we can also validate the indirect relation between the virtual environment chosen and the final purchase intention.

Manufacturers/vendors of home appliances or other similar products featuring this research, should consider that selling those products under a 3D virtual reality scenario will have a positive impact on the consumer shopping, compared to less realistic traditional web look and feel.

Regarding research limitations we may highlight the experiment was conducted in "laboratory" conditions, i.e., using dummy/mock online environments/stores rather than actual e-tailer websites, due to the lack of e-tailers that have incorporated 3D design into their online stores.

As a second limitation the questionnaire measured only consumers' "purchase intentions," which influence their expected behavior; however, this influence cannot

be verified with a "real" purchase, in which they would have to complete the proposed purchase making an actual payment.

There is a last limitation since. The study focuses exclusively on durable appliances, leaving many other consumer good product categories or families (staple goods, FMCGs, fashion and apparel, etc.) unaddressed.

Finally and as future research guideline, we should include expanding the study to encompass a multichannel environment, expanding the study to include other product categories and types, and studying the impact of social media on consumer behavior in a physical environment and of how activity and visibility on social media influence the offline channel.

References

1. Baek, E., Choo, H.J., Yoon, S.Y., Jung, H., Kim, G., Shin, H., Kim, H.: An exploratory study on visual merchandising of an apparel store utilizing 3D technology. J. Global Fashion Market. **6**(1), 33–46 (2015)
2. Breugelmans, E., Campo, K.: Effectiveness of in-store displays in a virtual store environment. J. Retail. **87**(1), 75–89 (2011)
3. Childers, T.L., Carr, C.L., Peck, J., Carson, S.: Hedonic and utilitarian motivations for online retail shopping behavior. J. Retail. **77**(4), 511–535 (2002)
4. Chin, C., Swatman, P.: The virtual shopping experience: using virtual presence to motivate online shopping. Australasian J. Inf. Syst. **13**(1) (2005)
5. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. **35**(8), 982–1003 (1989)
6. Eroglu, S.A., Machleit, K.A., Davis, L.M.: Atmospheric qualities of online retailing: a conceptual model and implications. J. Bus. Res. **54**(2), 177–184 (2001)
7. Falk, R.F., Miller, N.B.: A Primer for Soft Modeling. University of Akron Press, Akron, OH (1992)
8. Fornell, C., Larcker, D.: Evaluating structural equations models with unobservable variables and measurement error. J. Market. Res. **18**(1), 39–50 (1981)
9. Gabisch, J.A., Gwebu, K.L.: Impact of virtual brand experience on purchase intentions: the role of multichannel congruence. J. Electron. Commer. Res. **12**(4), 302 (2011)
10. Gefen, D., Karahanna, E., Straub, D.W.: Inexperience and experience with online stores: the importance of TAM and trust. IEEE Trans. Eng. Manag. **50**(3), 307–321 (2003)
11. Gurrea, R., Sanclemente, C.O.: El papel de la vivacidad de la información online, la necesidad de tocar y la auto-confianza en la búsqueda de información online-offline. Revista Española de Investigación en Marketing ESIC **18**(2), 108–125 (2014)
12. Ha, Y., Lennon, S.J.: Online visual merchandising (VMD) cues and consumer pleasure and arousal: purchasing versus browsing situation. Psychol. Market. **27**(2), 141–165 (2010)
13. Ha, Y., Lennon, S.J.: Consumer responses to online atmosphere: the moderating role of atmospheric responsiveness. J. Global Fashion Market. **2**(2), 86–94 (2011)
14. Ha, Y., Kwon, W.S., Lennon, S.J.: Online visual merchandising (VMD) of apparel web sites. J. Fashion Market. Manag. Int. J. **11**(4), 477–493 (2007)
15. Hassouneh, D., Brengman, M.: Retailing in social virtual worlds: developing a typology of virtual store atmospherics. J. Electron. Commer. Res. **16**(3), 218 (2015)
16. Im, H.J.: If i can't see well, i don't like the website: website design for both young and old. J. Korean Soc. Clothing Textiles **38**(4), 598–609 (2014)
17. Kim, H., Lennon, S.J.: E-atmosphere, emotional, cognitive, and behavioral responses. J. Fashion Market. Manag. Int. J. **14**(3), 412–428 (2010)

18. Kim, J.H., Lennon, S.J.: Information available on a web site: effects on consumers' shopping outcomes. J. Fashion Market. Manag. Int. J. **14**(2), 247–262 (2010)
19. Krishna, A.: An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. J. Consum. Psychol. **22**(3), 332–351 (2012)
20. Kwon, W.S., Lennon, S.J.: Reciprocal effects between multichannel retailers' offline and online brand images. J. Retail. **85**(3), 376–390 (2009)
21. Lorenzo, C., Constantinides, E., Gómez, E., Geurts, P.: Análisis del consumo virtual bajo la influencia de las dimensiones constituyentes de la experiencia web. Estudios sobre consumo **84**, 53–65 (2008)
22. Mehrabian, A., Russell, J.A.: An Approach to Environmental Psychology. MIT Press (1974)
23. Melis, K., Campo, K., Breugelmans, E., Lamey, L.: The impact of the multi-channel retail mix on online store choice: does online experience matter? J. Retail. **91**(2), 272–288 (2015)
24. Mollen, A., Wilson, H.: Engagement, telepresence and interactivity in online consumer experience: reconciling scholastic and managerial perspectives. J. Bus. Res. **63**(9), 919–925 (2010)
25. Peck, J., Childers, T.L.: Individual differences in haptic information processing: the "need for touch" scale. J. Consum. Res. **30**(3), 430–442 (2003)
26. Pine, B.J., Gilmore, J.H.: Welcome to the experience economy. Harvard Bus. Rev. **76**, 97–105 (1998)
27. Puccinelli, N.M., Goodstein, R.C., Grewal, D., Price, R., Raghubir, P., Stewart, D.: Customer experience management in retailing: understanding the buying process. J. Retail. **85**(1), 15–30 (2009)
28. Robert, D., John, R.: Store atmosphere: an environmental psychology approach. J. Retail. **58**, 34–57 (1982)
29. Shankar, V.: Shopper marketing 2.0: opportunities and challenges. In: Shopper Marketing and the Role of In-Store Marketing Review of Marketing Research, Vol. 11, pp. 189–208. Emerald Group Publishing Limited (2014)
30. Song, S.S., Kim, M.: Does more mean better? an examination of visual product presentation in e-retailing. J. Electron. Commer. Res. **13**(4), 345 (2012)
31. Spies, K., Hesse, F., Loesch, K.: Store atmosphere, mood and purchasing behavior. Int. J. Res. Market. **14**(1), 1–17 (1997)
32. Suh, K.S., Chang, S.: User interfaces and consumer perceptions of online stores: the role of telepresence. Behav. Inf. Technol. **25**(2), 99–113 (2006)
33. Suh, K.S., Lee, Y.E.: The effects of virtual reality on consumer learning: an empirical investigation. MIS Q. 673–697 (2005)
34. Turley, L.W., Milliman, R.E.: Atmospheric effects on shopping behavior: a review of the experimental evidence. J. Bus. Res. **49**(2), 193–211 (2000)
35. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: Four longitudinal field studies. Manag. Sci. **46**(2), 186–204 (2000)

European Mobile Phone Industry: Demand Estimation Using Discrete Random Coefficients Models

Kyung B. Kim and José M. Labeaga

Abstract In this paper, we combine the literature from marketing and industrial economics about the mobile industry to build up the basics for estimating heterogeneous demand models. We also had a look at the general statistics for better understanding the market and proposing logit and Random Coefficient Logit (RCL) models to estimate the demand equations. Based on the observation on the market and the aggregate format of the market intelligence data, we follow the methodology of [1, 2] since we consider the RCL an attractive approach for estimating discrete purchases demand from aggregate data. The significant contribution of our approach is to provide a practical method for estimating price elasticities for demand systems involving many similar datasets using market intelligence data. The applications to other related electronics industry would be quite simple, considering the similarities in characteristics, perishable nature of selling prices, heterogeneous demand, and demographic differences in tastes.

Keywords Random coefficients logit · European electronics market · Aggregated intelligence data

1 Introduction

As shown in much Industrial Organization (IO) literature and marketing literature, estimating demand systems is a crucial task for answering many managerial and policy-related questions. We are entering into the world of the internet of things (IoT); for the time being, smartphones are at the center of this world due to connectedness and infrastructures support smartphones to connect with other machines.

K. B. Kim (✉) · J. M. Labeaga
Department of Economic Theory and Mathematical Economics, UNED, Paseo Senda del Rey, 11, Madrid 28040, Spain
e-mail: kkim1@alumno.uned.es

J. M. Labeaga
e-mail: jlabeaga@cee.uned.es

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_12

Therefore estimation of demand systems of the mobile phone sector is relevant since it will continue to constitute or evolve into one of the most critical parts in the future. That is one of the reasons why the mobile phone sector is a field with high innovation and customer churn, where the struggle for market share has been highly fierce. s Discrete choice models have a long tradition in empirical research. They were developed initially by [3], starting with a rigorous theoretical model in which agents maximize utility (or profit) functions to analyze consumer choices with micro-level data. Here there are no "taste" impacts on consumers; this keeps the "representative agent" concept (henceforth "structural approach"). The observable characteristics are the same across all consumers. Though this hypothesis is sometimes reasonable and sometimes counterintuitive depending on the context, it makes the model simpler. In the classical work of [1] (BLP hereafter), they relaxed this hypothesis and introduce different tastes among consumers, i.e., unobserved heterogeneity in consumer valuations of product characteristics. They propose an RCL demand model that can be estimated with aggregate data on sales, prices, and product characteristics. The BLP method is a way to predict demand curves, an idea that lends itself to testing theories of IO. The beauty of BLP is that it can combine a variety of new econometric methods with many optimization techniques, and it is flexibly adaptable to other contexts. What we get using this structural approach is to ensure a self-consistent set-up and the ability to test much more elegant hypotheses about economic behavior at the cost of complexity and absence of robustness to specification errors.

Since random coefficients account for unobserved heterogeneity in consumer valuations of product characteristics, they create flexible substitution patterns between products. Since [2] (Nevo hereafter), the aggregate RCL model has become increasingly popular in IO, marketing, international trade, management, environmental economics, and many other areas of economics. BLP's RCL model generates a nonlinear aggregate market share system. Then, BLP shows how to invert the system to solve for the product-specific unobservables and estimate the model using Generalized Methods of Moments (GMM) estimators. In recent years, several papers have documented numerical difficulties with BLP's approach and attempted to formulate solutions, often based on Monte Carlo studies. This potential computational burden has brought about alternative estimation methods and computationally light procedures. Dube et al. [4] assess the performance of BLP's contraction mapping, which is a Nested Fixed Point (NFP) algorithm to invert the market share system. They propose an approach called Mathematical Programming with Equilibrium Constraints (MPEC). This algorithm virtually eliminates the inner loop contraction mapping and instead minimizes the GMM objective function subject to the market share system as constraints.

Many papers in marketing and IO deal with the theoretical basis of a company's product and pricing strategy. Anderson et al. [5] use the discrete choice model to describe various product firms in the context of spatial price discrimination. Moreover, they show that many strong properties of the standard homogeneous goods case are no longer valid. In particular, the social optimum as a market equilibrium is no longer sustainable unless products are either identical or else very different. Anderson et al. [6] presents the multinomial logit model to describe the demand of

a population of heterogeneous consumers for a set of differentiated goods. Gallego and Wang [7] study companies selling multiple differentiated substitutable products and customers whose purchase behavior follows a nested logit structure. Customers make purchasing decisions sequentially: they first select a nest of products and subsequently purchase a product within the selected class. It shows that each nest has an adjusted nest-level markup that is nest-invariant, which further reduces this problem to a single variable optimization of a continuous function over a bounded interval and provides conditions for this function to be unimodal. We also use those results to simplify the oligopolistic price competition and characterize the Nash Equilibrium (NE). Armstrong and Vickers [8] provide a relatively simple necessary and sufficient conditions for a multi-product demand system that can be generated using a discrete choice model with unit demands.

While many academic types of research deal with the mobile industry and mobile operators' and manufacturers' marketing strategy from the business perspective, many studies still focus on network operators due to their strong influence on the value chain. Suryanegara and Miyazaki [9] examines the Japanese mobile industry focusing on how mobile operators replied to market and technological changes, emphasizing how the brand image made created the most economic value and convergence on technological evolution. Freire Kastner [10] explores, in the context of UK mobile telecommunication companies, theoretical justifications of pricing, the current pricing practices in real business, as well as the connection to the obtained empirical evidence. He states that price is a crucial element for firms as it can influence demand and, consequently, profits. He emphasizes the importance of information in pricing behavior about customers and competitors. Bhargava and Gangwar [11] compare pricing strategies of post-paid dominated mature mobile markets such as in the US with prepaid dominated growing markets such as India, and they develop a model to explains the pricing strategies employed in India and propose evolutionary steps to adapt their strategy as the market matures. Kim and Lee [12] study the South Korean mobile market and investigate the key drivers that establish and maintain customer loyalty to mobile telecommunications service providers. Bidyarthi et al. [13] perform a case study of Nokia in India and show how pricing strategy can strengthen its market potential but also can cause room for new entries of competitors. Prasad and Sahoo [14] perform value chain analysis in the mobile phone industry and find out that under the given scenario, developing a core competence shall provide an immense boost to the companies' performance from the costing as well as a marketing point of view. Dedrick [15] uses quantitative analysis of value captured by firms in the supply chain of mobile phones. They find even though telecom operators still have a significant portion of the gross profit in comparison to other handset manufacturers, when it comes to operating profit level, brand-name handset manufacturers capture similar financial value from each phone than any of its suppliers. Therefore, many handset manufacturers try to enhance brand recognition via R&D and marketing. Kraemer et al. [16] show that Apple makes good examples of making profits and reinvesting in these sectors, trying to capture a large share of value from brand awareness. While these products, including those produced in China or other developing countries, the main benefits are reinvested in its product design, software

development, product management, and marketing. However, the main prerequisite of investment in both R&D and marketing is profit; therefore, many manufacturers strive to find room for more gains by reducing costs and increasing selling prices. Therefore, cost reduction via production optimization, streamlining processes, and logistics is an excellent subject in many electronic companies. However, companies have a great motivation to optimize their pricing strategies, which will allow them to reinvest and survive in the industry.

In this paper, we use readily available market intelligence data with information on consumer choice, and we show how they can help to identify demand parameters in a widely used class of differentiated product demand models. The demand framework in our paper follows the setup and method [1, 2], using GMM estimators of the aggregate RCL model. We have market intelligence data containing aggregate choices, prices, types, countries, and features of the brands. We use very common setups and structures of many random utility models in IO papers and adopt many aspects of the demand models of BLP/Nevo. In these models, products are described as bundles of characteristics, and consumers choose the product that maximizes utility derived from product characteristics. We seek to uncover demand parameters so that we can obtain a detailed analysis of past events to make realistic predictions.

Our primary purpose is to estimate demand systems for the mobile phone industry using market intelligence data from various sources. To answer any question in marketing and IO, we need to understand how consumers behave and make purchase decisions among many goods or services as a function of the market and individual characteristics. Estimating the underlying parameters will allow us to get many insights, as price elasticities. We like to show that it is possible, using aggregated data, to capture critical features of the corresponding distribution of consumers' willingness to pay; in other words, we want to estimate the proper demand system. Even though estimating and evaluating demand systems is itself a fascinating subject, it becomes even more interesting since it is necessary to obtain related parameters to get those demand systems. Moreover, this process is often used in various fields to answer further questions to provide a more comprehensive picture.

We want to estimate elasticities of demand (d) and supply (s), assuming linear relationships (we will relax this condition later). The implication is that we want to estimate β, γ are the demand parameters or individual-specific taste coefficients. θ , λ are the supply parameters, x are observed product attributes (of brand j in market m), p is the price and ϵ_d and ϵ_s are demand and supply idiosyncratic shocks:

$$d = \beta p + \gamma x + \epsilon_d,$$
$$s = \theta p + \lambda x + \epsilon_s.$$

The main issue of this set-up is endogeneity. Variables affecting the supply shock ϵ_s or the demand shock ϵ_d will also affect the equilibrium price. Therefore, p is endogenous in both equations. We need at least an instrumental variable, changing prices but independent of the demand curve, to estimate demand. Our basic framework of modeling and estimating follows the setup and method of BLP and Nevo. It uses a GMM estimator for the RCL model. In principle, BLP or RCL does not only

solve endogeneity. Although the basic idea goes around an IV-logit, it allows for a more general model without assuming linear demand and supply. However, in a more complicated demand system, we need to be more careful with the choice and use of the instruments. BLP demand function has its base in a Probit choice model (there is a continuation of agents receiving product specific shocks that determine their preferences). When integrating over agents, we arrive at a logistic demand system. This demand system is typically based on firms having linear cost functions and playing a differentiated Bertrand pricing game. For each guess of the parameters of the model, we should compute the probability for each price-quantity combination, as a function of observed covariates. Since we do not have exact solutions, we calculate the maximum likelihood at each outcome using GMM instead of 2SLS to estimate the parameters. The data contains aggregate choices, prices, types, countries, and features of the brands. Therefore, we think that the BLP setting is an excellent choice since its design serves to estimate demand in differentiated product markets using aggregate data.

Estimation of the parameters of a demand system in the mobile phone industry will provide us with a better understanding of the market and behavior of consumers. This paper adjusts a dynamic demand model in conjunction with important product characteristics that can change consumers' decisions in the market for mobiles. So, it contributes to the existing marketing and empirical IO literature on dynamic, durable goods in three ways. First, we use demographic data for random interactions for heterogeneity to give "taste" impacts and individual abnormalities. We relax the heterogeneity hypothesis of structural models, at the cost of making the solution more difficult. We show that the lack of microeconomic data can be dealt with by alternative approaches, and we employ demographics obtained and simulated from macroeconomic data. The BLP method that we use captures more abundant forms of product differentiation, as well as a more flexible distribution of consumer heterogeneity. Second, wildly available market intelligence data is used to estimate the complex demand structure in the mobile phone industry. As we will see in Sect. 2, from the viewpoints of importance, policy and market changes, profit and pricing, estimating demand systems of the mobile phone industry is an essential and meaningful field. Since our data contains the price and non-price related characteristics of cell phones, we would be able to get information about substitution patterns between products and success factors in the current market. Finally, we make use of cost data from the mobile phone industry as instruments to deal with price endogeneity and guarantee consistent estimation of the parameters. By instrumenting price using 2SLS (and GMM), we do not need to obtain the marginal costs from the model. We discuss the importance of handset manufacturers and handset prices as a crucial factor of success and provide a useful tool for policymakers. Moreover, many market intelligence data are available these days since PoS data, and big transaction data are available. Whereas traditional profit/non-profit agencies are publishing aggregate data, many of them are open source on the aggregate level. Therefore we also think that it is an interesting academic question of whether consumer heterogeneity can be recovered from combinations of many forms of aggregate data and micro-economic demographics.

The rest of the paper is as follows. In Sect. 2, we present the history and current trends of the mobile industry. Section 3 displays the model used to simulate the industry structure. Section 4 reports the data used, including short summaries of the raw data set, joint with the theoretical framework, and Sect. 5 provides results, discussions, and implications. Section 6 reviews study findings and concludes the paper with theoretical and managerial implications, limitations, and directions for future research.

2 Mobile Phone Industry

Network operators around the world used to play the leading role in the mobile phone industry up to the mid-2000s, mainly due to barriers to entry. They used to specify everything from the hardware to the applications and services included on the mobile handsets they sell [17]. Consequently, mobile manufacturers played a minor role, and they need to support these requirements and personalize devices for individual operators effectively. Therefore, the primary studies about the mobile industry have focused on network operators and consumer behavior. After the advent of iPhone and Android phones in the 2010s, mobile operators, as well as mobile manufacturers, are facing a new phase of the competition. Fully maturing mobile industry is now facing new changes and challenges. Due to economic policies favoring competition, Mobile Virtual Network Enablers (MVNE), and similar service quality as the infrastructures get better in the developed world, mobile operators are losing oligopolistic power to get more customers. Moreover, the arrival of Apple and Android smartphones made manufacturers' role more crucial in terms of competition. Even if network operators still hold the leading position in this battle of the brand, the importance of mobile manufacturers in this game is increasing and already influencing decisions of operators using their proper market power. Therefore, cooperation and competition between network operators and manufacturers are becoming critical factors in this industry.

Mergers and separations

The main reason is increasing competition due to merges and separations ([18]), complete or partial, separation options of the ownership [19]. (1) Legal separation (different legal entities under the same ownership); (2) business separation with different governance arrangements; (3) business separation with localized incentives; (4) virtual separation, and (5) separation through the creation of a wholesale accounting division. These separations can be divided into two dimensions:

1. **Horizontal**: More competition in the mobile network business and less differentiation in service quality by network providers. In a recommendation to national telecom regulators issued by the European Commission, they called on public mobile telephone networks in 2003 to increase the competitiveness of the market for wholesale access. Consequently, in several countries, including Ireland

and France, policymakers made new regulations forcing operators to open their network to mobile virtual network operators (MVNOs).[1]

2. **Vertical**: As vertical separation increases due to high competition (price as well as non-price), operator service quality (non-price), as well as price, is becoming less differentiable [18].

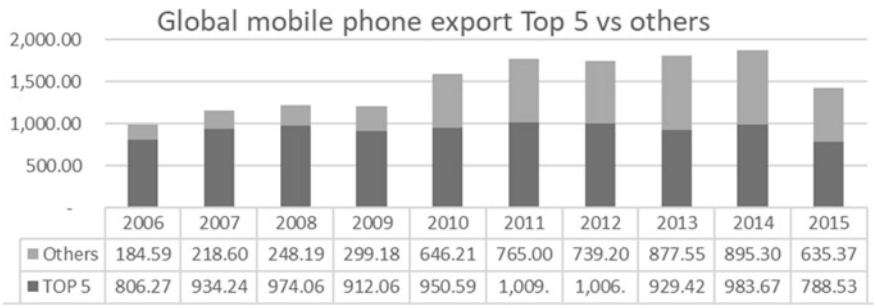
Market growth and competition

Over the last ten years, the mobile phone market is growing fast, and the sales reached almost 2 billion euros per year, only in the European Union (EU). Consequently, despite high entry barriers due to technology and capital-intensive requirements, more and more manufacturers are joining the market to gain market share. There were winners and losers in this industry from the beginning of the new millennium. However, after 2010, the competition is getting fierce. The main reason is due to Chinese manufacturers who are now well-equipped with high technology and experiences from their big domestic market, targeting the entire world. Even if growing seems to be stabilized and the market appears to be mature in 2015, expected demand from emerging economies and spillover effects (making smartphone facilitating related gadgets in the future) make this industry still attractive. For many manufacturers, the mobile phone market is not a cash cow or profit center anymore because of this ever-growing fierce competition. However, it is difficult to abandon the mobile market due to (1) big market size, (2) high sunk costs for investment, and (3) spillover effect to other related electronics future areas in the world of the IoT. In a perfectly competitive market like commodities, the microeconomic theory suggests that manufacturers cannot find room for profit. Therefore, a mobile phone manufacturer's long-term strategy for survival should be differentiation: (1) investing in R&D to make their product unique, and (2) investing in marketing to improve the perceived value of their products and to build a differentiating brand image for consumers to avoid competition, namely increasing brand awareness, product exposure (PR).

Price and profit

The globally harsh competition, mainly from China, has made the mobile market a red ocean, cutting down average consumer prices. In the EU telecommunications service, prices fell yearly by an average of 11–13% between 2006 and 2010 (in comparison, fixed-line prices fell by only 5% per year from 1998 to 2010). When it comes to mobile manufacturers, figures for 2015 suggest that only two leading mobile manufacturers (Apple and Samsung, 92% and 15%, respectively) made a profit from Mobile business, indicating other players account for −7% loss. Figure 1 shows the export proportion of the prominent five mobile manufacturers vis-à-vis the rest. It shows that the market share of 5 manufacturers accounted for more than 80% in 2006, but it became less than 50% after 2013.

[1] See also *La lecture par l'Autorité de régulation des télécommunications de l'article L.1425-1* (2005). An MVNO is a wireless communications services provider that does not own the wireless network infrastructure over which the MVNO provides services to its customers.



	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Others	184.59	218.60	248.19	299.18	646.21	765.00	739.20	877.55	895.30	635.37
TOP 5	806.27	934.24	974.06	912.06	950.59	1,009.	1,006.	929.42	983.67	788.53

Fig. 1 Global mobile phone market (big 5 vs. others)

Price discrimination is one of the most common forms of marketing practices and one of a direct way of regaining profit. In the electronics industry, the majority of manufacturers are still using cost-plus or competition-based pricing mainly due to the convention of being under high price competition even though many leading companies are now learning about the importance of value-based pricing and adopting it in their company [20]. There could be many ways of completing other manufacturers, and they cost reduction in production, economies of scales, efficiency improvement in the organization by reducing fixed costs and logistics. However, all of them only can be achieved via significant efforts, investments, and continuous improvements. All those goals can only be met through long-term efforts. The easiest way of increasing the short-mid-term profit of products in production or a near-production phase would be profit optimization. However, with the ever-growing competition, consumers may switch to alternatives if they notice this seller surplus. Thus, consumers will choose other options having higher customer values which will give high pressure again on the manufacturers and operators to invest more in R&D and marketing for innovation and better productivity as well as better market awareness.

3 The Model

Since our main objective is to estimate a complete demand model, we start with the set-up of the well-known classical aggregate logit model; then, we turn to the aggregate RCL model. We assume a set of markets, $m = 1, \ldots, M$, where the same set of products, $j = 1, \ldots, J$, is available in each market (or/and time). We denote not-purchasing by $j = 0$ and individuals by $i = 1, \ldots, N$ (N is large).

We express the utility of consumer i from purchasing product j in market m as:

$$u_{ijm} = x_{jm}\beta_{ijm} + \alpha_i(y_i - p_{jm}) + \xi_{jm} + \epsilon_{ijm}, \tag{1}$$

where x_{jm} are observed product attributes of brand j in market m; p_{jm} is the price of purchasing product j in market m; ξ_{jm} are unobserved product characteristics (demand

shifters different by brands and markets); y_i is the income of the consumer i; and ϵ_{ijm} is an idiosyncratic shock assumed to be Type I extreme value distribution, independently and identically distributed across products, consumers, and markets. The other terms of (1) are parameters, which can be defined: α_i is i-consumer's marginal utility of income, and β_{ijm} are demand parameters or individual-specific taste coefficients (varying by individuals, product, and markets). Finally, for identification purposes, we assume $u_{i0m} = \epsilon_{i0m}$.

As mentioned, x_{jm} is a vector of observed product attributes including a constant, and β_i represents the taste of consumer i, and it is assumed to follow a distribution $F(\cdot; \theta)$, where θ is a K-vector of parameters to be estimated. As expressed in (2), we can make consumer preferences vary with individual characteristics, as in [1] and [2], by introducing the concepts of observed "demographics", D_j, and "unobserved" tastes, v_j. Collecting terms, the model can be expressed in matrix notation as (2):

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi\, D_i + \Sigma v_i, \qquad \text{with} \quad v_i \sim P_v{}^*(v), \quad D_i \sim \hat{P}_v^*(D)\,, \tag{2}$$

where $P_v{}^*(v)$ is a parametric distribution and $\hat{P}_v^*(D)$ is a nonparametric distribution known from other data sources. If we define an outside good and normalize its utility to zero, we can split vector θ into the linear parameters $\theta_1 = (\alpha, \beta)$ and non-linear parameters $\theta_2 = (\Pi, \Sigma)$, and we can combine (1) and (2) to obtain the following expression:

$$u_{ijm} = \alpha_i y_i + \delta_{im}(x_{jm}, p_{jm}, \xi_{jm}; \theta_1) + \mu_{ijm}(x_{jm}, p_{jm}, v_i; \theta_2) + \epsilon_{ijm}, \tag{3}$$

and

$$\delta_{im} = x_{jm}\beta - \alpha p_{jm} + \xi_{jm}, \qquad \mu_{ijm} = \left[-p_{jm}, x_{jm}\right](\Pi D_i + \Sigma v_i),$$

with δ_{im} being an average utility (common to all customers), $\mu_{ijm} + \epsilon_{ijm}$ a mean-zero heteroscedastic deviation from average utility capturing the effects of the random coefficients, and the first term $\alpha_i y_i$ vanishes. Consumers choose the option of giving the highest utility. BLP illustrates that, by specifying ϵ_{ijt} as a Type I extreme value distribution, parameters in (3) can be estimated. From the distributional assumption of ϵ_{ijt}, the probability of consumer i purchasing product j in market m is given by:

$$\frac{\exp\left(x_{jm}\beta_i + \xi_{jm}\right)}{1 + \sum_{j'=1}^{J} \exp\left(x_{j'm}\beta_i + \xi_{j'm}\right)}, \tag{4}$$

where $x_j \equiv (x'_{1m}, \ldots, x'_{Jm})$ and $\xi_m \equiv (\xi'_{1m}, \ldots, \xi'_{Jm})$. The model aggregate market share function integrates over the consumer-specific choice probabilities, where we let $dF(\beta_i; \theta)$ denote the population distribution of consumer heterogeneity.

The prediction of the market share of product j in market m is then:

$$s_j\left(x_m, \xi_{im}; \beta_i\right) = \int \frac{\exp\left(x_{jm}\beta_i + \xi_{jm}\right)}{1 + \sum_{j'=1}^{J} \exp\left(x_{j'm}\beta_i + \xi_{j'm}\right)} dF(\beta_i; \theta). \tag{5}$$

Often, the evaluation of this integral requires simulation. If we generate β_i for $i = 1, \ldots, N$:

$$s_j\left(x_m, \xi_{im}; \beta_i, N_s\right) = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\exp\left(x_{jm}\beta_i + \xi_{jm}\right)}{1 + \sum_{j'=1}^{J} \exp\left(x_{j'm}\beta_i + \xi_{j'm}\right)}. \tag{6}$$

Moreover, the price elasticities of the market shares are defined by Eq. (5), and we can get them from the aggregate logit model:

$$\eta_{ikm} = \frac{\partial s_{jm}}{\partial p_{kt}} \frac{p_{km}}{s_{jm}} = \begin{cases} -\alpha p_{jm}(1 - s_{jm}), & \text{if } j = k, \\ \alpha p_{km} s_{km}, & \text{otherwise.} \end{cases} \tag{7}$$

So far, we have reviewed the concept of the underlying discrete choice random utility model for each good with extremum distribution of error leading to a logit probability of purchase. We now turn to elasticities in (7). This implies some potential issues when the distribution of ϵ_{ijt} is assumed to be Type I extreme value. First, since in most cases the market shares are small, the factor $\alpha p_{jm}(1 - s_{jm})$ is nearly constant; this provides own-price elasticities proportional to price. Therefore, the lower the price, the lower the absolute value of the elasticity, meaning the lower-priced brands would have a higher markup when priced based on a basic pricing model. This would be only true if the marginal costs of cheaper brands are lower (not just in absolute value, but as a percentage of price) than that of a more expensive product. The second problem, which has been stressed in the literature, concerns cross-price elasticities. If many products with similar or distinctive characteristics have similar market shares, then the substitution from one towards another will always be the same, regardless of the similarities in characteristics. In other words, the logit model restricts consumers to substitute towards other brands in proportion to market shares, regardless of characteristics. In general, this is called *Independence of Irrelevant Alternatives* (IIA) property, i.e., how this structure implies weird substitution patterns for goods which might be close to each other due to the iid assumption of the shocks. Since the problems with cross-price elasticities come from the iid structure of the random shock, allowing ϵ_{ijm} somewhat correlated among products would be an adequate approach.

Assuming Generalized Extreme Value (GEV hereafter) distributions is a suitable alternative in this set-up, the GEV models allow various substitution patterns to introduce correlations over other options. One of the most well-known and intuitive GEV models is the nested logit model where the alternatives are divided into subsets (nets), relaxing the IIA property since it adds individual-specific slopes to the utilities, resulting in better substitution patterns at the aggregate level. There are many papers on the nested logit and GEV models that deal with these two issues. However, in

our case and similar cases, we do not have a priori information about the market to perform the division of products into groups. Moreover, we have to assume the distribution of the shocks within a group when performing the demand estimation.

This is one of the main reasons why we propose to use RCL models. Now, we have to turn to the aggregate version of the RCL model. In this model, now we also try to estimate an unrestricted variance-covariance matrix of the shock, (ϵ_{ijm}), and this relaxes all above mentioned issues. However, this does not come without cost. First, we need to deal with "the curse of dimensionality." This occurs because of the number of parameters that we need to estimate will explode—a problem that aggravates with the number of products and markets. If there are N products in M markets, there are $N \cdot M$ demand curves, and since the demand curve of each product depends on the prices of the others, there are at least $N^2 \cdot M$ parameters in the model.[2] However, if we maintain Type I extreme value distribution of ϵ_{ijm}—and introduce demographics—, there will be some correlation between products with similar characteristics. This is how [2] deals with errors and also helps with the curse of dimensionality. In this setup, the consumers with similar demographics will have similar rankings of products and therefore similar substitution patterns; and the correlation between choices is captured through the term μ_{ijt}. The correlation is a function of both product and consumer characteristics. Therefore in this set-up now we have to estimate a smaller amount corresponding to an unrestricted variance-covariance matrix, rather than estimating a large number of parameters. With reduced settings, [2]'s setup facilitates in reducing errors and also helps with the curse of dimensionality significantly.

Now following [1, 2], the price elasticity of brand j in market m is:

$$\eta_{ikm} = \frac{\partial s_{jm}}{\partial p_{km}} \frac{p_{km}}{s_{jm}} = \begin{cases} -\dfrac{p_{jm}}{s_{jm}} \int \alpha_i s_{ijm}(1 - s_{jt})\, d\hat{P}_D^*(D)\, d\hat{P}_v^*(v), & \text{if } j = k, \\ \dfrac{p_{kt}}{s_{jm}} \int \alpha_i s_{ijm} s_{jkm}\, d\hat{P}_D^*(D)\, d\hat{P}_v^*(v), & \text{otherwise,} \end{cases} \tag{8}$$

where s_{ijm} is the probability of individual i purchasing product j:

$$s_{ijm} = \frac{\exp(\delta_{jm} + \mu_{ijm})}{1 + \sum_{k=1}^{K} \exp(\delta_{km} + \mu_{ikm})}. \tag{9}$$

[2] It might be better to feel the difficulty in the reality of our dataset. Even without considering characteristics and demographics, we will already have to estimate at least $N^2 \cdot M \cdot 36$ months (here N is the number of products and M is the number of markets). For our final clean dataset, this means 4,148,928 parameters.

4 Data and Estimation

4.1 The Data

We intend to estimate the model with monthly information on European mobile phone sales, prices, and product characteristics. After mandatory fields for models are defined, the database was constructed, clensed, and aggregated into one market intelligence data. The raw data comes from Eurostat, Statista, European G2k, ITC, and other third-party websites that are freely available on the web. Then, this data is combined into one relational database. It had 40 fields containing various information brand, model, price, and channel types. It initially gives us practical difficulties for data management. It is also computationally challenging because the numbers of product-market combinations exploded the dimensionality of the parameters to be estimated(more than two million parameters). To investigate further with combined data, we first define the limits of our study. In particular, we are interested in analyzing how manufacturers' product line-up, their products' characteristics (specs), and the promotions and subsidies of brands affect manufacturers' substitution patterns and pricing decisions. After running a panel survey and a series of reduced-form regressions, we have selected the following choice set. Our tables in Figs. 2, 3, 4, and 5 illustrate the evolution of our choice set in the aspects of countries, brands, price, and demographics, respectively. Also, Figs. 11, 12 and 13 in the Appendix shows the detailed specifications of selected variables during selected periods for some prominent selected brands.

The choice set

First, we had to define the set where consumers make choices. These computational challenges are dealt with by using new estimation algorithms and cloud computing, which is another way of dealing with big data in data science, but this is not within the scope of our paper. We are more interested in defining a proper choice set that would be representative enough to show the effectiveness of our setup. To define a choice set, we had to classify main models of interest into a list of distinct models and associated characteristics and quantities sold. The final list is determined by the market shares, expert panels, and lists from third party websites. Out of full data, for the reason of practicality and simplicity, we decide to use monthly data, therefore aggregated monthly to the product level across ten national markets from January 2014 to June 2016 per channel and brand (98 products), as described below. Our final sample has 36 months, ten countries, and 98 products, i.e., 29400 observations. We show some details and statistics per category below both in the main text and the Appendix. The primary data contains market shares and prices in each market, dummies, and model characteristics. Also, information on the distribution of demographics, some instruments that we consider correlated with the market share but uncorrelated with the price such as marketing mix variables, supply-side information (production cost), and some product characteristics (i.e., operating system). As for the product characteristics, which is one of the essential parts of any data set required to implement this

Country	2011	2012	2013	2014	2015	2016
Austria	791,508	794,001	892,178	860,162	994,822	1,078,943
Belgium	567,991	717,976	758,599	883,180	1,013,176	1,091,769
Czech Republic	415,665	415,059	451,472	500,241	622,231	657,507
France	4,708,800	4,660,583	4,667,221	5,258,093	4,937,294	5,066,069
Germany	6,033,516	7,358,477	8,340,855	9,007,173	9,978,009	9,666,165
Italy	3,453,280	4,230,630	5,124,371	4,993,160	5,888,468	6,197,628
Netherlands	1,478,885	1,628,413	1,660,531	1,826,805	2,243,070	2,435,477
Poland	1,260,335	1,380,539	1,389,297	1,404,264	1,834,921	2,130,351
Portugal	364,896	444,490	558,634	748,981	818,624	836,546
Spain	3,145,310	3,090,737	3,266,396	3,678,116	3,931,650	3,658,015
Grand Total	**22,220,188**	**24,720,905**	**27,109,553**	**29,160,175**	**32,262,265**	**32,818,470**

Fig. 2 Sales in selected countries (in thousand euros)

kind of model, we include physical product characteristics and market segmentation information. As [2] indicates, we use product characteristics to explain the average utility level δ_{im}, and cover the substitution patterns through μ_{ijm} in Eq. (3).

We mainly collect this information from manufacturers' official product descriptions or our prior experiences assisted by expert panels and external sources. Namely, during the data cleansing phase, we gathered and corrected these data directly from online catalogs and product descriptions from Amazon or other websites,[3] which in turn gather this information mainly from the websites of manufacturers. In the rare case of missing or conflicting information, we would directly refer to the manufacturers' homepages. We sometimes referred to the most similar models' specifications or excluded them entirely from the analysis.

As for demographics, we used the same settings as [2], who used for its examples, namely, income, income squared, age, and presence of children. Here, since our estimation will rely on assumed distributional assumptions, data is generated with distributional assumptions from macroeconomic data. As for income, we use the lognormal assumption. Statistics differ annually. We denote the estimated mean and the standard deviation by the Eurostat "Distribution of income by quantiles." This allows us to use the available information on income distribution to increase the efficiency of our estimation procedure.

Now, we summarize the variables included either as regressors or instruments in our analysis. Time variables are year and month. The model/channel is an operator, the month when the model is introduced, the family model, global region, brand, and country. The price is considered as segments (in euros) and tier by euro. Grouping is considered a channel, model, product ID, channel group code, channel group, and product group. We use monetary and physical sales (euros and units). Since the data is extensive, monthly data for 36 months are selected (January 2014–June 2016), and they are aggregated to the country-level (separated per channel, model, and brand) for ten selected countries. Figure 2 summarizes sales (in thousand euros) in these countries. There were about 22.2 billion euros worth mobile phones sold in 2011, while in 2016, the amount increases 48% to 32.8 billion. The estimation was limited for this period due to the complexity of showing elasticities since this is

[3] Some sites are www.idealo.de, www.gsmarena.com, www.kimovil.com, or www.bardtech.

	Austria	Belgium	Czech	France	Germany	Italy	Netherlands	Poland	Portugal	Spain
Alcatel	0.70	0.45	0.79	0.76	0.35	2.95	0.69	0.39	2.60	1.59
Apple	37.56	40.85	23.12	35.96	42.83	23.66	32.88	48.11	21.08	28.55
Blackberry	0.46	0.23	0.32	0.16	0.53	0.05	0.06	0.12	0.03	0.03
Htc	0.75	0.40	0.59	0.64	1.09	0.54	0.18	0.72	0.01	0.29
Huawei	10.13	10.12	12.52	3.81	5.04	8.69	14.95	4.60	13.68	15.87
Lg	2.59	1.10	2.09	1.13	1.35	4.39	2.30	2.17	2.59	4.24
Microsoft	1.24	1.35	2.18	1.16	1.21	1.46	1.22	1.35	1.30	0.19
Motorola	0.27	1.02	0.11	0.66	0.43	0.18	0.29	0.97	0.14	0.96
Nokia	0.54	0.87	1.45	0.21	0.17	1.76	0.40	0.21	1.05	0.06
Samsung	38.23	37.11	33.65	33.81	39.28	37.33	38.60	37.23	41.46	33.53
Sony	4.22	1.70	3.15	2.65	3.72	3.44	0.57	2.31	0.92	2.59
Wiko	0.13	1.56	0.00	6.77	0.81	0.00	2.08	0.33	4.91	1.31
Outside good	3.17	3.24	20.05	12.27	3.19	15.55	5.78	1.48	10.23	10.79
Total	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

Fig. 3 Market share average per brand during Jan 2014–Jun 2016 (%)

an illustrative application. However, once the estimation is made, the model can be dynamically updated by adding one more dataset per month per country, which might be meaningful especially mobile marketers who are mainly interested in changes in the demand curve (Fig. 3).

Since the data is aggregated based on the same price sold in transactions, we have to compute the average and standard deviation for grouped data to run an analysis of divergence in prices. First, we find the sales data of each channel, model per each time and period. Then, we multiply the sales figure by the frequencies of the corresponding record, and we divide it by the quantity sold. To find the value of variances in price, and since we assume that those are only sample data, we use the following basic formulas.

$$s^2 = \frac{\sum m^2 f - \frac{\left(\sum mf\right)^2}{n}}{n-1} \tag{10}$$

Since the data is in aggregated form, under some assumptions (below), standard deviations calculated using above mentioned formula will represent brand, model, and country-specific pricing strategies and price deviations at the end-consumers' level. Figures 18, 19, 20 and 21 in the Appendix provide some interesting summary statistics to illustrate raw data by brand, country, and price groups.

Assumption 1 A channel does not frequently change their prices within a month. Since we have monthly level aggregation data, "Frequently" changing prices without tendency (in consumables usually downwards) at the POS can make the standard deviation unreliable since data will only represent average prices. This can happen especially in Brick-and-Mortar stores where promotions exist, but this kind of actives are infrequent (twice a year) in the Mobile phone industry.

Assumption 2 Even if Assumption 1 does not hold, there are downwards tendencies of prices. Since downward price tendency exists, the up-and-down strategy of Brick-

and-Mortar stores can be captured by aggregation over months or a year and becomes negligible.

We define the following price variables—the price deviation index (Diff). The standard deviation of individual-level price data (by month, brand, model, channel, and country) are analyzed and aggregated again to the higher levels. There are two independent variables to represent price deviations that customers (or dealers) can capture in the long-run: (i) *Intra-country pricing variances*. They are calculated based on a mobile phone model price variation over ten countries. Standard deviations are calculated on the model level and aggregated back to brand level or brand-price tier level (we denote this variable Diff_C). (ii) *Time-series variation*. Time-series pricing variances are calculated based on a mobile phone model price variation over 12 months. Standard deviations are calculated on the model level and aggregated back to brand level or brand-price tier level (we denote this variable Diff_T). Monthly aggregated final price (p_{jm}) is the monthly average real price that consumers pay to the retailers. It is defined as sales revenue divided by the number of units sold and deflated by the consumer price index.

Physical product characteristics (specifications)

Here we discuss the definitions of physical product characteristics and their data structures that are used throughout the paper. The product characteristics in the vector (x_{jm}) we have selected from the data include Display Size, 1st Camera Resolution, 2nd Camera Resolution, Subsidy, Storage in GB, RAM in GB, HIG Resolution, CPU, CPU-CORES, Operating System, Display Techno (LCD, OLED etc.), max Generation, Keyboard, WiFi, GPS, Memory Card Slot, Bluetooth, Multi-Touch, and NFC. Display Size is the diagonal measure of the phone's display area in inches; 1st/2nd Camera Resolution is the number of megapixels on the mobile phone's camera (2nd: front/1st: back); subsidy is the money that manufactures and telecom companies for customers with contracts and marketing expenses (in euros); storage IN GB is the storage capacity of the smartphone in gigabytes (GB); RAM IN GB is the total RAM capacity of the smartphone in GB; resolution HIG is the number of pixels in the screen, CPU is the speed in megahertz (MHz) of the central processing unit; CPU-CORES is the number of processors in the CPU; Operating System and Display Techno (LCD, OLED, etc.), are categorical dummy variables (Android, I-OS, SImbian, etc.); max Generation is the max number of generation that the phone is compatible with (if is fourth-generation: $= 4$, if is third-generation: $= 3$, etc.), Keyboard, WiFi, GPS, Memory Card Slot, Bluetooth, Multi Touch, NFC are binary characteristics equal one when the phone has the technology and zero otherwise. Figure 12 in the Appendix provides some examples by brand and model.

Segmentation information

Regardless of the actual selling price to the end-customer, manufacturers' intention on market positioning (often observed by Recommended Retail Price) level or consumers' perception about the product is one of the product's non-physical characteristics. We name it "Pricing Tier" and it is used for detailed analysis and defined

Price Tier	Perceived value, €	2011	2012	2013	2014	2015	2016
Basic	~49	831,572	668,598	515,609	475,300	413,280	353,003
Basic	49.xx~99	2,466,517	2,022,796	1,709,771	1,710,751	1,711,648	1,313,021
Basic	99.xx~149	2,297,195	2,043,401	1,944,939	2,830,435	2,777,909	1,681,700
Basic	149.xx~199	2,023,004	2,127,943	2,503,964	2,613,989	3,052,613	3,694,499
Step-up	199.xx~249	1,708,019	1,843,908	1,930,798	1,746,353	1,794,736	1,940,465
Step-up	249.xx~299	1,383,555	1,302,008	1,551,755	1,497,427	1,705,285	1,762,605
Step-up	299.xx~349	965,906	908,810	1,375,948	1,831,513	1,153,360	1,739,553
High	349.xx~399	1,023,006	1,208,667	1,880,561	1,612,191	1,404,099	1,043,366
High	399.xx~449	1,296,489	1,138,897	1,378,845	1,123,721	1,516,964	881,411
High	449.xx~499	1,154,249	1,450,739	1,558,771	1,619,996	1,269,283	1,449,871
Premium	499.xx~599	2,989,614	4,104,295	3,559,221	3,166,976	2,044,846	2,696,831
Premium	599.xx~699	3,133,119	4,307,850	4,845,260	4,933,589	4,082,738	3,653,292
Premium	699.xx~	947,943	1,592,993	2,354,111	3,997,934	9,335,503	10,608,853
Total		22,220,188	24,720,905	27,109,553	29,160,175	32,262,265	32,818,470

Fig. 4 Sales in EU selected countries per pricing tier (thousand euros)

for the review. It consists of four tiers Basic: 0–199 euros, Step-up: 199–349 euros, High: 349–499 euros and Premium: 499 and more euros. This segmentation information is also used as an instrument in our estimation. Figure 4 shows sales in EU selected countries per pricing tier (thousand euros). In general, we can observe that Premium Tier's sales share increase during the period. This can reflect many aspects of changes in the market, but also different patterns in customer's tastes since the market is saturated and this category captures many early adopters without caring about owning more than one mobile phone. Moreover, customer surveys suggest that consumers tend to upgrade their phones when the previous one gets old or breaks. Therefore, many high tier consumers might have migrated to Premium Tier.[4] Figure 4 also shows product by brackets and pricing group. Again, these figures represent only the expected perceived value of the purchase.

Demographics

We used the same settings as [1, 2]. All these data have been download from Eurostat (http://ec.europa.eu/eurostat/data/database). We assumed the income distribution to be lognormal, and we estimate its parameters and the demographics drawn from Eurostat, as briefly mentioned in Sect. 2 and detailed below. Age (and child) are based on actual sample household data based on a questionnaire. To avoid disproportionate

[4] Further discussion would be out of the scope of our paper. However, actually many other interesting facts are found based on the analysis of the raw data. Summary statistics are shown in the Appendix, Figs. 18, 19, 20 and 21. We can observe, for example, that the number of models found in basic tier increase with time, while the actual sales portion of the same tier decrease. Then, there are high competition even though the demand is decreasing. This is due to low entry barriers for manufactures and saturation of European mobile market, making consumers prefer phones with better features. The number of models in Premium tier also increase with time, and consequently current sales portion of the same tier skyrocket. This explain why high-tier and premium markets are the focus of well-known firms (the cash-cow).

	income	income squared	age	child
Price	1	1	1	1
Display Size	1		1	1
1st Camera Resolution	1	1	1	
Subsidy	1			
Storage IN GB	1		1	
CPU	1	1		

Fig. 5 BLP demand estimation-combinations and demographic interactions

sampling among ages, we treated them to have a zero mean value, then reshaped to be fit to Eurostat population statistics. Therefore we use the log of income (Income), the log of income squared (Income Sq), to allow for non-linear profiles affecting demand, age, and a dummy variable equal to one of the individuals is aged less than nineteen (Child). The proxies for unobserved demographics, v_i, are drawn from a standard normal distribution. For each market, 20 individuals are drawn. We assume that the v_i are random draws from a distribution obtained with zero mean value and an identity covariance matrix independent of the level of consumer's income (y_i). The interaction distribution is assumed to be lognormal, and we estimate its parameters and the demographics drawn from Eurostat. In detail, a log of income variable (Income) and the log of income squared variable (Income Sq) are based on the "Income per sex, age" per country and year. To mimic microeconomic decisions varying by income, we sample individuals populated for 300 (market size) · 20 (individuals per market). With the same principle and based on the same data source from Eurostat, the age variable is populated per sex, age individuals are sampled for 300 (market size) · 20 (individuals per market). Child variable is extracted from the demographics that are from age variables and is a dummy variable made by the criteria, if the age is under or same as 18, 0; otherwise1.

For explanatory variables, over 200 pre-simulations with many demographic combinations have been tested. In Fig. 5, we summarize some selected interactions with combinations that show exciting results. Several interactions between observed demographic attributes and some product characteristics stand out, including age, a child with display size, income squared with 1st Camera and CPU, Storage with age. Income is assumed to interact with almost all product characteristic variables. Price is assumed to interact with all demographic variables. We will discuss other combination alternatives in Sect. 5.

4.2 Estimation

We are going to estimate the parameters by GMM. We assume that the supply and demand unobservables are mean independent of both observed product characteristics and cost shifters. On top of the discrete choice model of demand, we want to estimate a structural model of manufacturer and model pricing in this channel. The

context of our work is vertical channels consisting of competing for mobile manufacturers who sell through many conventional distributors, online and offline retailers. As we will see below, consumer heterogeneity is modeled with a finite number of discrete segments. This setup lets us better control for the potential endogeneity due to unmeasured product characteristics that can be contained in the aggregated data. We use demographics for instruments as well. We include cost data in the model of the distribution channel that captures essential features of the aggressive price-setting behavior of manufacturers using public cost information and profit published by manufacturers.

4.3 The Estimation Set-Up and Assumptions

As briefly mentioned, unobservables are mean independent of both observed product characteristics ξ_j and w_j supply shifters, i.e.,

$$\mathsf{E}\big[\xi_j \mid (x_1, w_1), (x_2, w_2), \ldots (x_J, w_J)\big] = 0,$$
$$\mathsf{E}\big[w_j \mid (x_1, w_1), (x_2, w_2), \ldots (x_J, w_J)\big] = 0.$$

Of course, we can note that price or quantity are not included in the conditioning vector, since they are partially determined by ξ_j and w_j. We assume that product characteristics x_j and cost shifters w_j are exogenous.

Estimation Algorithm

Inverting the demand equations gives:

$$\delta_j(s^{\text{obs}}, \delta) = x_j\beta - \alpha p_j + \xi_j. \tag{11}$$

Suppose $H_j(x, w)$ is a matrix of instruments:

$$\mathsf{E}\left[H_j(x, w)\begin{pmatrix}\xi_j(s, \theta)\\ w_j(s, \theta)\end{pmatrix}\right] = 0. \tag{12}$$

Our goal is here to estimate the demand parameters α, β, γ, and θ; α_i is the consumer i's marginal utility from income, and β_{ijm} are demand parameters or individual-specific taste coefficients (for individual i purchasing j in market m). The moment conditions are:

$$\mathsf{E}\left[H_j(x, w)\begin{pmatrix}\delta_j(s, \theta) - x_j\beta + \alpha p\\ p_j - c_j(q_j, w_j) - \frac{1}{\alpha}\left[\frac{s_j}{\partial s_j/\partial \delta_j}\right]\end{pmatrix}\right] = 0. \tag{13}$$

The estimation is carried out as a nested procedure: First, we make an outer loop for searching the minimum of the GMM objective function in the parameter space.

The linear parameters α, β, γ are easy to obtain using matrix algebra, while for estimating the non-linear ones need numerical methods. Second, we make an inner loop for each θ, use contraction mapping to solve for δ_j.

4.4 Instruments

We use instruments to deal with the endogeneity of prices, the classic problem faced when estimating demand models. Since price and other explanatory variables are correlated, variances in demand parameters are not captured by the model. We usually assume that these parameters in the model to be a function of consumers' perceptions about the product, which vary by countries, over time, and with demographics. This heterogeneity cause differences in the demand system and prices per period and market. Therefore, we need to find instruments for both the demand and pricing equations. Any variables that are correlated with prices and explanatory variables but are not correlated with the demand market share are appropriate instruments. The next step is specifying a list of variables that are mean independent and are relevant to the variables to be instrumented. The unobservables for both supply and demand are mean independent of both observed product characteristics and cost shifters. References [2, 4] summarizes some of the solutions offered in the literature and suggests many types of instruments that can be used in this set-up. We use the same approach to select adequate instrumental variables. The instrumental variables must be associated with a corresponding product j and include functions of the characteristics and cost shifters of all other products. Finally, we test the validity of the instrumental variables. In order to obtain consistent parameter estimates, tests of overidentifying restrictions should inform about their validity (non-correlation with the error term). The value of the tests does not reject the null that the instruments are valid in any of the estimated specifications.

Supply-side: Cost Data (w, supply shifters)

The first suggestion for demand-side instrumental variables following [2, 4] is to look for variables that are uncorrelated with the demand shock and are able to shift cost. Since most of the products are not produced in European countries, the wage-setting process affects only marketing and logistic parts of the cost. For the demand for a specific brand and model, we need to find more model-specific information. It is quite challenging to obtain cost-related data because they are often confidential. We refer to some public data found on for two major brands (Apple and Samsung) extrapolated by disassembling parts and adding up costs with the reasonable hypothesis (see https://www.fairphone.com or http://gizmodo.com). Some other independent organizations perform cost analysis based on this kind of reverse-engineering. Figure 6 shows the results of this kind of analysis. While controlling brand and demographics, the average price of the other nine countries (excluding the country to be instrumented) will be uncorrelated with market-specific valuation.

Smartphone	Production Cost	Retail Price	Profit margin (%)
Apple iPhone 7 (32GB)	$224.80	$649.00	65.4
Apple iPhone 6S Plus (16GB)	$236.00	$749.00	68.5
Apple iPhone 6S Plus (64GB)	$253.00	$849.00	70.2
Apple iPhone 6S (16GB)	$211.50	$649.00	67.4
Apple iPhone 6 Plus (16GB)	$215.60	$749.00	71.2
Apple iPhone 6 Plus (128GB)	$263.00	$949.00	72.4
Apple iPhone 6 (16GB)	$200.10	$649.00	69.2
Apple iPhone 6 (128GB)	$247.00	$849.00	70.9
Apple iPhone 5C (16GB)	$183.00	$549.00	67.0
Apple iPhone 5S (16GB)	$199.00	$649.00	69.0
Apple iPhone 5S (64GB)	$218.00	$849.00	74.0
Apple iPhone 5 (16GB)	$207.00	$649.00	68.0
Apple iPhone 4 (16GB)	$188.00	$599.00	69.0
Apple iPhone 4S (16GB)	$188.00	$599.00	69.0
Google Pixel XL (32GB)	$285.75	$769.00	62.8
Samsung Galaxy Note 3 (32GB)	$232.50	$699.00	69.6
Samsung Galaxy S3 (16GB)	$213.00	$549.00	61.0
Samsung Galaxy S4 (16GB HSPA+)	$244.00	$579.00	58.0
Samsung Galaxy S5 (32GB)	$256.00	$569.00	55.0
Samsung Galaxy S6 (32GB)	$275.50	$699.99	60.6
Samsung Galaxy S6 Edge (64GB)	$290.45	$799.99	63.7
Samsung Galaxy S7 (32GB)	$255.00	$599.00	57.4
Samsung Galaxy S8 (64GB)	$307.50	$720.00	57.3

Fig. 6 Examples of production costs, retail prices, and margins

Manufacturer (Brand) dummies

The most popular identifying assumption used to deal with the above endogeneity problem is to assume that the location of products in the characteristics space is exogenous or predetermined to the revelation of the consumers' valuation of the unobserved product characteristics. This assumption can be combined with a specific model of competition and functional-form assumptions to generate an implicit set of instrumental variables as in [21–23]. If a manufacturer (brand) b sets prices of its products, it will try to maximize its profit, therefore the markup. However, those mobile phones with many comparable phones cannot have high markups due to high competition. Since b has better information about its own products' prices and markups, the own markups vis-à-vis those of competitors' products will be different. Therefore, the optimal brand dummies will distinguish between the characteristics of products produced by the same manufacturer versus the features of products manufactured by others. Let J_b denote the set of all products produced by the manufacturer (brand) b. The suggestion of BLP for the instruments of x_{jb} (the kth characteristic of product j by the manufacturer or brand b) are:

$$x_{jk}, \quad \sum_{r \neq j,\ r \in J_b} x_{rk}, \quad \sum_{r \notin J_b} x_{rk}.$$

Hausman type instruments

Similarly, as in the seminal paper by [24], which was also used in [2], the notations and the model we describe here are in the same context. We mainly make the most of

I1	I2	I3	I6	I7	I10	I11	I14
2nd Camera MP	CPU-CORES	RAM IN MB	Operating System	Keyboard	Bluetooth	Generation Total	NFC
D1~D7 Brand dummies (8 Brands - 1)							

Fig. 7 Some selected instruments

the nature of the panel data and construct instruments for the price by averaging the price of product j in other markets different from the one studied. We make here an assumption of oligopolistic competition (not unrealistic for the mobile phone market). Controlling for brand-specific interceptions and demographics helps to identify an assumption, as country-specific product valuations are independent of cities, but allowed to correlate within a country over some time. Under this assumption, prices of the same brand in other countries are valid instruments. Prices of brand j in two countries are correlated by the average marginal cost, but due to the assumption of independence, they are not correlated with the market-specific customer perception of the product.

Demographics and other instruments

Demographics are assumed exogenous given the source of information used. We also assume that the location of products in the characteristics space is exogenous, or at least predetermined to the revelation of the consumers' valuation of the unobserved product characteristics. This assumption can be combined with a specific model of competition and functional-form assumptions to generate an implicit set of instrumental variables as in [22, 23], at least predetermined in an econometric sense. Below in Fig. 7, we summarize some combinations selected, which show relatively good performance.

5 Results

We now report our findings for the model presented in Sect. 3 and applied to the data described in Sect. 4. In Figs. 14 and 15, we report the parameter estimates for various specifications. Due to big raw data and future consideration for the dynamic update, open-source statistic software R depending mainly on BLPestimatoR-package,[5] though not exclusively, and most of the optimization processes were run on Amazon Web Service (https://aws.amazon.com/). The routine uses analytic gradients and offers a large number of implemented integration methods and optimization routines. We use a Wald test for goodness of fit. It tests whether the determinants explain the actual market shares. Figure 8 provides some combinations selected for running the

[5] We use R 3.4.0 for the analysis, using status-of-the-art BLPestimatoR-package for the analysis. This package provides the estimation algorithm described in [2].

Name	Price	x1	x2	x3	x4	x5
detail	Average selling price	Display Size	1st Camera Resolution	Subsidy	Storage IN GB	CPU
Type	endogenous	exogenous, random	exogenous, random	exogenous, random	exogenous, random	exogenous, random

Fig. 8 Some selected combinations

simulations (more details are Figs. 11, 12, 13, 14 and 15 in the Appendix). Here we only present product characteristics x1 to x5 (Display Size, 1st Camera Resolution, Subsidy, Storage in GB, and CPU) in addition to the price.[6] Interactions of these variables with demographics are also included.

5.1 Estimation Without Country Dummies and Without Demographics

We start with the simplest model, where we assume that countries are homogeneous, and there is no heterogeneity among individuals in a country. The parameter estimates for the RCL demand model given by Eq. (13) are summarized in Fig. 13 of the Appendix. As expected, the price parameter is negative and statistically significant, implying that higher prices reduce utility for mobile phone consumers. CPU partly captures the price effect, but the other included characteristics are non-significant. Price captures almost all the impact of mobile phone consumers' utility. The model and country dummy variables are, in general, highly significant. Once they are included, product characteristics do not affect consumers' utility. This might be related to the significant differences in tastes by brand and country. In case the tastes per market are different, we would expect improvement in our model's performance by introducing demographic information. In the case of high brand effects, organic products compared to non-organic products as consumers are price-sensitive, so higher prices result in lower utility. These results suggest that average consumers are highly interested in specific brands, especially well-known brands. The primary driver of the market share is branded (3–10%), which explains 1.4–1.6% of its market share.

5.2 Estimation with Demographics

By relaxing homogeneity in tastes, we can observe different decisions among consumers in various countries with different levels of income and average age. The

[6] In many of the estimations with more than 5 variables, we do not get convergence of the GMM method or the results are meaningful at the optimum. We opt for choosing only 6 variables (price plus 5 product characteristics) entering linearly the utility function, out of the 21 possible explanatory variables (1 price + 20 product characteristics).

parameter estimates for the RCL demand model given by Eq. (13) are summarized in Fig. 14 of the Appendix. The distribution of marginal utilities is estimated by minimum-distance and are presented in columns under the heading "Std. Error". The level of significance is presented next and grouped with different colors. The parameter estimates for demographic variables allow calculating individual taste variations depending on them. The price Display Size, 1st Camera Resolution, Subsidy, Storage IN GB, and CPU enter the mean utility variation linearly and also interact with demographics. Since all parameters are average values of the combination of parameters and demographics, we only discuss average impacts (i.e., effects around the mean).

At the aggregate level, we can make several observations. All brand dummy variables except for "branded" are non-statistically significant. This might be related to the fact that, on average, consumers do not show preferences for specific brands, except in the case of Apple and Samsung. Moreover, country effects might be captured by country-level demographics, which should be observed by market and product level. Most of the intercepts are not significant, and this might be because of the large portion of "outside goods" since consumers buy a product in the generic group. The price parameter is negative and statistically significant, implying, again, that higher prices reduce utility for mobile phone consumers. In general, as one might expect, being under the 19-year impact on the consumer's decisions negatively and make them heavily dependent on the price of the product. However, when a young consumer decides to purchase, he first considers the camera's resolution as a relevant factor. Older consumers value other different characteristics of mobile phones. Income shows, as expected, a positive relationship with the purchase decision. The higher the income is, the higher the utility that the consumer attains. However, the marginal impact of income decreases with increased age.

Now we discuss the estimates in conjunction with demographics. A general observation about all of these results: tastes about mobile phones vary hugely with the country and demographic factors. The results represent our natural expectation about price elasticity of common goods—average consumers prefer massive subsidy from phone manufacturers and telecommunication companies at the purchasing, and these subsidies boost sales. However, here are some remarkable points—on average to high-income demographics are interacting more with these subsidies than low-income demographics—we think it is because high-income individuals buy relatively pricy phones. Therefore, the absolute amount per handset should be high. Coefficients on the interaction of price with demographics are significant, while the estimate of the standard deviation suggests that most of the heterogeneity is explained by the demographics. Underage and above-average-income consumers tend to be less price-sensitive. The rest of the estimates have some different interpretations per combinations of explanatory variables. Figure 9 reports a summary of the effects (signs and significance).

	Mean	Income	Age	Age under 18
Price	-	+	+	-*
Display Size	+	-*	+*	-*
1st Camera Resolution	+	-*	+*	-*
Subsidy	+	+*	-	-*
Storage IN GB	+	-*	+*	+*
CPU	+	+*	-*	+*
Apple/Samsung	+*	+*	+*	+*

*statistically significant; + positive correlation; - Negative correlation

Fig. 9 Signs of parameters and statistical significance

Average values for the ten countries of own and cross-price elasticities based on Eq. (8) are summarized in Fig. 16. The model gives more than 345,000 values for each country (98^2 times 36 months). These results would allow manufacturers to develop a customized and complex strategy per market and model and would allow policymakers to analyze substitution patterns per brand, period, and market. However, here, we will only mention some general findings from the full set of results. Firstly, we comment on own-price elasticities. The majority of them are negative, as expected, and they range from −19.13 to 1.27. It implies that mobile demand is elastic. This result is consistent with the BLP estimations for the differentiated demand results. Figure 16 report the minimum, average, and maximum values of own-brand-price elasticities for the ten countries. Most of them are negative Fig. 16, and the range of variation goes from −8.65 to 0.24. The major brands, Apple, Samsung, Sony, HTC, and LG, show high own-brand substitution patterns Fig. 16. In essence, it seems logical that a company with a high market share would also have its own-brand competition. However, for one of those brands in a country with higher substitution patterns, product managers must re-think their product portfolios trying to avoid their own-cannibalization. Cross-price elasticities show interesting results. First, they show lower figures (in absolute value) than own-price elasticities, suggesting that consumers tend to have some degree of product and brand loyalty for mobile phones. On average for all ten selected countries, demand for Apple's iPhone-16-Gb-LTE and iPhone-64Gb-LTE models tend to be more elastic than the rest of Apple products, with cross-price elasticities in the range 0.21–0.39. The implication would be that if Apple increases the price of iPhone-16-Gb-LTE and iPhone-64Gb-LTE by 1%, the other aforementioned products will increase demand by 0.21–0.39%, the highest figure among all selected models. Apple's Iphone5-16Gb-LTE, Samsung's GalaxyS4-16G-LTE, GalaxyA7-16Gb-LTE, Sony's One-32Gb-LTE, Onem8S-16GB-LTE, Htc LUMIA1020-32Gb-LTE, and Lg D855-G3-32Gb-LTE have very low average cross-price elasticities (around 0). This result suggests that consumers are less sensitive to price changes in those products.

Due to the heterogeneity of markets, many of the remarkable implications of our estimation results would be best observed at a model/country level of aggregation. To give some idea of these implications without overwhelming the reader with details, we display them in Fig. 10 only for the illustrative examples of 14 models for a country (The Netherlands) and a month (January 2016). Also, in Figs. 22, 23, 24, 25, 26, 27, 28, 29, 30 and 31, we provide country-level aggregation results for other countries. To show how these results are used, some comments about them follow. In almost all countries except the Czech Republic and Poland, Apple and Samsung show high brand-own cross-price elasticities, implying many lines of products are cannibalizing themselves. For those brands, product characterization, different positioning of models, and price differentiation within the brands would be helpful to increase sales and profits in many European countries. And in France, OneM8-S16Gb-Lte had on average a positive own-price sensitivity (due to its low price), so even price increases lead to increasing demand (kind of Giffen good). In Germany, Apple and Samsung' cross-product elasticity (0.09 on average) is higher than in other countries (0.05 on average), a kind of income effect. In the Netherlands, Apple Iphone5S-16Gb-LTE faces high cross-product elasticities against 8 Samsung and 2 Htc products. In Spain, Sony and Htc models show very low cross-price elasticities (around 0), while Samsung's' GalaxyS4-16Gb-LTE and Sony's XperiaZ2-LTE have positive own-price elasticity ranging from 1.11 to 2.25. In Austria, Apple iPhone6-16-Gb-LTE and iPhoneS5-16-Gb-LTE face extremely high cross-product elasticities against 8 Samsung 2 Sony and 4 Htc products. However, almost all Samsung's cross-product elasticity versus all other Android phones (0.02 on average) shows this brand is almost the unique alternative to the iPhone for the majority of consumers in this country. Galaxy-alpha-32Gb-850F-NFC-LTE, Galaxy-S4Active-16GB-I9295 -LTE, and ONE-32GB-NFC-LTE had, on average a positive own-price sensitivity, so even price increases lead to increasing demand.

We can also provide some results by any variable as a country, time, brand, etc. So, we are going to finish this section with an example (Fig. 10) of the distribution of price sensitivity attending the market share of the product (choosing those with the highest market shares). Most of the coefficients are statistically significant. There are some implications for the managers of the firm, at least at the moment they decide to launch new products. They should examine the kind of competition they face, the possibilities for substituting products among consumers, or the sort of portfolio that would prevent cannibalization within a brand. Moreover, the different models of the same brand can be characterized as price-elastic or price-inelastic, and a possible strategy consists in focusing on increasing sales (and the market share) according to these elasticities. Managers can identify other essential characteristics of the product for the consumer choice regardless of its price. Of course, there are also implications

Brand	Display Size	1st Camera Resolution	Storage IN GB	Screen Resolution HIG	Model	Sahre	IPHONE5S16	IPHONE616G	IPHONE664G	IPHONE4S8G	IPHONE5C8G	GALAXYSIII	GALAXYS416	GALAXYS516	GALAXYS632	GALAXYS4MI	GALAXYS5NE	GALAXYA316	NEXUS516BG
APPLE	4	7990	16	1136	IPHONE5S16	6.19%	-1.151	0.041	0.028	-0.022	-0.008	-0.010	0.038	0.010	0.011	-0.036	-0.019	-0.024	-0.007
APPLE	5	7990	16	1334	IPHONE616G	3.54%	0.079	-1.084	0.028	-0.022	-0.008	-0.010	0.038	0.011	0.011	-0.036	-0.019	-0.024	-0.007
APPLE	5	7990	64	1334	IPHONE664G	1.51%	0.079	0.041	-1.770	-0.022	-0.008	-0.010	0.038	0.011	0.011	-0.036	-0.019	-0.024	-0.007
APPLE	4	7990	8	960	IPHONE4S8G	6.19%	0.084	0.042	0.029	-0.024	-0.009	-0.010	0.036	0.010	0.010	-0.037	-0.017	-0.025	-0.007
APPLE	4	7990	8	1136	IPHONE5C8G	2.36%	0.081	0.041	0.028	0.317	-0.008	-0.010	0.037	0.010	0.011	-0.036	-0.019	-0.024	-0.007
SAMSU	5	7990	16	1280	GALAXYSIII	2.70%	0.080	0.041	0.028	-0.022	0.325	-0.010	0.038	0.010	0.011	-0.036	-0.019	-0.024	-0.007
SAMSU	5	12780	16	1920	GALAXYS416	6.33%	0.073	0.039	0.026	-0.020	-0.008	-0.009	-0.572	0.012	0.012	-0.034	-0.022	-0.022	-0.007
SAMSU	5	15872	16	1920	GALAXYS516	2.47%	0.070	0.037	0.025	-0.019	-0.007	-0.009	0.041	-0.431	0.013	-0.033	-0.023	-0.022	-0.007
SAMSU	5	15872	32	2560	GALAXYS632	2.62%	0.069	0.037	0.025	-0.019	-0.007	-0.009	0.041	0.012	-0.434	-0.032	-0.023	-0.022	-0.007
SAMSU	4	7990	8	960	GALAXYS4MI	6.67%	0.079	0.041	0.028	0.022	0.008	0.010	-0.038	-0.011	-0.011	-0.486	0.019	0.024	0.007
SAMSU	5	15872	16	1920	GALAXYS5NE	4.16%	0.068	0.036	0.025	0.019	0.007	0.009	-0.041	-0.012	-0.013	0.032	-0.471	0.021	0.007
SAMSU	5	7990	16	960	GALAXYA316	2.27%	0.078	0.041	0.028	0.022	0.008	0.010	-0.038	-0.011	-0.011	0.036	0.020	-0.993	0.007
LG	5	7990	16	1920	NEXUS516BG	1.72%	0.078	0.041	0.028	0.022	0.008	0.010	-0.038	-0.011	-0.011	0.035	0.020	0.023	-0.408

Fig. 10 Price elasticities among models with high market share (The Netherlands, January 2016)

for policy-makers concerning the monopolistic/oligopolistic status of a brand in a country. This fact, together with additional information as production costs and marginal profits, could generate cases of dumping. It is policy-makers' responsibility to evaluate the pricing strategies of companies and to introduce regulations to correct market failures, when necessary.

6 Concluding Remarks

In this paper, we studied the demand for several EU countries' mobile industry using proposals from the literature on marketing and IO. The methods and results from our paper are meaningful for many people who are related to the mobile industry or demand estimation, such as mobile phone markers or European policymakers. We present logit and RCL following the BLP/Nevo methodology of discrete purchases from aggregate data. We base our specification on the observations of the market, microeconomic demographics, and aggregate market intelligence data. We made the most of these data from many sources to understand static models of demand in the mobile phone industry. We find that the methodology is computationally quite challenging. We used macroeconomic data matching the distributional assumption for fitting demographics in each country. Raw data was carefully analyzed, and we select data of 98 products in 12 European countries for 36 months. We made some assumptions for Hausman type and supply-side information for building suitable instruments.

The RCL model reveals impressive results. Estimations without demographics show mobile phones in European countries, in general, are well sold if their prices are low. The results, including demographics, show several different conclusions per country and demographics. However, they also imply that heterogeneity in tastes is significant for consumer behavior in the country-time-specific mobile phone industry. Results indicate that consumers are price and income-sensitive. Moreover, the results suggest that age (both in number and under 18/not) is a crucial determinant of consumer's tastes. We presented some findings of our RCL demand model given

by Eq. (13) at the aggregate level and the market level. In the setting of [1, 2], we also estimated price elasticities as a by-product of estimating demand parameters. We showed that our approach and analysis is not only interesting for product managers who should make product prices and portfolios, but also for policymakers who should objectively evaluate the market's regulation.

The significant contribution of our approach is to provide a practical method for estimating price elasticities for demand systems involving many similar datasets using market intelligence data. The applications to other related electronics industry would be quite simple (see [25]), considering the similarities in characteristics, perishable nature of selling prices, heterogeneous demand, and demographic differences in tastes. The literature shows that researchers and companies have used statistical techniques and random coefficient logit (RCL) models the estimating demand systems, using historical sales series as their primary source of data. The paper shows an excellent example of applying the classical BLP model with smart use of estimation and instrumental variable constraints, but at the same time making the most of new technologies such as Amazon Web Service cloud, and the state-of-the-art BLPestimatoR-package. Of course, other approaches are also making the most of new technologies, such as demand forecasts using Machine Learning(ML), which is a buzz word these days. ML demand forecasts sometimes can show better quality results than traditional techniques [26]. However, the ML models can not replace the RCL model. Nevertheless, it can be accompanied by BLP because BLP shows more energetic performances in many contexts [27]. Moreover, interpretability is one of the essential advantages of our RCL model and structural models in comparison to machine learning. As we showed in Chaps. "Collaborative Innovation of Spanish SMEs in the European context: A compared study" and "Haar systems, KMS states on von Neumann algebras and C* -algebras on dynamically defined groupoids and Noncommutative Integration" in our paper, RCL can extract useful information from aggregate data such as coefficients, the relationship among variables, as well as own/cross-product price elasticities.

Acknowledgements We thank Alberto A.Álvarez-López for his insightful comments and encouragement, but also for some useful comments and suggestions, which incentivized us to widen our research from various perspectives.

Appendix

See Figs. 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 and 31.

Brand	Model	1st Camera Resolution (K Pixel)	2nd Camera MP	CPU-CORES	RAM IN MB	Storage IN GB	Resolution HIG
APPLE	IPHONE5S16GBLTE	7990	1.2	2	1024	16	1136
	IPHONE616GBNFCLTE	7990	1.2	2	1024	16	1334
	IPHONE4S16GB	7990	0.3	2	512	16	960
	IPHONE516GBLTE	7990	1.2	2	1024	16	1136
	IPHONE664GBNFCLTE	7990	1.2	2	1024	64	1334
	IPHONE416GB	5018	0.3	1	512	16	960
	IPHONE48GB	5039	0.3	1	512	8	960
	IPHONE5S32GBLTE	7990	1.2	2	1024	32	1136
	IPHONE4S8GB	7990	0.3	2	512	8	960
	IPHONE5C16GBLTE	7990	1.2	2	1024	16	1136
	IPHONE532GBLTE	7990	1.2	2	1024	32	1136
	IPHONE6PLUS16GBNFCLTE	7990	1.2	2	1024	16	1920
	IPHONE5C8GBLTE	7990	1.2	2	1024	8	1136
	IPHONE6PLUS64GBNFCLTE	7990	1.2	2	1024	64	1920
	IPHONE4S32GB	7990	0.3	2	512	32	960
	IPHONE6128GBNFCLTE	7990	1.2	2	1024	128	1334
	IPHONE5S64GBLTE	7990	1.2	2	1024	64	1136
	IPHONE564GBLTE	7990	1.2	2	1024	64	1136
	IPHONE6PLUS128GBNFCLTE	7990	1.2	2	1024	128	1920
	IPHONE4S64GB	7990	0.3	2	512	64	960
	IPHONE5C32GBLTE	7990	1.2	2	1024	32	1136
SAMSUNG	GALAXYSIII16GBI9300NFC	7990	1.9	4	1024	16	1280
	GALAXYS416GBI9505NFCLTE	12780	2	4	2048	16	1920
	GALAXYS516GBG900NFCLTE	15872	2	4	2048	16	1920
	GALAXYS632GBG920NFCLTE	15872	5	8	3072	32	2560
	GALAXYS4MINI8GBI9195NFCLTE	7990	2	2	1536	8	960
	GALAXYS6EDGE32GBG925FNFCLTE	15872	5	8	3072	32	2560
	GALAXYNOTE332GBN9005NFCLTE	12780	2	4	3072	32	1920
	GALAXYS5MINI16GBG800FNFCLTE	7990	2.1	4	1536	16	1280
	GALAXYNOTE4N91032GBNFCLTE	15872	3.7	4	3072	32	2560
	GALAXYNOTEII16GBN7100NFC	7990	1.9	4	2048	16	1280
	GALAXYA5A500FNFCLTE	12780	5	4	2048	16	1280

Fig. 11 Some selected models and characteristics (1)

Brand	Model	1st Camera Resolution (K Pixel)	2nd Camera MP	CPU-CORES	RAM IN MB	Storage IN GB	Resolution HIG
	GALAXYS5NEO16GBG903FNFCLTE	15872	5	8	2048	16	1920
	GALAXYNOTE16GBN7000	7990	2	2	1024	16	1280
	GALAXYA316GBA300FNFCLTE	7990	5	4	1536	16	960
	GALAXYS6EDGEPLUS32GBG928NFCL	15872	5	8	4096	32	2560
	GALAXYALPHA32GBG850FNFCLTE	11944	2.1	8	2048	32	1280
	GALAXYS6EDGE64GBG925NFCLTE	15872	5	8	3072	64	2560
	GALAXYS416GBI9515NFCLTE	12780	2	4	2048	16	1920
	GALAXYS664GBG920NFCLTE	15872	5	8	3072	64	2560
	GALAXYNOTE3NEON7505NFCLTE	7990	2	6	2048	16	1280
	GALAXYS4ACTIVE16GBI9295NFCLTE	7990	2	4	2048	16	1920
	GALAXYNOTEII16GBN7105NFCLTE	7990	1.9	4	2048	16	1280
	GALAXYNOTEEDGE32GBN915NFCLTE	15925	3.7	4	3072	32	2560
	GALAXYXCOVER3G388FNFCLTE	5039	2	4	1536	8	800
	GALAXYGRAND2G7105NFCLTE	7990	1.9	4	1536	8	720
	GALAXYMEGA6.38GBI9205LTE	7990	1.9	2	1536	8	720
	GALAXYGRANDDUOSI9082	7990	2	2	1024	8	800
	GALAXYA716GBA700NFCLTE	12780	5	8	2048	16	1920
	GALAXYS4ZOOM8GBC1010NFC	17818	1.9	2	1536	8	960
SONY	XPERIAZNFCLTE	12780	2.2	4	2048	16	1920
	XPERIAZ3NFCLTE	20656	2.2	4	3072	16	1920
	XPERIAZ3COMPACTD5803NFCLTE	20656	2.2	4	2048	16	1280
	XPERIAZ2NFCLTE	20656	2.2	4	3072	16	1920
	XPERIAZ1COMPACTD5503NFCLTE	20656	2	4	2048	16	1280
	XPERIASPNFCLTE	7990	0.3	2	1024	8	1280
	XPERIAM4AQUA8GB	12780	5	8	2048	8	1280
	XPERIATLT30PNFC	12780	1.3	2	1024	16	1280
	XPERIAT3D5103NFCLTE	7990	1.1	4	1024	8	1280
	XPERIAVNFCLTE	13129	0.3	2	1024	8	1280
	XPERIAZULTRA16GBNFCLTE	7990	2	4	2048	16	1920
HTC	ONE32GBNFCLTE	4086	2.1	4	2048	32	1920
	ONEM816GBNFCLTE	4086	5	4	2048	16	1920
	ONES	7990	0.3	2	1024	16	960
	ONEX32GBNFC	7990	1.3	4	1024	32	1280

Fig. 12 Some selected models and characteristics (2)

Brand	Model	1st Camera Resolution (K Pixel)	2nd Camera MP	CPU-CORES	RAM IN MB	Storage IN GB	Resolution HIG
	ONEMINILTE	4086	1.6	2	1024	16	1280
	ONEM932GBNFCLTE	20171	4	8	3072	32	1920
	ONEMINI2NFCLTE	12780	5	4	1024	16	1280
	DESIRE820LTE	12979	8	4	2048	16	1280
	ONEM8S16GBNFCLTE	12780	5	8	2048	16	1920
NOKIA	LUMIA920NFCLTE	7990	1.3	2	1024	32	1280
	LUMIA93032GBNFCLTE	18690	1.2	4	2048	32	1920
	LUMIA92516GBNFCLTE	7990	1.3	2	1024	16	1280
	LUMIA102032GBNFCLTE	41484	1.2	2	2048	32	1280
	LUMIA830NFCLTE	8580	0.9	4	1024	16	1280
	LUMIA1320LTE	5039	0.3	2	1024	8	1280
	LUMIA735NFCLTE	6621	5	4	1024	8	1280
LG	D855G316GBNFCLTE	12979	2.1	4	2048	16	2560
	D802G216GBNFCLTE	12780	2.1	4	2048	16	1920
	H815G432GBNFCLTE	15872	8	6	3072	32	2560
	NEXUS516BGNFCLTE	7990	1.3	4	2048	16	1920
	D722G3SNFCLTE	7990	1.3	4	1024	8	1280
	D855G332GBNFCLTE	12979	2.1	4	3072	32	2560
	E960NEXUS416GBNFC	7990	1.3	4	2048	16	1280
	E975OPTIMUSGNFC	12979	1.3	4	2048	32	1280
BLACKBERRY	BOLD9900NFC	5039	0	1	768	8	640
	Z10NFCLTE	7990	2	2	2048	16	1280
	Q10NFCLTE	7990	2	2	2048	16	720
	CLASSIC16GBNFCLTE	7990	2	2	2048	16	720
	Z30NFCLTE	7990	2	2	2048	16	1280
	LEAP16GBLTE	7990	2	2	2048	16	1280
	Q5NFCLTE	5039	2	2	2048	8	720
HUAWEI	P8LITE16GBDUALNFCLTE	12979	5	8	2048	16	1280
	P816GBNFCLTE	12979	8	8	3072	16	1920
	ASCENDG7NFCLTE	12979	5	4	2048	16	1280
	ASCENDP7LTE	12979	8	4	2048	16	1920
	ASCENDP68GB	7990	5	4	2048	8	1280
	ASCENDMATE716GBNFCLTE	12780	5	8	2048	16	1920

Fig. 13 Some selected models and characteristics (3)

Variable	Estimate	Std. Error	Pr(>|t|)
price	-0.895	0.266	0.000
dummy1	-3.429	0.150	0.000
dummy2	-3.077	0.144	0.000
dummy3	-0.931	0.152	0.000
dummy4	-2.543	0.150	0.000
dummy5	-1.796	0.134	0.000
dummy6	-1.143	0.198	0.000
dummy7	-3.74	0.155	0.000
country1	-0.535	0.274	0.051
country2	-1.085	0.280	0.000
country3	-1.414	0.270	0.000
country4	0.895	0.266	0.000
country5	-1.236	0.271	0.000
country6	-0.190	0.281	0498
country7	-1.134	0.269	0.000
country8	-1.618	0.274	0.000
country9	-1.353	0.278	0.000
price	-141.842	6.338	0.000
Display Size	0.008	0.648	0991
1st Camera Resolution	168.144	1860.847	0928
Subsidy	5.207	25.945	0.841
Storage IN GB	1.268	0.256	0.000
CPU	436.439	26.611	0.000

Pr(>|t|)	Pr(>|t|)	Pr(>|t|)	Pr(>|t|)
0.2	0.1	0.05	0.01

Fig. 14 Results in models without demographics

	income			age			underage		
	Estimate	Std. Error	Pr(>|t|)	Estimate	Std. Error	Pr(>|t|)	Estimate	Std. Error	Pr(>|t|)
price	356.92	81.22	0.000	55.87	193.16	0.772	-412.86	118.22	0.000
Display Size	-0.42	0.20	0.037	1.30	0.50	0.009	-0.88	0.35	0.011
1st Camera Resolution	-7073.48	2712.37	0.009	12116.98	3778.99	0.001	-5043.53	1779.66	0.005
Subsidy	87.85	25.35	0.000	-35.22	64.16	0.583	-52.61	45.14	0.244
Storage IN GB	-192.27	94.03	0.004	141.87	68.82	0.039	50.39	27.00	0.0620
CPU	940.96	176.15	0.000	-1956.45	408.95	0.000	1015.47	413.76	0.014
Wald test: 626.8766 on 16 DF, p-value: 0.000									

Pr(>|t|)	Pr(>|t|)	Pr(>|t|)	Pr(>|t|)
0.2	0.1	0.05	0.01

Fig. 15 Results in models with demographics

Country	Stat	Apple	Samsung	Sony	HTC	Nokia	LG	Huawei
Average	Min	-19.13	-14.33	-4.86	-8.39	-0.40	-3.83	-0.90
	Max	-2.15	1.27	-1.70	0.82	0.89	0.10	-0.90
	Avg	-2.88	-2.32	-1.09	-1.19	0.24	-1.63	-0.90
France	Min	-19.67	-14.93	-6.70	-8.92	-1.30	-4.59	-1.83
	Max	1.59	-0.07	-3.72	0.65	-1.07	-1.79	-1.83
	Avg	-2.97	-2.64	-1.74	-1.31	-1.18	-2.80	-1.83
Germany	Min	-17.50	-14.79	-3.61	-7.68	0.39	-4.22	0.74
	Max	-1.24	2.94	-0.24	1.89	2.71	1.13	0.74
	Avg	-2.70	-1.98	-0.64	-0.65	1.55	-1.15	0.74
Spain	Min	-19.52	-14.08	-5.94	-10.76	-0.20	-6.08	-1.27
	Max	-2.70	1.11	-3.24	2.25	-0.04	-0.39	-1.27
	Avg	-3.05	-2.32	-1.53	-1.24	-0.12	-2.67	-1.27
Italy	Min	-20.32	-17.45	-3.64	-8.75	-1.24	-5.58	-2.41
	Max	-3.51	0.17	-1.42	-1.79	-0.90	-0.85	-2.41
	Avg	-3.37	-3.01	-0.84	-1.94	-1.07	-2.72	-2.41
Austria	Min	-17.18	-12.87	-3.64	-6.39	0.20	-1.99	0.74
	Max	-1.45	0.27	-0.29	0.40	1.12	0.83	0.74
	Avg	-2.59	-1.96	-0.65	-0.74	0.66	-0.45	0.74
Belgium	Min	-18.56	-13.34	-4.71	-9.39	0.25	-2.93	-0.75
	Max	-0.98	0.32	-1.38	1.25	2.35	-0.90	-0.75
	Avg	-2.85	-2.30	-1.01	-1.14	1.30	-2.04	-0.75
Czech Republic	Min	-19.28	-15.65	-6.67	-8.18	-0.30	-2.02	-1.50
	Max	-1.06	1.95	-0.65	1.52	1.34	2.33	-1.50
	Avg	-2.59	-2.50	-1.22	-1.29	0.52	-0.24	-1.50
Netherlands	Min	-18.69	-12.41	-3.28	-7.69	0.07	-3.45	-0.39
	Max	-0.57	2.21	-0.93	1.75	1.21	0.54	-0.39
	Avg	-2.65	-1.98	-0.70	-0.79	0.64	-1.04	-0.39
Poland	Min	-20.35	-13.18	-4.73	-9.04	-0.61	-3.36	-0.98
	Max	-1.20	3.25	-1.64	2.20	2.22	1.80	-0.98
	Avg	-3.11	-1.88	-1.06	-1.66	0.81	-0.88	-0.98
Portugal	Min	-20.21	-14.72	-5.73	-7.35	-1.86	-4.68	-1.40
	Max	0.23	0.55	-3.48	0.90	0.48	-0.85	-1.40
	Avg	-2.95	-2.59	-1.53	-1.12	-0.69	-2.34	-1.40

Fig. 16 Own and cross-product elasticities (average for 10 countries)

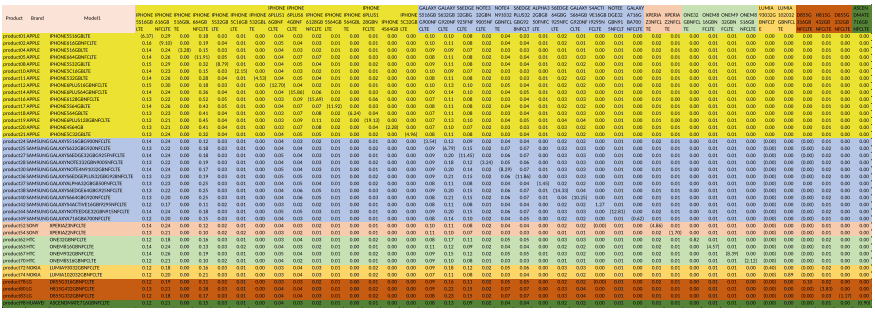


Fig. 17 Brand own-price elasticities—statistics for 10 countries

12 Brand	(Multiple Items)				
Count of Model	**Tier**				
Brand	**Basic**	**Step-Up**	**High**	**Premium**	**Grand Total**
ALCATEL	398	48	14		460
APPLE	14	21	26	38	99
BLACKBERRY	56	55	36	29	176
HTC	150	152	119	79	500
HUAWEI	162	78	39	32	311
LG	371	142	62	44	619
MICROSOFT	21	10	1	7	39
MOTOROLA	212	78	60	35	385
NOKIA	407	181	84	46	718
SAMSUNG	783	368	190	138	1479
SONY	245	156	79	52	532
WIKO	89	27	4		120
Grand Total	**2908**	**1316**	**714**	**500**	**5438**

Fig. 18 Number of products per brand and pricing group

Country	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**GrandTotal**
Austria	128	163	161	181	215	212	171
Belgium	148	176	217	234	263		206
Czechia	134	136	147	148	173	192	150
France	86	115	140	153	169	169	134
Germany	114	139	170	172	198	204	159
Italy	131	162	172	191	218	242	176
Netherlands	232	251	288	310	344	356	286
Poland	27	40	51	73	94	103	59
Portugal	88	104	144	159	174	191	137
Spain	75	94	114	115	144	155	107
Switzerland	167	170					168
Average	**107**	**131**	**153**	**163**	**187**	**196**	**149**
ASP Growth rate		23%	17%	7%	14%	5%	13%

Fig. 19 Development of Shares of Tiers

Country	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**GrandTotal**
Austria	128	163	161	181	215	212	171
Belgium	148	176	217	234	263		206
Czechia	134	136	147	148	173	192	150
France	86	115	140	153	169	169	134
Germany	114	139	170	172	198	204	159
Italy	131	162	172	191	218	242	176
Netherlands	232	251	288	310	344	356	286
Poland	27	40	51	73	94	103	59
Portugal	88	104	144	159	174	191	137
Spain	75	94	114	115	144	155	107
Switzerland	167	170					168
Average	**107**	**131**	**153**	**163**	**187**	**196**	**149**
ASP Growth rate		23%	17%	7%	14%	5%	13%

Fig. 20 Average sales price of handsets in Europe (ITC)

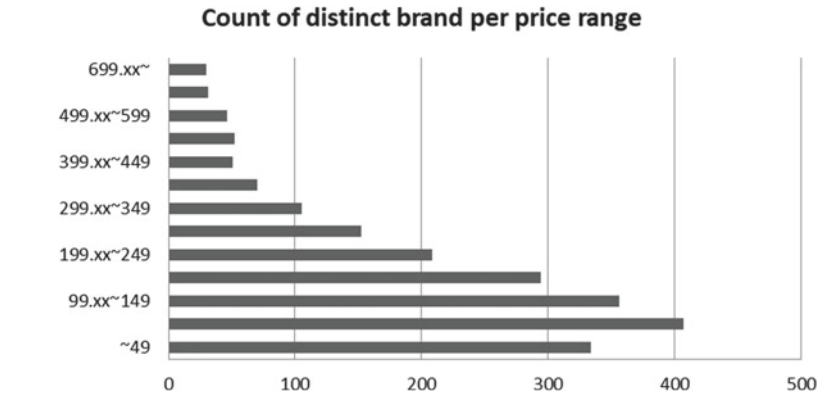


Fig. 21 Number of brands per pricing group in Europe (G2K)

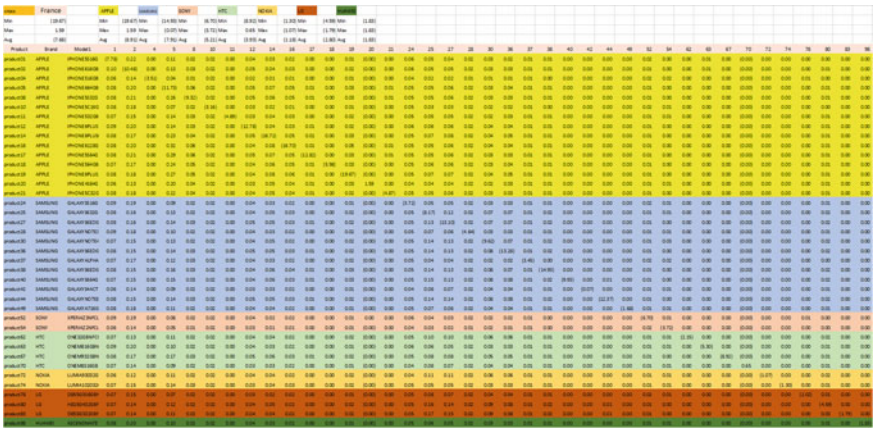


Fig. 22 Cross product-price elasticities of some products, France

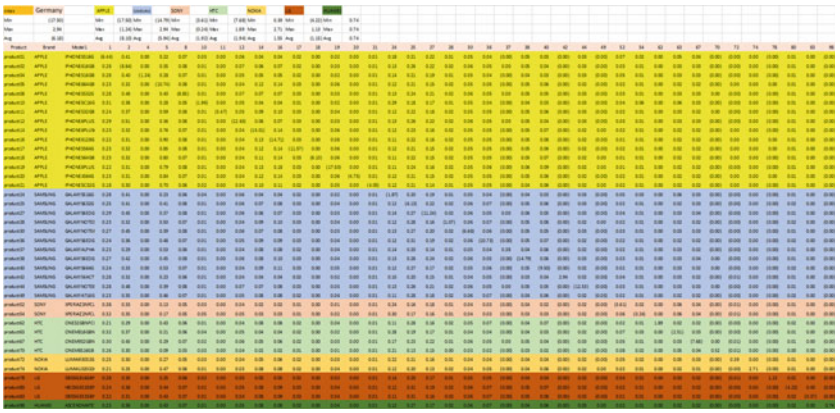


Fig. 23 Cross product-price elasticities of some products, Germany

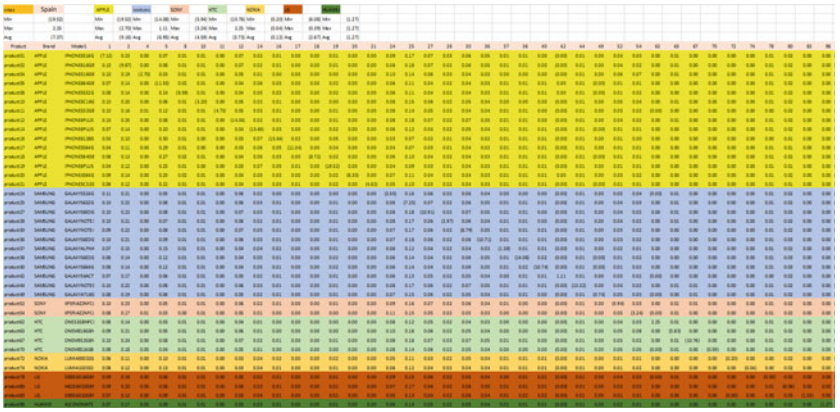


Fig. 24 Cross product-price elasticities of some products, Spain

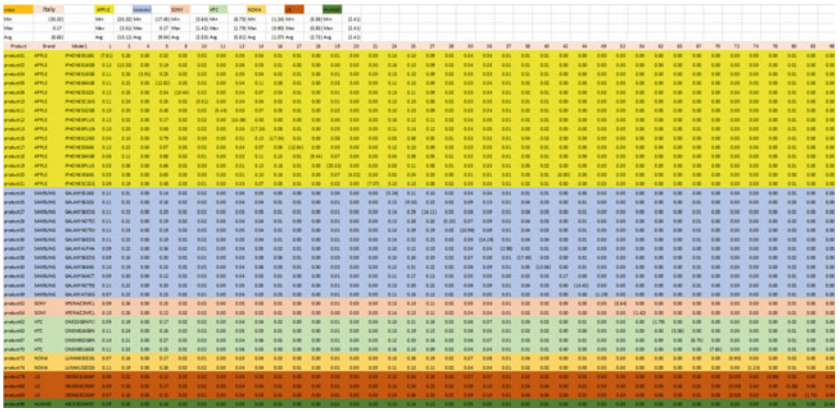


Fig. 25 Cross product-price elasticities of some products, Italy

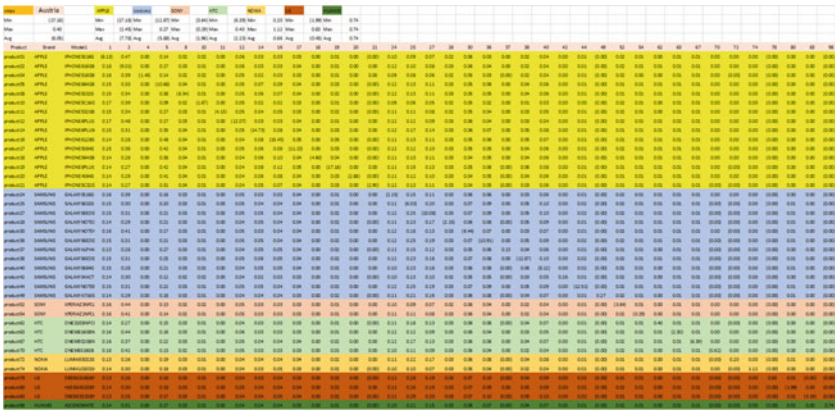


Fig. 26 Cross product-price elasticities of some products, Austria

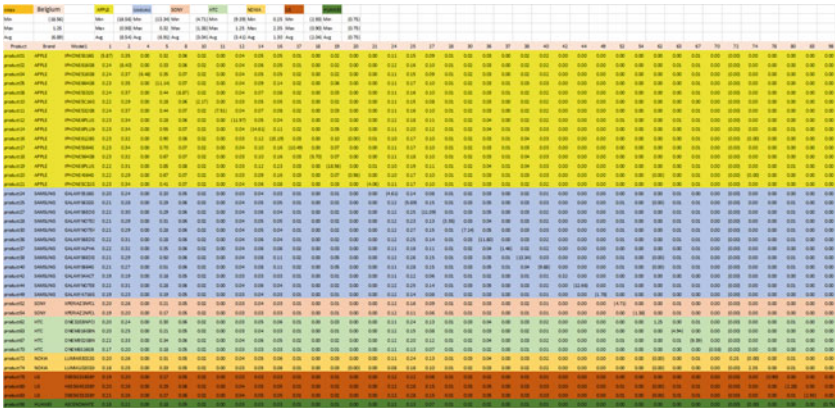


Fig. 27 Cross product-price elasticities of some products, Belgium

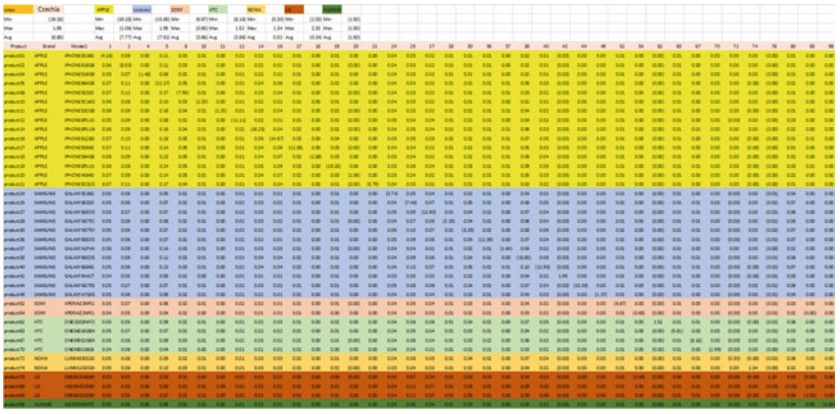


Fig. 28 Cross product-price elasticities of some products, Croatia

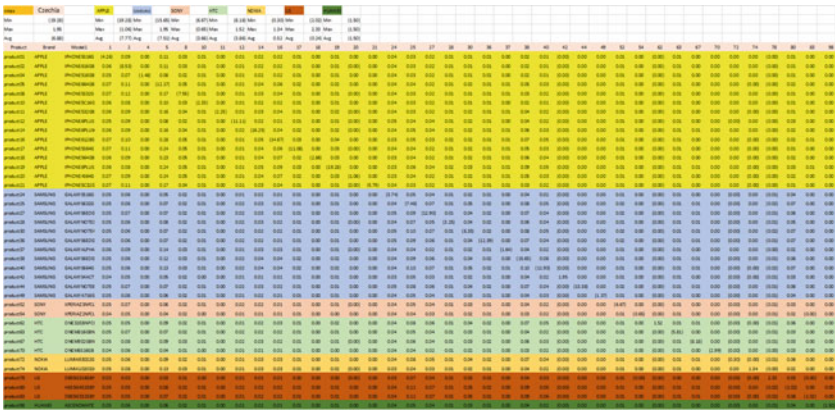


Fig. 29 Cross product-price elasticities of some products, Czech Republic

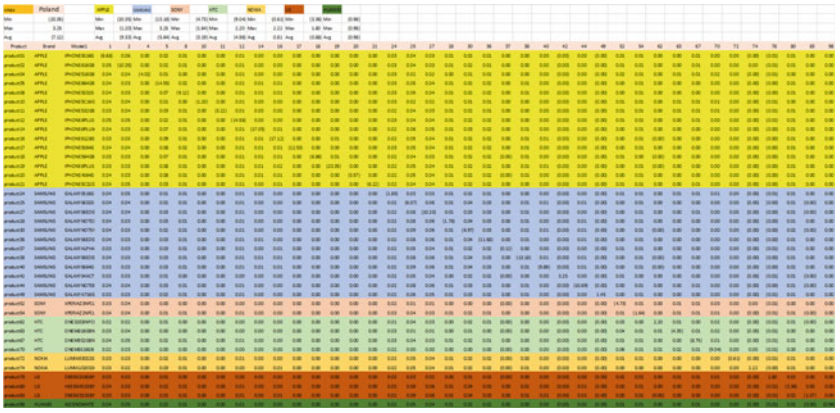


Fig. 30 Cross product-price elasticities of some products, Poland

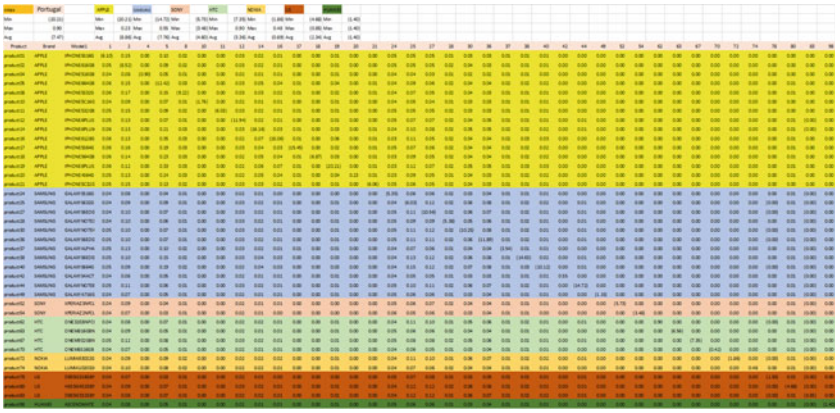


Fig. 31 Cross product-price elasticities of some products, Portugal

References

1. Berry, S., Levinsohn, J., Pakes, A.: Automobile prices in market equilibrium. Econometrica **63**(4), 841–90 (1995)
2. Nevo, A.: A practitioner's guide to estimation of random-coefficients logit models of demand. J. Econ. Manag. Strategy **9**(4), 513–548 (2000)
3. McFadden, D.: The measurement of urban travel demand. J. Public Econ. **3**(4), 303–328 (1974)
4. Dubé, J.P., Fox, J.T., Su, C.L.: Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. Econometrica **80**(5), 2231–2267 (2012)
5. Anderson, S.P., De Palma, A.: Spatial price discrimination with heterogeneous products. Rev. Econ. Stud. **55**(4), 573–592 (1988)
6. Anderson, S.P., De Palma, A., Thisse, J.F.: Discrete Choice Theory of Product Differentiation. MIT press (1992)
7. Gallego, G., Wang, R.: Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. Oper. Res. **62**(2), 450–461 (2014)

8. Armstrong, M., Vickers, J.: Which demand systems can be generated by discrete choice? J. Econ. Theory **158**, 293–307 (2015)
9. Suryanegara, M., Miyazaki, K.: A challenge towards 4g: the strategic perspective of japanese operators in a mature market. In: Technology Management for Global Economic Growth (PICMET). IEEE, pp. 1–9 (2010)
10. Freire Kastner, B.A.: Pricing: A Theoretical Approach and Modern Business Practices. Master's thesis, King's College London, Londres (2012)
11. Bhargava, H.K., Gangwar, M.: Mobile telephony pricing in emerging markets. In: INFORMS Conference on Information Systems and Technology, Minneapolis, MN (2013)
12. Kim, Y.E., Lee, J.W.: Relationship between corporate image and customer loyalty in mobile communications service markets. African J. Bus. Manag. **4**(18), 4035–4041 (2010)
13. Bidyarthi, H.J., Srivastava, A.K., Bokad, P., Deshmukh, L.: Case study-nokia's strategies in indian mobile handsets markets during 2002 to 2006. Int. J. **6**(2), 178–188 (2011)
14. Prasad, V.V., Sahoo, P.: Competitive advantage in mobile phone industry. Int. J. Comput. Sci. Commun. **2**(2), 615–619 (2011)
15. Dedrick, J., Kraemer, K.L., Linden, G.: The distribution of value in the mobile phone supply chain. Telecommun. Policy **35**(6), 505–521 (2011)
16. Kraemer, K.L., Linden, G., Dedrick, J.: Capturing value in global networks: Apple's ipad and iphone. In: Research supported by grants from the Alfred P Sloan Foundation and the US National Science Foundation (CISE/IIS) (2011)
17. Peppard, J., Rylander, A.: From value chain to value network: insights for mobile operators. Euro. Manag. J. **24**(2–3), 128–141 (2006)
18. Cave, M.: Six degrees of separation: operational separation as a remedy in european telecommunications regulation. Commun. Stratégies **2006**(64), 89–103 (2006)
19. ITU (2008) Practice note—ICT regulation toolkit. www.ictregulationtoolkit.org. Accessed 18 04 2020
20. Liozu, S.M., Hinterhuber, A.: Pricing orientation, pricing capabilities, and firm performance. Manag. Decision **51**(3), 594–614 (2013)
21. Bresnahan, T.F.: Departures from marginal-cost pricing in the american automobile industry: estimates for 1977–1978. J. Econ. **17**(2), 201–227 (1981)
22. Bresnahan, T.F.: Competition and collusion in the american automobile industry: the 1955 price war. J. Indus. Econ. 457–482 (1987)
23. Reynaert, M., Verboven, F.: Improving the performance of random coefficients demand models: the role of optimal instruments. J. Econ. **179**(1), 83–98 (2014)
24. Hausman, J., Leonard, G., Zona, J.D.: Competitive analysis with differentiated products. Ann. Econ. Stat. **34**, 143–157 (1994)
25. Aguirregabiria, V., Ho, C.Y.: A dynamic oligopoly game of the us airline industry: estimation and policy experiments. J. Econ. **168**(1), 156–173 (2012)
26. Xu, C., Ji, J., Liu, P.: The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. Transp. Res. part C: Emerging Technol. **95**, 47–60 (2018)
27. Badruddoza, S., Amin, M.D.: Determining the importance of an attribute in a demand system: structural versus machine learning approach. In: Annual Meeting, p. 291210. Atlanta, Georgia, Agricultural and Applied Economics Association (2019)

On Bertelson-Gromov Dynamical Morse Entropy

Artur O. Lopes and Marcos Sebastiani

Abstract In this mainly expository paper we present a detailed proof of several results contained in a paper by M. Bertelson and M. Gromov on Dynamical Morse Entropy. This is an introduction to the ideas presented in that work. Suppose M is compact oriented connected C^∞ manifold of finite dimension. Assume that $f_0 : M \to [0, 1]$ is a surjective Morse function. For a given natural number n, consider the set M^n and for $x = (x_0, x_1, \ldots, x_{n-1}) \in M^n$, denote $f_n(x) = \frac{1}{n} \sum_{j=0}^{n-1} f_0(x_j)$. The Dynamical Morse Entropy describes for a fixed interval $I \subset [0, 1]$ the asymptotic growth of the number of critical points of f_n in I, when $n \to \infty$. The part related to the Betti number entropy does not requires the differentiable structure. One can describe generic properties of potentials defined in the XY model of Statistical Mechanics with this machinery.

Keywords Bertelson-Gromov dynamical morse entropy · Asymptotic growth of critical points · Singular homology · Betti number entropy

1 Introduction

We follow the main guidelines and notation of [1].

A Morse function is a smooth function such all critical points are not degenerate (see [2]).

Suppose M is compact oriented C^∞ manifold of dimension $q \geq 1$. Assume that $f_0 : M \to [0, 1]$ is a surjective Morse function and Γ is a free group with basis $\gamma_1, \ldots, \gamma_n$. We assume that f_0 has p critical points ($p \geq 2$).

Suppose $\Omega \subset \Gamma$ is a finite non-empty set. If $x \in M^\Omega$ we denote $x_\gamma \in M, \gamma \in \Omega$, the corresponding coordinate.

Then, we define $f_\Omega : M^\Omega \to [0, 1]$ by the expression

Instituto de Matematica—UFRGS—Brasil A. O . Lopes was partially supported by CNPq and INCT.

A. O. Lopes (✉) · M. Sebastiani
Instituto de Matemática e Estatística - UFRGS, Porto Alegre, Brazil

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_13

$$f_\Omega(x) = \frac{1}{|\Omega|} \sum_{\gamma \in \Omega} f_0(x_\gamma),$$

where $|\Omega|$ is the cardinality of Ω. This function f_Ω is also a surjective Morse function.

2 The *X Y* Model

As a particular case we can consider $\Gamma = \mathbb{Z}$, the set $M^{\mathbb{Z}}$ and for $x = (x_j)_{j \in \mathbb{Z}} \in M^{\mathbb{Z}}$, $n > 0$, $f_0 : M \to \mathbb{R}$, and

$$f_n(x) = -\frac{1}{n} \sum_{j=0}^{n-1} f_0(x_j).$$

We mention this case because it is a more well known model in the literature and we want to trace a parallel to what will be done here.

The question about the minus sign in front of the sum is not important but if we want that f_0 represents a kind of Hamiltonian (energy) we will keep the—(at least in this section).

In this model it is natural to consider that adjacent molecules in the lattice interact via a potential (an Hamiltonian) which is described by the smooth function of two variables f_0. The mean energy up to position n is described by f_n. The points $x \in M^n$ where the mean n-energy is lower or higher are of special importance. We are interested here, among other things, in the growth of the number of critical values, when $n \to \infty$. The critical points are called the stationary states (see [1]).

Denote by $\mathrm{Cri}_n(I)$ the number of critical points of f_n in a certain interval $f^{-1}(I)$. Roughly speaking the purpose of [1] is to provide for a fixed value $c \in [0, 1]$ a topological lower bound for

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{\log(\mathrm{Cri}_n(I))}{n}, \text{ where } I = (c - \delta, c + \delta),$$

in terms of a certain strictly positive concave function (a special kind of entropy). This is done by taking into account the homological behavior of the functions f_n.

The so called classical XY model consider the case where $M = S^1$ (see for instance [3–9] or [10]). A function $A : (S^1)^{\mathbb{Z}} \to \mathbb{R}$ describes an interaction between sites on the lattice $\mathbb{Z}$ where the spins are on S^1. One is interested in equilibrium probabilities $\hat{\mu}$ on $(S^1)^{\mathbb{Z}}$ which are invariant for the shift $\hat{\sigma} : (S^1)^{\mathbb{Z}} \to (S^1)^{\mathbb{Z}}$. A point x on $(S^1)^{\mathbb{Z}}$ is denoted by $x = (\ldots, x_{-2}, x_{-1} \,|\, x_0, x_1, x_2, \ldots)$.

In the case the potential A depend just on the first coordinate $x_0 \in S^1$, that is $A(x) = f_0(x_0)$, then the setting described above applies.

In the case the potential A depend just on the two first coordinate $x_0, x_1 \in S^1$, that is $A(x) = f_0(x_0, x_1)$, then, we claim that the setting described above in the introduction applies. This is the case when $f_0 : S^1 \times S^1 \to \mathbb{R}$. Indeed, in this case

one can take $M = S^1 \times S^1$ and consider that f_0 acts on M. In this case we can say that f_0 depends just in the first coordinate on $M^{\mathbb{Z}} = (S^1 \times S^1)^{\mathbb{Z}}$ and adapt the general formalism we describe here.

Therefore, we will state all results for $f_0 : M \to \mathbb{R}$, that is, the case the potential on $M^{\mathbb{Z}}$ depends just on the first coordinate.

In the case $\hat{\mu}$ is ergodic the sequence f_n describes Birkhoff means which are $\hat{\mu}$ almost everywhere constant. We are here interested more in the topological and not in the measure theoretical point of view.

In the measure theoretical (or Statistical Mechanics) point of view, if one is interested in equilibrium states at positive temperature $T = 1/\beta$, then, is natural to consider expressions like $\int e^{\sum_{j=0}^{n-1} -\beta f_0(x_j)} dx_0 dx_1 \dots d_{x_{n-1}}$ (or, when the set of spins is finite: $\sum e^{\sum_{j=0}^{n-1} -\beta f_0(x_j)}$) and its normalization (see [11–13]) which defines the partition function.

By the other hand if one is interested in the zero temperature case (see for instance [14]), then, expressions like $-\sum_{j=0}^{n-1} f_0(x_j)$ are the main focus. For instance, if f_0 has a unique point of minimum $x^- \in S^1$, then $\delta_{(x^-)^\infty}$ defines the ground state (maximizing probability). In the generic case the function f_0 has indeed a unique point of minimum.

Given $f_0 : M \times M \to \mathbb{R}$ and n one can also consider periodic conditions. In this case we are interested in sums like

$$\tilde{f_n}(x) = -\frac{1}{n}\,(f_0(x_0) + f_0(x_1) + \cdots + f_0(x_{n-2}) + f_0(x_0)),$$

or

$$-(f_0(x_0) + f_0(x_1) + \cdots + f_0(x_{n-2}) + f_0(x_0)).$$

In the case we want to get Gibbs states via the Thermodynamic Limit (see for instance [11] or [13]), given a natural number n, we have to look for the probability μ on M^n (absolutely continuous with respect to Lebesgue probability) which maximizes

$$\int e^{-\sum_{j=0}^{n-1} \beta f_0(x_j)}\, d\,\mu(dx_0,\, dx_1, \dots, d_{x_{n-1}}),$$

or, at zero temperature the periodic probability μ on M^n which maximizes

$$-\int \sum_{j=0}^{n-1} f_0(x_j)\; d\,\mu(dx_0,\, dx_1, \dots, d_{x_{n-1}}).$$

One can easily adapt the reasoning of [15] to show that for a generic f_0 we get that $\tilde{f_n}$ is a Morse function for all n.

When f_0 is not generic several pathologies can occur (see for instance [3, 5, 10]).

Suppose the case when there is a unique point x^- of minimum for f_0. For each $\beta > 0$ and n denote by $\mu_{n,\beta}$ the absolutely continuous with respect to Lebesgue probability which maximizes

$$\int e^{-\sum_{j=0}^{n-1}\beta\, f_0(x_j)}\, d\,\mu(dx_0,\, dx_1, \ldots, d_{x_{n-1}}).$$

By the Laplace method (adapting Proposition 3 in [7] or Lemma 4 in [8]) we get that when $\beta \to \infty$ and $n \to \infty$ the probability $\mu_{n,\beta}$ converges to the Dirac delta on $(x^-)^\infty$. Therefore, in the generic case this last probability is the ground state (zero temperature limit).

3 The General Model—The Dynamical Morse Entropy

From now we forget the—sign in front of f_0. For instance, $f_n(x) = \frac{1}{n}\sum_{j=0}^{n-1} f_0(x_j, x_{j+1})$.

Given $c \in [0, 1]$ and $\delta > 0$, take $N_\Omega(c, \delta)$ the number of critical points of f_Ω in $f_\Omega^{-1}[c-\delta, c+\delta]$. Note that if f_0 has p critical points then f_Ω has $p^{|\Omega|}$ critical points.

Consider the cylinder sets

$$\Omega_i \,=\, \{a_1\,\gamma_1 + \cdots + a_n\gamma_n\,;\ |a_j| \le i,\ 1 \le\ j \le n\,\},\ \ i = 1, 2, \ldots,$$

where a_j are integers.

Denote $N_i(c, \delta) = N_{\Omega_i}(c, \delta)$. Then, of course, $N_i(c, \delta)$ for c fixed decrease with δ.

For a fixed $0 \le c \le 1$, we denote the entropy by

$$\varepsilon(c) = \lim_{\delta\to 0}\left(\liminf_{i\to+\infty} \frac{\log(N_i(\,c,\,\delta)\,)}{|\Omega_i|}\right).$$

The above limit exists and it is bounded by $\log p$ but in principle could take the value $-\infty$. We call $\varepsilon(c)$ the **dynamical Morse entropy** on the value c.

In the case $\Gamma = \mathbb{Z}$ as we mentioned before $\varepsilon(c)$ is described by

$$\varepsilon(c) = \lim_{\delta\to 0}\left(\liminf_{n\to+\infty} \frac{\log(\text{number of critical points of } f_n \text{ in } f_n^{-1}[c-\delta, c+\delta]\,)}{n}\right).$$

Later we introduce a function $b(c)$ (see Definition 3 and also Definition 2), which will be a topological invariant of f_0. The function $b(c)$ is defined in terms of rank of linear operators and Cohomology groups.

We will show later that

(1) $0 \le b(c) \le \varepsilon(c),\ 0 \le c \le 1$;

(2) $b(c)$ is continuous and concave;

(3) $b(c)$ is not constant equal to 0.

Finally, in the case $M = S^1$ (the unitary circle) and f_0 has just two critical points, we show in Sect. 7 that

$$\varepsilon(c) = b(c) = -c \log c - (1-c)\log(1-c).$$

$b(c)$ is sometimes called the **Betti entropy** of f_0.

Our definition of $b(c)$ is different from the one in [1] but we will show later (see Sect. 8) that is indeed the same.

A key result in the understanding of the main reasoning of the paper is Lemma 6 which claims that for any Morse function f, given $a, b \in \mathbb{R}$, $a < b$, the number of critical points of f in $f^{-1}[a, b]$ is bigger or equal to the dimension of the vector space

$$\frac{H^*(f^{-1}(\infty, b))}{H^*(f^{-1}(-\infty, a))},$$

where H^* denotes the corresponding cohomology groups which will be defined in the following paragraphs (see also [16] for basic definitions and properties).

$H^*(X, \mathbb{R})$ denotes the usual cohomology. Note that H^* will have another meaning (see Definition 1).

4 Cohomology

Suppose X is a metrizable, compact, oriented topological manifold C^∞ manifold. We will consider the singular homology. Suppose $U \subset X$ is an open set and $a \in H^*(X, \mathbb{R})$. The meaning of the statement supp $a \subset U$ is: there exist an open set $V \subset X$, such that, $X = U \cup V$, and $a|_V = 0$.

Definition 1 $H^*_X(U) = \{a \in H^*(X, \mathbb{R}) : \text{supp } a \subset U\}$, where U is an open subset of X. When X is fixed we denote $H^*_X(U) = H^*(U)$.

Remember (see for instance [16]) that when $U \subset X$ is open we get the exact cohomology sequence:

$$... \to H^{k-1}(X-U, \mathbb{R}) \to H^k_c(U, \mathbb{R}) \to H^k(X, \mathbb{R}) \to H^k(X-U, \mathbb{R}) \to H^{k+1}_c(U, \mathbb{R}) \to ... \tag{1}$$

where H^*_c denotes the compact support cohomology.

Lemma 1 *If U is an open set, then*

$$H^*(U) = Im(\ H^*_c(U, \mathbb{R}) \to H^*(X, \mathbb{R})\ ) = Ker(\ H^*(X, \mathbb{R}) \to H^*(X-U, \mathbb{R})\ ).$$

Proof The second equality follows from the fact that the above sequence is exact.

We will prove that

$$\text{Im}(\ H_c^*(U,\mathbb{R}) \to H^*(X,\mathbb{R})\ ) \subset H^*(U) \subset \text{Ker}(\ H^*(X,\mathbb{R}) \to H^*(X-U,\mathbb{R})\ ).$$

Let $a \in \text{Im}(\ H_c^*(U,\mathbb{R}) \to H^*(X,\mathbb{R})\ )$. Then, a is represented by a cocycle α with compact support $K \subset U$. Therefore, $a\,|(X-K) = 0$.

Defining $V = X - K$ we have that $U \cup V = X$ and $a\,|V = 0$. Then, $a \in H^*(U)$.

Let be $\alpha \in H^*(U)$. Let $V \subset X$ be an open set such that $U \cup V = X$ and $\alpha\,|V = 0$.

Since $X - U \subset V$, we have $\alpha\,|(X-U) = 0$.

Then, $\alpha \in \text{Ker}(\ H^*(X,\mathbb{R}) \to H^*(X-U,\mathbb{R})\ )$. □

Lemma 2 *If U is an open set then $H^*(U)$ is a graded ideal of the ring of cohomology of X.*

Proof This follows at once from Lemma 1. □

Now we consider a continuous function $f : X \to \mathbb{R}$.

Definition 2 Given $\delta > 0$ and $c \in \mathbb{R}$ we define

$$b'_{c,\delta} = Dim\left(\frac{H^*(f^{-1}(-\infty, c+\delta)))}{H^*(f^{-1}(-\infty, c-\delta))}\right).$$

Proposition 1 *Suppose X and Y are metrizable compact, oriented topological manifolds, moreover take $f : X \to \mathbb{R}$, $g : Y \to \mathbb{R}$ continuous functions. If we define $f \oplus g : X \times Y \to \mathbb{R}$, by $(f \oplus g)(x,y) = f(x) + g(y)$, then, if $c, c' \in \mathbb{R}$, $\delta, \delta' > 0$, we get*

$$b'_{c,\,\delta}(f)\ b'_{c',\,\delta'}(g)\ \leq b'_{c+c',\,\delta+\delta'}(f \oplus g). \tag{2}$$

Before the proof of this important proposition we need two more lemmas.

As it is known (see [16]) the cup product $\vee$ defines an isomorphism

$$\mu : H^*(X,\mathbb{R}) \otimes H^*(Y,\mathbb{R}) \ \to\ H^*(X \times Y, \mathbb{R}).$$

Lemma 3 *If $U \subset X$ and $V \subset Y$ are open sets, then*

$$\mu(\ H_X^*(U) \otimes H^*(Y,\mathbb{R}) + H^*(X,\mathbb{R}) \otimes H_Y^*(V)\ )\ = H_{X\times Y}^*(\ (U \times Y) \cup (X \times V)\ ).$$

Proof By Lemma 1 we get

$$H_{X\times Y}^*(\ (U \times Y) \cup (X \times V)\ ) = \text{Ker}(\ H^*(X \times Y,\mathbb{R}) \to H^*((X-U) \times (Y-V),\mathbb{R}).$$

Then,

$$H_{X\times Y}^*(\ (U \times Y) \cup (X \times V)\ ) =$$

$$\mu(\ \text{Ker}(\ H^*(X,\mathbb{R}) \otimes H^*(Y,\mathbb{R}) \ \to H^*(X-U,\mathbb{R}) \otimes H^*(Y-V,\mathbb{R}))\ ).$$

From simple Linear Algebra arguments the claim follows from Lemma 1. □

Lemma 4 *If $U \subset X$ and $V \subset Y$ are open sets then*

$$\mu(\, H_X^*(U) \otimes H_Y^*(V)\,) = H_{X\times Y}^*(U \times V).$$

Proof The $\vee$ product defines a natural isomorphism

$$H^*(X, X-U, \mathbb{R}) \otimes H^*(Y, Y-V, \mathbb{R}) \to H^*(X\times Y, (X\times(Y-V)\cup(X-U)\times Y, \mathbb{R}) =$$

$$H^*(\, X\times Y, (X\times Y) - (U\times V, \mathbb{R})\,).$$

By Lemma 1 and the exact relative cohomology sequence we get:

$$H_X^*(U) = \mathrm{Im}\,(\, H^*(X, X-U, \mathbb{R}) \to H^*(X, \mathbb{R})\,),$$

$$H_Y^*(V) = \mathrm{Im}\,(\, H^*(Y, Y-V, \mathbb{R}) \to H^*(Y, \mathbb{R})\,),$$

and

$$H_{X\times Y}^*(U\times V) = \mathrm{Im}\,(\, H^*(X\times Y, (X\times Y) - (U\times V, \mathbb{R})) \to H^*(X\times Y, \mathbb{R})\,).$$

From this the claims follows at once. □

Now we will present the proof of Proposition 1.
Take $h = f \oplus g$ and denote

$$A^- = f^{-1}(-\infty, c-\delta),\ B^- = g^{-1}(-\infty, c'-\delta'),\ C^- = h^{-1}(-\infty, (c+c') - (\delta+\delta'))\ ,$$

and

$$A^+ = f^{-1}(-\infty, c+\delta),\ B^+ = g^{-1}(-\infty, c'+\delta'),\ C^+ = h^{-1}(-\infty, (c+c') + (\delta+\delta'))\ .$$

Note that

$$A^+ \times B^+ \subset C^+ \subset (A^+ \times Y) \cup (X \times B^+)$$

$$A^- \times B^- \subset C^- \subset (A^- \times Y) \cup (X \times B^-).$$

Consider the commutative diagram

$$
\begin{array}{ccc}
H^*(X,\mathbb{R})\otimes H^*(Y) & \to \text{ (using } \mu\text{)} & H^*(X\times Y,\mathbb{R}) \\
\cup & & \cup \\
H^*_X(A^+)\otimes H^*_Y(B^+) \to H^*_{X\times Y}(C^+) & \subset & H^*_{X\times Y}((A^+\times Y)\cup(X\times B^+)) \\
\cup & & \cup \\
H^*_X(A^+)\otimes H^*_Y(B^-)+H^*_X(A^-)\otimes H^*_Y(B^+) \to H^*_{X\times Y}((A^-\times Y)\cup(X\times B^-)) & & \\
\cup & & \cup \\
H^*_X(A^-)\otimes H^*_Y(B^-) \to H^*_{X\times Y}(C^-). & &
\end{array}
$$

From this follows the linear transformation

$$\tilde{\mu}:\ \frac{H^*_X(A^+)\otimes H^*_Y(B^+)}{H^*_X(A^-)\otimes H^*_Y(B^-)} \to \frac{H^*_{X\times Y}(C^+)}{H^*_{X\times Y}(C^-)}.$$

By the other hand

$$(H^*_X(A^+)\otimes H^*_Y(B^+)\cap\mu^{-1}(H^*_{X\times Y}(C^-))\subset$$

$$(H^*_X(A^+)\otimes H^*_Y(B^+)\cap\mu^{-1}(H^*_{X\times Y}((A^-\times Y)\cup(X\times B^-))) =$$

$$(H^*_X(A^+)\otimes H^*_Y(B^+))\cap(H^*_X(A^-)\otimes H^*(Y,\mathbb{R})+H^*(X,\mathbb{R})\otimes H^*_Y(B^-)) =$$

$$H^*_X(A^-)\otimes H^*_Y(B^+)\ +\ H^*_X(A^+)\otimes H^*_Y(B^-).$$

The first equality above follows from Lemma 3; the second follows from Linear Algebra; namely, if $E_2\subset E_1\subset E$ and $F_2\subset F_1\subset F$, then

$$(E_1\otimes F_1)\cap(E_2\otimes F+E\otimes F_2)=E_2\otimes F_1+E_1\otimes F_2.$$

From the above it follows that

$$\operatorname{Ker}\tilde{\mu}\ \subset\ \frac{H^*_X(A^-)\otimes H^*_Y(B^+)\ +\ H^*_X(A^+)\otimes H^*_Y(B^-)}{H^*_X(A^-)\otimes H^*_Y(B^-)}.$$

Therefore,

$$b'_{c+c',\delta+\delta'}\ =\ \dim\frac{H^*_{X\times Y}(C^+)}{H^*_{X\times Y}(C^-)}\geq\dim(\operatorname{Im}\tilde{\mu})\geq$$

$$\dim \frac{H_X^*(A^+) \otimes H_Y^*(B^+)}{H_X^*(A^-) \otimes H_Y^*(B^+) + H_X^*(A^+) \otimes H_Y^*(B^-)} =$$

$$\dim \left(\frac{H_X^*(A^+)}{H_Y^*(A^-)} \otimes \frac{H_Y^*(B^+)}{H_Y^*(B^-)} \right) = b'_{c,\delta}(f)\ b'_{c',\delta'}(g). \qquad \square$$

5 Critical Points

In what follows X is a compact, oriented C^∞ manifold and $f : X \to \mathbb{R}$ is a Morse function.

Lemma 5 *Suppose X is a compact, oriented C^∞ manifold and $U \subset X$ is an open set. If $a \in H^*(X, \mathbb{R})$, then, supp $a \subset U$, if and only if, there exists a closed C^∞ differentiable form w such that supp $w \subset U$, and a is the de Rham cohomological class of w.*

Proof If there exists $w \in a$, such that supp $w \subset U$, then

$$a|_{(X - \text{supp } w)} = 0 \text{ and } U \cup (X - \text{supp } w) = X.$$

If there exists an open set $V \subset X$ such that $U \cup V = X$ and $a|_V = 0$, then, there exist a C^∞ form η on V such that $d\eta = w|_V$ where $w \in a$.

Let W be an open set such that $\overline{W} \subset V$ and $W \cup U = X$. Take a C^∞ function $\varphi : X \to [0, 1]$ such that $\varphi|_{\overline{W}} = 1$ and $\varphi|_{X-K} = 0$, where K is compact set such that $\overline{W} \subset K \subset V$. Then, $\varphi\, \eta$ has an extension to X and $(w - d(\varphi\, \eta)) \in a$. But,

$$\text{supp } (w - d(\varphi\, \eta)) \subset X - W \subset U. \qquad \square$$

Lemma 6 *Given $a, b \in \mathbb{R}$, $a < b$, then, the number of critical points of f in $f^{-1}[a, b]$ is bigger or equal that*

$$dim\ \frac{H^*(f^{-1}(\infty, b))}{H^*(f^{-1}(-\infty, a))}.$$

Proof Without lost of generality we can assume that a and b are regular values of f (decrease a and increase b a little bit).

Given $c_1 < c_2 < \cdots < c_m$, the critical values of f in (a, b), take

$$a = d_0 < c_1 < d_1 < c_2 < d_2 < \cdots < d_{m-1} < c_m < d_m = b.$$

By Proposition 3 and Lemma 8, the number of critical points in $f^{-1}(c_i)$, $i = 1 =$, $2, \ldots, m$, is bigger or equal to

$$dim\ \frac{H^*(f^{-1}(\infty, d_i))}{H^*(f^{-1}(-\infty, d_{i-1}))}.$$

Finally consider the filtration

$$H^*(f^{-1}(\infty, a) = H^*(f^{-1}(-\infty, d_0)) \subset H^*(f^{-1}(-\infty, d_1)) \subset ...$$

$$\subset H^*(f^{-1}(-\infty, d_{m-1}) \subset H^*(f^{-1}(-\infty, d_m)) = H^*(f^{-1}(-\infty, b)).$$

□

Now we denote $b'_\Omega(c, \delta) = b'_{c,\delta}(f_\Omega)$ and $b'_i(c, \delta) = b'_{\Omega_i}(c, \delta)$, $0 \le c \le 1$, $\delta > 0$.

Corollary 1 *$b'_i(c, \delta) \le N_i(c, \delta)$ for all $i = 1, 2, 3, \ldots$ and $0 \le c \le 1$, $\delta > 0$.*

Now we define the function b using Proposition 3(a)

Definition 3

$$b(c) = \lim_{\delta \to 0} \liminf_{i \to \infty} \frac{\log(b'_i(c, \delta))}{|\Omega_i|},\ 0 \le c \le 1.$$

We will show that in above definition we can change the lim inf by lim.

Lemma 7

$$b(c) \le \varepsilon(c) \le \log(\textit{the number of critical points of } f_0).$$

Proof The first inequality follows from Corollary 1. From the definition is easy to see that $\varepsilon(c)$ is smaller than log of the number of critical points of f_0. □

We denote $B(\Gamma)$ a family of finite subsets of Γ and $B_N(\Gamma)$, $N \in \mathbb{N}$, the family of sets $\Omega \in B(\Gamma)$ such that $|\Omega| > N$.

Proposition 2 *Suppose $\Omega', \Omega'' \in B(\Gamma)$ are disjoint not empty sets. Then,*

$$b'_{\Omega \cup \Omega''}(\alpha c_1 + (1 - \alpha)c_2, \delta) \ge b'_{\Omega'}(c_1, \delta)\, b'_{\Omega''}(c_2, \delta),$$

where $0 \le c_1, c_2 \le 1$, $\delta > 0$ and $\alpha = \frac{|\Omega'|}{|\Omega'| + |\Omega''|}$.

Proof By definition

$$f_{\Omega' \cup \Omega''} = \alpha f_{\Omega'} \oplus (1 - \alpha) f_{\Omega''}.$$

By Proposition 1, as $\delta = \alpha\, \delta + (1 - \alpha)\delta$, then

$$b'_{\alpha\, c_1 + (1-\alpha)\, c_2, \delta}(f_{\Omega' \cup \Omega''}) \ge b'_{\alpha\, c_1, \alpha\, \delta}(\alpha\ f_{\Omega'})\, b'_{(1-\alpha)\, c_2, (1-\alpha)\, \delta}((1 - \alpha)\ f_{\Omega''}) =$$

$$b'_{c_1, \delta}(f_{\Omega'})\ \ b'_{c_2, \delta}(f_{\Omega''}).$$

□

Lemma 8 *Suppose the interval* $[a, b]$ *does no contains critical values of* f. *Then,*

$$H^*(\, f^{-1}(-\infty, a\,)\,) = H^*(\, f^{-1}(-\infty, b\,)\,).$$

Proof This follows from Lemma 1 and the fact that $f^{-1}[b,\ \infty)$ is a deformation retract of $f^{-1}[a,\ \infty)$. □

Definition 4 Given $c \in \mathbb{R}$ we define

$$\tilde{b}_c(f) = \lim_{\delta \to 0} b'_{c,\delta}(f).$$

Proposition 3 *For a fixed c we have*
(a) $b'_{c,\delta}(f)$ *decreases with* δ *and* $b'_{c,\delta}(f) = \tilde{b}_c(f)$ *for all* δ *small enough.*
(b) $\tilde{b}_c(f) = 0$ *if c is not a critical value of f*
(c) $\tilde{b}_c(f)$ *is smaller than the number of critical points of f in* $f^{-1}(c)$
(d) $\sum_c \tilde{b}_c(f) = Dim\ H^*(X)$.

Proof (a) follows from the above definitions and Lemma 8.
(b) follows from Lemma 8
For the proof of (c) consider the exact diagram

$$\begin{array}{ccc} & & H^*(X, \mathbb{R}) \\ & \downarrow r_1 & \qquad r_2 \searrow \\ H^*(f^{-1}[c-\delta, \infty)),\, f^{-1}(c+\delta, \infty), \mathbb{R}\,) & \to H^*[\,f^{-1}(c-\delta, \infty), \mathbb{R}) & \to H^*(f^{-1}[c+\delta, \infty), \mathbb{R}), \end{array}$$

where r_1 and r_2 are the restriction homomorphisms.
By Lemma 1

$$H^*(f^{-1}(-\infty, c+\delta)) = \mathrm{Ker}\, r_2 \quad \text{and} \quad H^*(f^{-1}(-\infty, c-\delta)) = \mathrm{Ker}\, r_1.$$

From this follows that

$$b'_{c,\delta}(f) = \ \mathrm{Dim}\ (r_1(\mathrm{Ker}\,(r_2))\ \leq\ \mathrm{Dim}\,(\,H^*(f^{-1}(c-\delta, \infty)),\, f^{-1}(c+\delta, \infty)\,), \mathbb{R}\,)$$

because the above sequence is exact.
In order to finish the proof we apply Morse Theory (see [2]) with δ small enough.
For the proof of (d) suppose $c_1 < c_2 < \cdots < c_m$ are the critical values of f. Now, consider

$$d_0 < c_1 < d_1 < c_2 < d_2 < \cdots < d_{m-1} < c_m < d_m.$$

Now, from (a) and Lemma 8 we have

$$\tilde{b}_{c_i}(f) = \mathrm{Dim}\,(\frac{H^*(\,f^{-1}(-\infty, d_i)\,)}{H^*(\,f^{-1}(-\infty, d_{i-1})\,)}), \quad i = 1, 2, \ldots, m.$$

Finally, note that

$$0 = H^*(\, f^{-1}(-\infty, d_0)) \subset H^*(\, f^{-1}(-\infty, d_1)) \subset ... \subset$$

$$H^*(\, f^{-1}(-\infty, d_m)) = H^*(X).$$

□

Lemma 9 *Given $\delta > 0$, there exists an integer N such that: $b'_\Omega(c, \delta) \geq 1$ for all $c \in [0, 1]$ and all $\Omega \in B_N(\Gamma)$. Therefore, $b(c) \geq 0$, for all $0 \leq c \leq 1$.*

Before the Proof of Lemma 9 we need two more lemmas.

Lemma 10 *Suppose X is a compact oriented C^∞ manifold and $f : X \to \mathbb{R}$ is a Morse function. Then, for all $\delta > 0$*

$$b'_{a_1,\delta}(f) \geq 1 \ \text{ and } \ b'_{a_2,\delta}(f) \geq 1,$$

where a_1 and a_2 are respectively the maximum and minimum of f.

Proof If δ is small enough, $f^{-1}(-\infty, a_2 + \delta)$ is the disjoint union of a finite number of open discs and $f^{-1}(-\infty, a_2 - \delta) = \emptyset$.

If n is the dimension of X, then, it follows from Lemma 1 that

$$H^n(X, \mathbb{R}) \subset H^*(f^{-1}(-\infty, a_2 + \delta)\,) \neq 0$$

and

$$H^*(f^{-1}(-\infty, a_2 - \delta)\,) = 0.$$

Then, $b'_{a_2,\delta}(f) \geq 1$, if $\delta > 0$ is small enough. Therefore, this claim is also true for any $\delta > 0$ by Proposition 3(a).

In a similar way we have that for small $\delta > 0$

$$H^0(X, \mathbb{R}) \subset H^*(f^{-1}(-\infty, a_1 + \delta)\,)$$

and

$$H^0(X, \mathbb{R}) \text{ is not contained } H^*(f^{-1}(-\infty, a_1 - \delta)\,).$$

From this the final claim is proved. □

Lemma 11 *Consider $\Omega \in B(\Gamma)$ where $|\Omega| = m \geq 1$, then, $b'_\Omega(k/m,\ \delta) \geq 1$, for all $\delta > 0$ and $k = 0, 1, 2, \ldots, m$.*

Proof If $k = 0$, or m, the claim follows from Lemma 10 with $X = M^\Omega$, $f = f_\Omega$.

Given $0, k, m$, $0 < k < m$, take $\Omega = \Omega' \cup \Omega''$, where Ω', Ω'' are disjoints and $k = |\Omega'|$.

By Proposition 2 with $c_1 = 1$ and $c_2 = 0$ we get

$$b'_{\Omega}(k/m,\ \delta) \geq b'_{\Omega'}(1,\ \delta)\, b'_{\Omega''}(0,\ \delta) \geq 1.$$

Yet from last lemma. □

Now we will prove Lemma 9.

Proof Take $N > \frac{2}{\delta}$, $\Omega \in B_N(\Gamma)$, $|\Omega| = m > N$ and k such that $\frac{k}{m} \leq c < \frac{k+1}{m}$,
By definition,

$$b'_{c,\delta}(f_\Omega) \geq b'_{k/m,\,\delta/2}(f_\Omega),$$

since $c - \delta < k/m - \delta/2$ and $c + \delta > k/m + \delta/2$.
Therefore, $b'_{\Omega}(c, \delta) \geq b'_{\Omega}(k/m,\ \delta/2) \geq 1$ by Lemma 11. □

Proposition 4

$$0 \leq b(c) \leq \varepsilon(c) \leq \log(\textit{number of critical points of } f_0),\ \ 0 \leq c \leq 1.$$

Proof This follows from Lemmas 7 and 9 □

Lemma 12 *Given $c \in [0, 1]$ and $\delta > 0$, consider a non-empty set $\Omega \in B(\Gamma)$ and $\gamma \in \Gamma$. Then,*

$$b'_{\Omega}(c, \delta) = b'_{\Omega+\gamma}(c, \delta).$$

In the case $\Gamma = \mathbb{Z}$ we have that for any $\Omega = \{1, 2, \ldots, k\}$

$$b'_{\Omega}(c, \delta) = b'_{\hat{\sigma}(\Omega)}(c, \delta),$$

where $\hat{\sigma}$ is the shift acting on $M^{\mathbb{Z}}$.

Proof For fixed γ consider the transformation $x \in M^{\Omega} \to y \in M^{\Omega+\gamma}$, such that $y_w = x_{w-\gamma}$, which is a diffeomorphism which commutes $f_{\Omega+\gamma}$ with f_Ω.
The result it follows from this fact. □

We will show now that indeed one can change lim inf by inf in Definition 3. In order to do that we need the following proposition which describes a kind of subadditivity.

Proposition 5 *Given an integer number $N > 0$ take $h : B_N(\Gamma) \to \mathbb{R}$, $h \geq 0$, which is invariant by Γ and such that*

$$h(\Omega' \cup \Omega'') \geq h(\Omega') + h(\Omega''),$$

if Ω', $\Omega'' \in B_N(\Gamma)$, are disjoint. Then, the limiit

$$\lim_{i\to\infty} \frac{h(\Omega_i)}{|\Omega_i|} \geq 0 \quad \textit{exists: finite or } +\infty\,.$$

From this follows:

Corollary 2 *For $c \in [0,1]$ and $\delta > 0$,*
(a) there exist the limit

$$\lim_{i\to\infty} \frac{\log b_i'(c,\delta)}{|\,\Omega_i\,|} = b'(c,\delta).$$

(b) $0 \le b'(c,\delta) \le \log$ (number of critical points of f_0),
(c) $b(c) = \lim_{\delta\to 0} b'(c,\delta)$

Proof The claim (a) follows from last proposition applied to $h(\Omega) = \log b'_\Omega(c,\delta)$, by Lemma 9, Proposition 2 taking $c_1 = c_2 = c$ and also by Lemma 12.
Item (b) follows from Lemma 2 and Corollary 1.
Item (c) follows from item (a) and the definition of $b(c)$. □

Before the proof of Proposition 5 we need two lemmas.

Lemma 13 *Given an integer positive number k, then for each $i > (3k+1)$ there exists $\Omega_{k,i} \in B(\Gamma)$ such that: (a) $\Omega_{k,i} \subset \Omega_i$; (b) $\Omega_{k,i}$ is a disjoint union of a finite number of translates of Ω_k ; (c) $\lim_{i\to\infty} \frac{|\,\Omega_{k,i}|}{|\,\Omega_i\,|} = 1$; (d) $|\Omega_i| - |\Omega_{k,i}| \ge (2k+1)^n$, where n is the number of generators of Γ.*

Proof For the purpose of the proof we can assume that $\Gamma = \underbrace{\mathbb{Z}\oplus\mathbb{Z}\oplus\cdots\oplus\mathbb{Z}}_{n}$ and take $\gamma_1, \gamma_2, \ldots, \gamma_n$ the canonical basis.
Take $m \ge 1$ an integer such that

$$k + m\,(2\,k + 1) \le i < k = (m+1)\,(2\,k+1),$$

and

$$\Omega_{k,i} = \cup\,\{\ \Omega_k + (\ j_1(2k+1), \ldots, j_n(2k+1)\ )\ |$$

$$-m \le j_1, \ldots, j_n \le m,\ (j_1, \ldots, j_n) \ne (0, \ldots, 0)\ \}.$$

It is easy to see that the sets $\Omega_{k,i}$ satisfy all the above claims. □

Lemma 14 *Given real numbers $x_i \ge 0$ $i = 1, 2, 3, \ldots$, suppose that for each k and each $\varepsilon > 0$ there exist $N_{k,\varepsilon}$ such that*

$$x_i \ge x_k(1-\varepsilon) \text{ if } i \ge N_{k,\varepsilon}.$$

Then, the $\lim_{i\to\infty} x_i$ exists (can finite or $+\infty$).

Proof Take $L = \limsup_{i\to\infty} x_i$ and $a \in \mathbb{R}$, $a < L$. Then, there exists $x_k > a$. Therefore, $x_i \ge a$, if i is very large. Then, $\liminf_{i\to\infty} x_i \ge a$. From this follows the claim. □

Now we will prove Proposition 5.

Proof Suppose k is such that $(2k+1)^n > N$. Take $i > 3k+1$, then, $|\Omega_{k,i}| \geq (2k+1)^n > N$ and $|\Omega_i - \Omega_{k,i}| \geq (2k+1)^n > N$.

Then, $h(\Omega_i) = h(\Omega_{k,i} \cup (\Omega_i - \Omega_{k,i})) \geq h(\Omega_{k,i})$.

Moreover, each translate of Ω_k has cardinality $(2k+1)^n$. Therefore,

$$h(\Omega_{k,i}) \geq \frac{|\Omega_{k,i}|}{|\Omega_k|} h(\Omega_k).$$

From this follows that

$$\frac{h(\Omega_i)}{|\Omega_i|} \geq \frac{h(\Omega_{k,i})}{|\Omega_{k,i}|} \frac{|\Omega_{k,i}|}{|\Omega_i|} \geq \frac{h(\Omega_k)}{|\Omega_k|} \frac{|\Omega_{k,i}|}{|\Omega_i|},$$

and the claim is a consequence of Lemmas 13 and 14. □

The next lemma will be used later

Lemma 15 *Under the hypothesis of Proposition 5 consider*

$$\Omega_i' = (\Omega_i + (2i+1)\gamma_1) \cup \Omega_i, \quad i = 1, 2, 3, \ldots$$

Then,

$$\lim_{i\to\infty} \frac{h(\Omega_i')}{|\Omega_i'|} = \lim_{i\to\infty} \frac{h(\Omega_i)}{|\Omega_i|}.$$

Proof If $i > N$, then $|\Omega_i| > N$. Therefore,

$$h(\Omega_i') \geq h(\Omega_i + (2i+1)\gamma_1) + h(\Omega_i) = 2h(\Omega_i).$$

From this follows

$$\frac{h(\Omega_i')}{|\Omega_i'|} \geq \frac{h(\Omega_i)}{|\Omega_i|}.$$

Therefore,

$$\liminf_{i\to\infty} \frac{h(\Omega_i')}{|\Omega_i'|} \geq \liminf_{i\to\infty} \frac{h(\Omega_i)}{|\Omega_i|}.$$

We assume that $\Gamma = \underbrace{\mathbb{Z} \oplus \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}}_{n}$ and $\gamma_1, \gamma_2, .., \gamma_n$ is the canonical basis.

Take k such that $(2k+1)^n > N$. For $i > 5k+2$, take $m > 1$ such that $k + m(2k+1) \leq i \leq k + (m+1)(2k+1)$.

Consider

$$\Omega'_{k,i} = \cup \{ \Omega'_k + (j_1(2k+1), \dots, j_n(2k+1)) \mid j_1 \text{ is even}, -m \le j_1 \le m-1,$$

$$-m \le j_2, .., j_n \le m, \ (j_1, j_2, \dots, j_n) \ne (0, \dots, 0) \ \}.$$

Then, $\Omega'_{k,i} \subset \Omega_i$, and $\Omega'_{k,i}$ is a finite union of disjoints translates of Ω'_k. Moreover $\lim_{i\to\infty} \frac{|\Omega'_{k,i}|}{|\Omega_i|} = 1$,

$$| \Omega'_{k,i} | \ge 2\,(2\,k+1)^n > N \text{ and } | \Omega_i - \Omega'_{k,i} | \ge 2\,(2\,k+1)^n > N \ .$$

From this follows that

$$h(\Omega_i) = h(\Omega'_{k,i} \cup (\Omega_i - \Omega'_{k,i})) \ge h(\Omega'_{k,i}),$$

By the other hand, all translate of Ω'_k has cardinality bigger than N.
Therefore,

$$h(\Omega'_{k,i}) \ge \frac{| \Omega'_{k,i} |}{| \Omega'_k |} h(\Omega'_k).$$

Then,

$$\frac{h(\Omega_i)}{| \Omega_i |} \ge \frac{h(\Omega'_{k,i})}{| \Omega_i |} \ge \frac{1}{|\Omega_i|} \frac{| \Omega'_{k,i} | h(\Omega'_k)}{| \Omega'_k |} = \frac{| \Omega'_{k,i} |}{| \Omega_i |} \frac{h(\Omega'_k)}{| \Omega'_k |}.$$

Now, for a fixed k, taking $i \to \infty$ in the above inequality we get

$$\lim_{i\to\infty} \frac{h(\Omega_i)}{| \Omega_i |} \ge \frac{h(\Omega'_k)}{| \Omega'_k |}.$$

From this follows that

$$\lim_{i\to\infty} \frac{h(\Omega_i)}{| \Omega_i |} \ge \limsup_{k\to\infty} \frac{h(\Omega'_k)}{| \Omega'_k |}.$$

□

6 Properties of $b(c)$

Lemma 16 *There exists $c \in [0, 1]$ such that*

$$b(c) \ge \log(dim\, H^*(M, \mathbb{R})) > 0$$

Proof Note that dim $(H^*(M))\geq 2$ because dim $M\geq 1$. Let q be the number of connected components of M.

If $|\Omega_i|=m_i$, take $0=t_0<t_1<\cdots<t_{m_i}=1$, a partition of $[0,1]$ in m_i intervals of the same size. By Lemma 1

$$H^*(f_{\Omega_i}^{-1}(-\infty,t_{m_i}))=\oplus_{r>0}H^r(M^{\Omega_i},\mathbb{R}),$$

Denote A_{ij} a supplement of $H^*(f_{\Omega_i}^{-1}(-\infty,t_{j-1}))$ in $H^*(f_{\Omega_i}^{-1}(-\infty,t_j))$, $1\leq j\leq m_i$. Then,

$$\sum_{j=1}^{m_i}\dim A_{ij}=\dim H^*(f_{\Omega_i}^{-1}(-\infty,t_{m_i}))=\dim H^*(M^{\Omega_i},\mathbb{R})-q.$$

Therefore, there exists a certain $A_{i\,j_i}=A_i$, such that,

$$\dim A_i\geq\frac{(\dim H^*(M,\mathbb{R}))^{m_i}-q}{m_i}.$$

Denote s_i the middle point of $(t_{j_i-1},t_{j_i}]$ and $\delta_i=\frac{1}{2m_i}$.

Then, by definition of $b_i'(s_i,\delta_i)=\dim A_i$.

There exists a subsequence $s_{i_k}\to c\in[0,1]$, when $k\to\infty$.

Given $\delta>0$, there exists a $K>0$ such that $\delta_{i_k}<\delta/2$ and $|s_{i_k}-c|<\delta/2$, if $k>K$.

This means $c-\delta<s_{i_k}-\delta_{i_k}$ and $s_{i_k}+\delta_{i_k}<c+\delta$.

From this follows that $b'_{i_k}(c,\delta)\geq b'_{i_k}(s_{i_k},\delta_{i_k})=\dim A_{i_k}$.

Finally, we get

$$\frac{\log(b'_{i_k}(c,\delta))}{|\Omega_{i_k}|}\geq\frac{1}{m_{i_k}}\log\frac{(\dim H^*(M,\mathbb{R}))^{m_{i_k}}-q}{m_{i_k}}.$$

Now, taking limit in $k\to\infty$ in the above expression we get

$$b'(c,\delta)\geq\log(\dim(H^*(M,\mathbb{R})).$$

□

Lemma 17 *The function $b(c)$ is upper semicontinuous.*

Proof Suppose c_k, $k\in\mathbb{N}$ is a sequence of points in $[0,1]$ such that, $c_k\to c$.

Given $\varepsilon>0$, take $\delta>0$, such that, $b'(c,\delta)<b(c)+\varepsilon$. There exists a $N>0$ such that $|c-c_k|<\delta/2$, if $k\geq N$. Then, $c-\delta<c_k-\delta/2$ and $c_k+\delta/2<c+\delta$, if $k\geq N$.

Then, $b_i'(c,\delta)\geq b_i'(c_k,\delta/2)$, if $k\geq N$, for all $i=1,2,3,\ldots$

From this follows that $b'(c,\delta) \geq b'(c_k, \delta/2)$. Therefore,

$$b(c) + \varepsilon > b'(c,\delta) \geq b'(c_k, \delta/2) \geq b(c_k), \text{ if } k \geq N.$$

Therefore

$$\limsup_{k\to\infty} b(c_k) \leq b(c) + \varepsilon,$$

for any $\varepsilon > 0$. From this it follows the claim. □

Lemma 18 *The function $b(c)$ is concave.*

Proof Consider $0 \leq c_1 < c_2 \leq 1$ and $0 \leq t \leq 1$, we will show that

$$b(t\,c_1 + (1-t)\,c_2) \geq t\,b(c_1) + (1-t)\,b(c_2).$$

First we will show the claim for $t = 1/2$. Denote $\tilde{\Omega}_i = \Omega_i + (2\,i+1)\gamma_1$ and $\Omega_i' = \Omega_i \cup \tilde{\Omega}_i$.

By Proposition 2 and Lemma 12 we get:

$$b'_{\Omega_i'}(1/2\,c_1 + 1/2\,c_2,\ \delta) \geq b'_{\Omega_i}(c_1,\delta)\,b'_{\tilde{\Omega}_i}(c_2,\delta) = b'_i(c_1,\delta)\,b'_i(c_2,\delta),$$

for all $\delta > 0$.

Now, applying Lemma 15 to $h(\Omega) = \log b'_\Omega(1/2\,c_1 + 1/2\,c_2, \delta)$, we get $b'(1/2\,c_1 + 1/2\,c_2, \delta) \geq 1/2\,b'(c_1,\delta) + 1/2\,b'(c_2,\delta)$.

Now, taking $\delta \to 0$, we get $b(1/2\,c_1 + 1/2\,c_2) \geq 1/2 b(c_1) + 1/2\,b(c_2)$.

The inequality we have to prove is true for a dense set of values of t in $[0, 1]$. Then, by Lemma 17 is true for all $t \in [0, 1]$. □

Corollary 3 *The function $b(c)$ is continuous for $c \in [0,]$.*

Proof This follows from Lemmas 17 and 18. □

We summarize the above results in the following theorem.

Theorem 1 *(a)* $0 \leq b(c) \leq \varepsilon(c) \leq \log($ *number of critical points of* $f_0)$*, for all* $0 \leq c \leq 1$.

(b) $b(c)$ is continuous on $[0, 1]$

(c) $b(c)$ is concave, that is, its graph is always above the cord

(d) $b(c)$ is not constant equal zero. Moreover, there exists a point c where $b(c) \geq \log(\dim H^*(M, \mathbb{R})) > 0$.

7 An Example

The next example shows that the item (d) in the above theorem can not be improved.

Take $M = S^n$, $n \geq 1$, and a Morse function $f_0 : M \to [0, 1]$ which is surjective with only two critical points. Suppose x_- is the minimum and x_+ the maximum of f_0. We will compute $b(c)$ and $\varepsilon(c)$.

Take $\Omega \in B(\Gamma)$ with $|\Omega| = m \geq 1$. For each $\Omega' \subset \Omega$ consider the canonical projection $p_{\Omega'} : M^\Omega \to M^{\Omega'}$. Now, take

$$\mu^{\Omega'} = p^*_{\Omega'}([\, M^{\Omega'}\,]) \in H^{n\,|\,\Omega'|}(M^\Omega, \mathbb{R}),$$

where [ ] represents fundamental class. Then,

$$\{\, \mu^{\Omega'} \,:\, \Omega' \subset \Omega\}$$

is a $\mathbb{R}$-homogeneous basis of $H^*(M^\Omega, \mathbb{R})$.

For $0 \leq d \leq 1$ denote

$$L_d = \{x \in M^\Omega \,:\, f_\Omega(x) < d\,\} \subset M^\Omega.$$

For $x \in M^\Omega$ we denote by x_γ the corresponding coordinate, where $\gamma \in \Gamma$.

Lemma 19 *If $0 \leq d \leq 1$, where d is not rational, then*

$$\{\mu^{\Omega'} \,:\, |\Omega'| > m\,(1-d)\,\}$$

is a basis of $H^*(L_d)$.

Proof Take $K_d = M^\Omega - L_d$. By Lemma 1

$$H^*(L_d) = \mathrm{Ker}(\, H^*(M^\Omega, \mathbb{R}) \;\to\; H^*(K_d, \mathbb{R})\,) \text{ (natural restriction).}$$

The claim follows from

(1) $H^k(M^\Omega, \mathbb{R}) \to H^k(K_d, \mathbb{R})$ is zero if $k > m\,(1-d)\,n$, and

(2) $H^k(M^\Omega, \mathbb{R}) \to H^k(K_d, \mathbb{R})$ is injective if $k < m\,(1-d)\,n$.

Now we prove (1) and (2).

(1) Suppose $\Omega' \subset \Omega$ is such that $\mu^{\Omega'} \in H^k(M^\Omega)$ where $k > m\,(1-d)\,n$. Then, $|\Omega'| > m\,(1-d)$. Suppose

$$F_{\Omega'} = \{x \in M^\Omega \,:\, x_\gamma = x_-\,,\ \text{if}\, \gamma \in \Omega'\}.$$

If $x \in F_{\Omega'}$, then $f_\Omega(x) \leq \frac{1}{m}\,(m - |\Omega'|) < d$. Then, $F_{\Omega'} \cap K_d = \emptyset$. This means that: if $x \in K_d \;\to\; x_\gamma \neq x_-$ for some $\gamma \in \Omega'$. Then, $K_d \subset p^{-1}_{\Omega'}(M^{\Omega'} - \{z\})$ where $z_\gamma = x_-$ for all $\gamma \in \Omega'$.

From this follows

$$\mu^{\Omega'} \mid K_d \;=\; p^*_{\Omega'}\,(\,[\,M^{\Omega'}\,]\,) \mid K_d \;= 0,\;\; \text{because}\, [\,M^{\Omega'}\,] \mid (\,[\,M^{\Omega'}\,] - \{z\}) = 0.$$

(2) Denote $T = \{x \in M^{\Omega} \;:\; \text{cardinality}(\,\{\gamma \;:\; x_\gamma = x^+\}) \;>\; m\,d\,\}$. The set T is closed.

If $x \in T$, then $f_\Omega(x) > \frac{1}{m}\, m\, d = d$. Then, $T \subset K_d$.

We have to show that

$$H^k(M^{\Omega}, \mathbb{R}) \to H^k(T, \mathbb{R}) \text{ is injective if } k < m\,(1-d)\,n.$$

As we had seen before $H^k(M^{\Omega}, \mathbb{R}) = 0$ if k is not multiple of n. Then, we can assume that $k = q\,n$, if $q = 0, 1, 2, \ldots$. The claim follows from the next lemma, taking s the integer part of $m\,d$, by the exact sequence of homology, given that $U = U_s(\Omega)$.

Lemma 20 *Suppose $s = 0, 1, 2, .., m$. Suppose*

$$U_s(\Omega) \;=\; \{x \in M^{\Omega} \;:\; card\,(\,\{\gamma \;:\; x_\gamma = x^+\}\,) \;\le s\},$$

then, $H^k_c(U_s(\Omega), \mathbb{R}) = 0$, if $k < (m-s)\,n$.

Proof The claim is trivial for $s = 0$ or $s = m$ ($U_0(\Omega)$ is homeomorphic to $(\,\mathbb{R}^n\,)^m$).

The proof is by induction in m. The claim for $m = 1$ is trivial. Suppose is true for $m - 1 \ge 1$. Take $0 < s < m$. Fix $w \in \Omega$ and take $\Omega' = \Omega - \{w\}$.

Consider $\varphi : M^{\Omega'} \to M^{\Omega}$ and $\psi : M^{\Omega'} \times (M - \{\,x^+\,\}) \to M^{\Omega}$, where for a given x we define $\varphi(x)$ by $x_\omega = x^+$ if $x \in M^{\Omega'}$, and $\psi(x, u)$ is defined by $x_w = u$ if $x \in M^{\Omega'}$ and $u \in M$, $u \neq x^+$.

ψ identifies $U_s(\Omega') \times (\,M - \{x^+\}\,)$ with an open set A contained in $U_s(\Omega)$.

Moreover, φ identifies $U_{s-1}(\Omega')$ with the complement of this open set A in $U_s(\Omega)$.

As $M - \{x^+\}$ is homeomorphic to $\mathbb{R}^n$ and by recurrence we get that

$$H^k_c(\,U_s(\Omega') \;\times\; (\,M - \{x^+\}\,)\,, \mathbb{R}) = 0,$$

if $k < (m-1-s)\,n + n = (m-s)\,n$ and, moreover, $H^k_c(\,U_{s-1}(\Omega', \mathbb{R})\,) = 0$, if $k < (\,(m-1) - (s-1)\,)\,n = (m-s)\,n$.

Now, using the exact sequence of homology we finish the proof. □

Now we fix irrationals d_1, d_2, $0 < d_1 < d_2 < 1$. Denote $a_m = m\,(1-d_1)$, $b_m = m(1-d_2)$, and, $c_m = \dim\,(\,H^*(L_{d_2})/H^*(L_{d_1})\,)$.

By Lemma 19 we get

$$c_m = \sum \{ \binom{m}{j} \,:\, b_m < j < a_m\}.$$

Assume m is much bigger than $(d_2 - d_1)$.

Take an integer j_m, such that $b_m < j_m < a_m$,

$$\binom{m}{j_m} = \sup\{\binom{m}{j} : b_m < j < a_m\}.$$

Then,

$$\binom{m}{j_m} \leq c_m \leq (a_m - b_m + 1)\binom{m}{j_m}.$$

By Stirling formula:

$$\frac{1}{m}\log\binom{m}{j} \sim \frac{1}{m}\log\left(\frac{m^{m+1/2}}{j^{j+1/2}(m-j)^{m-j+1/2}}\right) =$$

$$\frac{1}{m}\log\left(m^{-1/2}\left(\frac{j}{m}\right)^{-1/2}\left(1-\frac{j}{m}\right)^{-1/2}\left(\frac{j}{m}\right)^{-j}\left(1-\frac{j}{m}\right)^{-m+j}\right).$$

Therefore,

$$\frac{1}{m}\log\binom{m}{j_m} \sim \frac{1}{m}\log\left(\left(\frac{j_m}{m}\right)^{-j_m}\left(1-\frac{j_m}{m}\right)^{-m+j_m}\right) =$$

$$-\frac{j_m}{m}\log\left(\frac{j_m}{m}\right) - \left(1-\frac{j_m}{m}\right)\log\left(1-\frac{j_m}{m}\right),$$

when $m \sim \infty$.

As $1 - d_2 < \frac{j_m}{m} < 1 - d_1$, then (changing x by $(1-x)$) we get

$$\limsup_{m\to\infty}\frac{1}{m}\log\binom{m}{j_m} \leq \sup_{d_1<x<d_2}(-x\log(x) - (1-x)\log(1-x)),$$

and

$$\liminf_{m\to\infty}\frac{1}{m}\log\binom{m}{j_m} \geq \inf_{d_1<x<d_2}(-x\log(x) - (1-x)\log(1-x)).$$

From this follows

$$\limsup_{m\to\infty}\frac{\log c_m}{m} \leq \sup_{d_1<x<d_2}(-x\log(x) - (1-x)\log(1-x)),$$

and

$$\liminf_{m\to\infty}\frac{\log c_m}{m} \geq \inf_{d_1<x<d_2}(-x\log(x) - (1-x)\log(1-x)).$$

Proposition 6

$$\varepsilon(c) = b(c) = -\,c\,\log c - \,(1-c)\,\log(1-c),\;\; 0 \le c \le 1.$$

Proof Given $0 < c < 1$, there exists small $\delta > 0$ such that

$$0 < c - \delta < c < c + \delta < 1 \;\text{ and }\; c - \delta, c + \delta \text{ are not in } \mathbb{Q}.$$

From the above for $d_1 = c - \delta$ and $d_2 = c + \delta$ we get

$$\inf_{d_1 < x < d_2} (-x\log(x) - (1-x)\,\log(1-x)\,) \le b'(c,\delta) \le$$

$$\sup_{d_1 < x < d_2} (-x\log(x) - (1-x)\,\log(1-x)\,).$$

Now, taking $\delta \to 0$, we get

$$b(c) \;=\; (-c\log(c) - (1-c)\,\log(1-c)\,).$$

For $c = 0$ or $c = 1$ the result follows from continuity.

Now we will estimate $\varepsilon(c)$.

The critical values of f_Ω are $0, \frac{1}{m}, \frac{2}{m}, \ldots, 1$.

To the critical values $\frac{j}{m}$ $(j = 0, 1, 2, \ldots, m)$ corresponds $\binom{m}{j}$ critical points.

Therefore, given $d_1, d_2 \in \mathbb{R}$ $d_1 < d_2$, the number c'_m of critical points of f_Ω in $f_\Omega^{-1}(d_1, d_2)$ is

$$c'_m = \sum \{\binom{m}{j} \,:\, d_1 < \frac{j}{m} < d_2\,\} =$$

$$\sum \{\binom{m}{j} \,:\, m\,(1-d_2) < \,j\, < m\,(1-d_1)\,\}.$$

The computation of $\varepsilon(c)$ is analogous to the one for $b(c)$. This also follows from the last Theorem and the fact that $H^*(M)$ = number critical points of f_0 in the present case. □

8 About the Definition of $b(c)$

We will show that the definition of $b(c)$ presented here coincides with the one in [1].

Suppose X is a compact connected oriented C^∞ manifold.

Lemma 21 *Given an open set V in X consider $\alpha \in H^*(X, \mathbb{R})$ such that $\alpha|_V \neq 0$. Then, there exists $\beta \in H^*(V)$ such that $\alpha \wedge \beta \neq 0$.*

Proof Take $w \in \alpha$. As $\alpha|_V \neq 0$, then there exists a cycle z on V such that $\int_z w \neq 0$.

Suppose w' is a closed form with compact support on V such that its cohomology class in $H_c^*(V, \mathbb{R})$ is the Poincare dual of the homology class of z in $H_*(V, \mathbb{R})$.

w' can be extended to a closed form on X (putting 0 where needed) and by Poincare duality:

$$0 \neq \int_z w = \int_V w \wedge w' = \int_X w \wedge w'.$$

Therefore, $w \wedge w'$ is not exact on X.

Denote $\beta \in H^*(X, \mathbb{R})$ the cohomology class of w'. By Lemma 1 we have that $\beta \in H^*(V)$. As $w \wedge w'$ is not exact we get that $\alpha \wedge \beta \neq 0$. □

Notation: if $S \subset X$, then $\mathscr{H}^*(S) = \cap\{H^*(W) : W \subset X$ is an open set and $S \subset W\}$.

Lemma 22 *Suppose $U, V \subset X$ are open sets and $X = U \cup V$. Take $K = U - V$ and $\alpha \in H^*(U)$. Then, $\alpha \wedge \beta = 0$ for all $\beta \in H^*(V)$, if and only if, $\alpha \in \mathscr{H}^*(K)$.*

Proof Suppose $\alpha \in \mathscr{H}^*(K)$ and take $\beta \in H^*(V)$. By Lemma 5 there exists $w \in \beta$ such that supp $w \subset V$.

Take $W = X -$ supp w (which contains K). By definition we get that $\alpha \in H^*(W)$. Then, by Lemma 5, there exists $w' \in \alpha$ such that supp $w' \subset W$. Therefore, $w \wedge w' = 0$, and finally it follows that $\alpha \wedge \beta = 0$.

Reciprocally, suppose that $\alpha \wedge \beta = 0$ for all $\beta \in H^*(V)$. By Lemma 21 we have that $\alpha|V = 0$. Take $W \supset K$, then $V \cup W = X$. Therefore, by definition $\alpha \in H^*(W)$. □

Lemma 23 *Take $K \subset X$ a compact submanifold with boundary such that $K - \delta K$ is an open subset of X.*

Then,

$$\mathscr{H}(K) = Ker\ (\ H^*(X, \mathbb{R}) \to H^*(X - K, \mathbb{R})\ )\ \textit{restriction.}$$

WE leave the rest of the proof for the reader.

Proof Take W an open set by adding a necklace to K. Then, $X - K$ can be retracted by deformation over $X - W$.

Then, if $\alpha \in H^*(X, \mathbb{R})$, we get that $\alpha|_{X-K} = 0$ is equivalent to $\alpha|_{X-W} = 0$.

Now, the claim follows from Lemma 1 and by the definition of $\mathscr{H}(K)$. □

Corollary 4 *Under the same hypothesis of last lemma it also follows that $\mathscr{H}(K) = H^*(\ int\ (K)\ )$.*

Proof This follows from the fact that $H^*(X - \text{int}\,(K)\,, \mathbb{R}) \to H^*(X - K, \mathbb{R})$ is an isomorphism. □

Proposition 7 *Suppose U, V are open sets such that $X = U \cup V$ and moreover that $\overline{U}, \overline{V}$ are submanifolds with boundary of X.*

Consider the linear transformation L such that

$$L : H^*(U) \to Hom\,(\,H^*(V),\, H^*(U \cap V)\,),$$

where, $a \to (\,b \to a \wedge b\,)$.

Then, the rank of L is dim $(\,H^*(U)/H^*(M - \overline{V})\,)$.

Proof By Lemma 22 we get that Ker $L = H^*(X - V)$. Finally, by the last corollary $H^*(X - V) = H^*(M - \overline{V})$. □

Consider now a Morse function $f : X \to \mathbb{R}$ and $c \in \mathbb{R}, \delta > 0$.

Definition 5 $b_{c,\delta}(f)$ is the rank of the linear transformation

$$H^*(\,f^{-1}(-\infty, c + \delta)\,) \to \text{Hom}\,(\,H^*(f^{-1}(\,c - \delta, \infty\,),\, H^*(f^{-1}(c - \delta, c + \delta)\,)\,),$$

where $a \to (b \to a \wedge b)$.

Note that $b_{c,\delta}(f)$ decreases with δ.

Lemma 24 *If $c - \delta$ and $c + \delta$ are regular values of f, then*

$$b_{c,\delta}(f) = b'_{c,\delta}(f).$$

Proof Just apply Proposition 7 to $U = f^{-1}(-\infty, c + \delta)$ and $V = f^{-1}(\,c - \delta, \infty\,)$. □

Note that $b_\Omega(c, \delta) = b_{c,\delta}(f_\Omega)$, where $\Omega \in B(\Gamma)$ and $\Omega \neq \emptyset$, and moreover that $b_i(c, \delta) = b_{\Omega_i}(c, \delta)$. The next limit exists (see [1]).

Definition 6

$$b(c, \delta) = \lim_{i \to \infty} \frac{\log(\,b_i(c, \delta)\,)}{|\,\Omega_i\,|}.$$

The set $S \subset [0, 1]$ of all critical values of all f_Ω is countable. By Lemma 24 we get that $b'_i(c, \delta) = b_i(c, \delta)$ if $c - \delta \notin S$ and $c + \delta \notin S$. Therefore, $b'(c, \delta) = b(c, \delta)$ if $c - \delta \notin S$ and $c + \delta \notin S$.

Finally,

$$\lim_{\delta \to 0} b'(c, \delta) = \lim_{\delta \to 0} b(c, \delta)$$

because both limits exist.

Therefore the function $b(c)$ we define coincides with the one presented in [1].

References

1. Bertelson, M., Gromov, M.: Dynamical Morse Entropy, Modern Dynamical Systems and Applications, pp. 27–44. Cambridge University Press, Cambridge (2004)
2. Milnor, J.: Morse Theory. Princeton University Press, Princeton (1963)
3. Baraviera, A.T., Cioletti, L., Lopes, A.O., Mohr, J., Souza, R.R.: On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. Rev. Math. Phys. **23**(10), 1063–1113, 82Bxx (2011)
4. Chou, W., Griffiths, R.: Ground states of one-dimensional systems using effective potentials. Phys. Rev. B **34**(9), 6219–6234 (1986)
5. Coronel, D., Rivera-Letelier, J.: Sensitive dependence of Gibbs measures. J. Stat. Phys. **160**, 1658–1683 (2015)
6. Fukui, Y., Horiguchi, M.: One-dimensional chiral XY model at finite temperature. Interdiscip. Inf. Sci. **1**(2), 133–149 (1995)
7. Lopes, A.O., Mohr, J., Souza, R.R., Thieullen, P.: Negative Entropy, Zero temperature and stationary Markov chains on the interval. Bull. Soc. Bras. Math. **40**(1), 1–52 (2009)
8. Lopes, A.O., Mengue, J.K., Mohr, J., Souza, R.R.: Entropy and variational Principle for one-dimensional lattice systems with a general a-priori probability: positive and zero temperature. Ergod. Theory Dyn. Syst. **35**(6), 1925–1961 (2015)
9. Thompson, C.: Infinite-spin ising model in one dimension. J. Math. Phys. **9**(2), 241–245 (1968)
10. van Enter, A.C.D., Ruszel, W.M.: Chaotic temperature dependence at zero temperature. J. Stat. Phys. **127**(3), 567–573 (2007)
11. Cioletti, L., Lopes, A.: Interactions, Specifications, DLR probabilities and the Ruelle operator in the one-dimensional lattice. Discret. Contin. Dyn. Syst.-Ser. A **37**(12), 6139–6152 (2017)
12. Cioletti, L., Lopes, A.: Phase Transitions in one-dimensional translation invariant systems: a Ruelle operator approach. J. Stat. Phys. **159**(6), 1424–1455 (2015)
13. Sarig, O.: Lecture notes on thermodynamic formalism for topological Markov shifts. Penn State (2009)
14. Baraviera, A., Leplaideur, R., Lopes, A.O.: Ergodic Optimization, Zero temperature limits and the Max-Plus Algebra, mini-course in XXIX Colóquio Brasileiro de Matemática - IMPA - Rio de Janeiro (2013)
15. Asaoka, M., Fukaya, T., Mitsui, K., Tsukamoto, M.: Growth of critical points in one-dimensional lattice systems. J. d'Anal. Math. **127**, 47–68 (2015)
16. Massey, W.: Homology and Cohomology, M. Dekker (1978)

Synchronisation of Weakly Coupled Oscillators

Rogério Martins

Abstract The synchronization phenomenon was reported for the first time by Christiaan Huygens, when he noticed the strange tendency of a couple of clocks to synchronise their movements. More recently this phenomena was shown to be ubiquitous in nature and it is broadly studied by its applications, for example in biological cycles. We consider the problem of synchronization of a general network of linearly coupled oscillators, not necessarily identical. In this case the existence of a linear synchronization space is not expected, so we present an approach based on the proof of the existence of a synchronization manifold, the so-called generalised synchronization. Based on some results developed by R. Smith and on Wazewski's principle, a general theory on the existence of invariant manifolds that attract the solutions of the system that are bounded in the future, is presented. Applications and estimates on parameters for the existence of synchronization are presented for several examples: systems of coupled pendulum type equations, coupled Lorenz systems of equations, and oscillators coupled through a medium, among many others.

Keywords Synchronisation · Coupled oscillators · Invariant manifolds · Dissipative systems

1 Introduction

Christiaan Huygens was a central figure in the creation and development of mechanical clocks, he developed what we still know today as pendulum clocks. At the beginning of his work, clocks had measurement errors in the order of 15 min per day,

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UIDB/00297/2020 (Centro de Matemática e Aplicações).

R. Martins (✉)
Departamento de Matemática, FCT, UNL, Centro de Matemática e Aplicações (CMA), FCT, UNL, Lisbon, Portugal
e-mail: roma@fct.unl.pt

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_14

at the end of his life that error was reduced to about 1 min per day. One of Huygens' developments was the introduction of a restriction, in the form of a cycloid curve, at the base of the pendulum, in order to make it isochronous [2].

Interestingly, Huygens' great motivation was not to create wall clocks for everyday use. Nowadays we cannot imagine life without clocks, in order to know when to enter the next meeting or what time do we have to pick up kids from school. Apparently, in the seventeenth century, this was not a serious need in people's lives. One of the central problems at the time was the question of determining the longitude at the sea, the solution of which was believed to be in the creation of sufficiently accurate clocks that could be used on board. That was the problem that Huygens wanted to solve. In fact, this issue was so central to navigation that there were several cash prizes for those who manage to solve this problem, no matter the method. An excellent reference for this scientific achievement is the book [20], which tells the story of Huygens, but mainly of John Harrison, who dedicated 45 years of his life to the problem.

In 1665, Huygens was recovering from an illness when he noticed something curious: the two clocks he had at home, on top of a bookshelf, were oscillating synchronously. After doing some experiments he found that it was not a coincidence, no matter the position in which the pendula start the movement, after about half an hour they start to oscillate in unison. After doing several experiments in what he initially called "an unusual kind of sympathy", he realised that the phenomenon was due to a small oscillation on the shelf, which created the coupling between the two clocks.

Today we call this phenomenon *synchronization of coupled oscillators*. It is a well-known case of serendipity, something that was discovered by chance. These clocks were built in pairs in order to be used in the open ocean, in this way, even if one of them was malfunctioning, or some type of maintenance was needed, the other clock could keep the time. Note that, at sea, if the longitude is not known, there is no way to set a clock. So, it was fortunate that there were two identical clocks side by side on the same shelf.

Interestingly, Huygens was never rewarded for his contributions to the problem of longitude. One of the arguments against his clocks was precisely the fact that they could not be reliable as they were so easily driven by the clock on the side. However, he was the first person to report something that today is known to be present in so many different scenarios. Today it is known that certain species of fireflies tend to flash synchronously. The circadian cycle synchronises with the solar cycle, causing the jet lag in the absence of synchronization. Our neurons are subject to several synchronization phenomena. The financial markets exhibit synchronization phenomenon. Among many other cases.

Nowadays, the phenomenon of synchronization of oscillators is an extremely active field of research, largely due to its applications. The body of living organisms is full of biological cycles and nature seems to have used this phenomenon to make these systems more stable. If the function of a small organ depends on a synchronisation process, for example the synchronisation of several cells, this makes the functioning of that organ more stable, since an individual failure or imprecision of some of

the cells does not compromise the functioning of the organ as a whole. For more examples of synchronisation phenomena see [22].

The concept of synchronization is very broad and encompasses many situations and systems of a very different nature, hence the definition of synchronization is usually given on a case-by-case basis. We say that there is synchronization when some type of adjustment of rhythms is observed. Another way to look at synchronization is to think about the possibility of making predictions about the state of one of the oscillators knowing the asymptotic behaviour of the other. Depending on the type of adjustment, several types of synchronization can be considered: complete/identical synchronisation, phase synchronisation, generalised synchronization, gap synchronisation, just to name a few. In the following, we will consider two or more weekly coupled oscillators and mainly consider the identical and the generalised case.

The best way to understand what we are talking about is to observe a concrete example. Consider a system of two coupled pendula

$$\begin{cases} x_1'' + \gamma x_1' + \sin(x_1) = f(t) + L(x_2 - x_1) \\ x_2'' + \gamma x_2' + \sin(x_2) = f(t) + L(x_1 - x_2) \end{cases},$$

where γ and L are two positive parameters. Each of the variables x_1 and x_2 measures the angular position of each pendulum. The terms $\gamma x_1'$ and $\gamma x_2'$ are a very rough representation of friction, with γ representing the magnitude of that friction. On the right hand side of the equation we have a continuous $f(t)$ that represents an external force, something that feeds the system with energy. In a very crude way we can consider that it represents the pendulum clock's spring. Finally, the second terms on the right hand side create a coupling, L represents the magnitude of this coupling. If $L = 0$ we would have two decoupled mathematical pendula equations with friction, in which case the solution of each equation is independent of the other. For $L \neq 0$ this system models the movement of two pendula coupled by a spring. This type of coupling is quite simple, is far from modelling the Huygens' case, yet it is interesting as a first example.

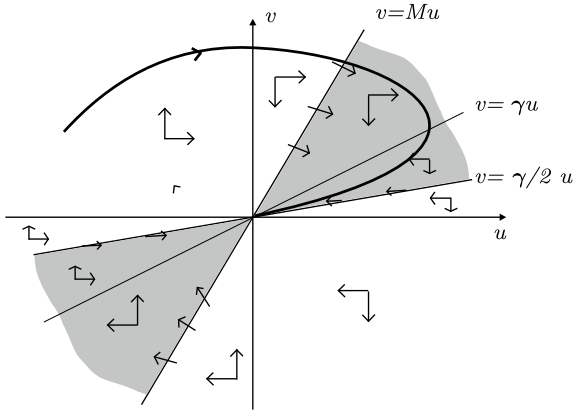
We can turn this into a first order system in the usual way

$$\begin{cases} x_1' = v_1 - \gamma x_1 \\ v_1' = -\sin(x_1) + f(t) + L(x_2 - x_1) \\ x_2' = v_2 - \gamma x_2 \\ v_2' = -\sin(x_2) + f(t) + L(x_1 - x_2) \end{cases}. \tag{1}$$

The first thing we notice is that the subspace $S = \{x_1 = x_2, v_1 = v_2\}$ is invariant for the solutions of this equation. By existence and uniqueness of solution, an orbit with initial conditions in S stays in S for all $t \in \mathbb{R}$. As we will see, under certain conditions, this subspace attracts all the orbits of the system, in this case we say that the two oscillators synchronise and that S is a synchronisation manifold.

We will now show in which situations this system synchronises. Given a solution $(x_1, v_1, x_2, v_2)^T$ of (1) we consider $(u, v)^T = (x_1 - x_2, v_1 - v_2)^T$ that is a solution of

Fig. 1 Phase space of (2)



$$\begin{cases} u' = v - \gamma u \\ v' = -\sin(x_1(t)) + \sin(x_2(t)) - 2Lu \end{cases}. \tag{2}$$

If we now consider $\xi = v/u$, for $u \neq 0$, that verifies

$$\xi' = -\frac{\sin(x_1(t)) - \sin(x_2(t))}{x_1(t) - x_2(t)} - 2L - \xi^2 + \gamma\xi,$$

which is essentially a Ricatti equation. Since the first term on the right side of the equation is bounded,

$$-1 < \frac{\sin(x_1(t)) - \sin(x_2(t))}{x_1(t) - x_2(t)} < 1,$$

we note that if $\xi = M$, with M large enough, we have $\xi' < 0$. On the other hand, if $\xi = \gamma/2$ then

$$\xi' = -\frac{\sin(x_1(t)) - \sin(x_2(t))}{x_1(t) - x_2(t)} - 2L + \frac{\gamma^2}{4} > -1 - 2L + \frac{\gamma^2}{4} > 0, \tag{3}$$

if

$$\frac{\gamma^2}{4} > 1 + 2L.$$

This shows that, if the previous equation is verified, the shaded area in the next graph is positively invariant.

On the other hand, $u' > 0$ above the line $v = \gamma u$ and $u' < 0$ below. Finally, from the second equation in (2), if $L > 1/2$, we have

$$v' = -\left(\frac{\sin(x_1(t)) - sin(x_2(t))}{x_1(t) - x_2(t)} + 2L\right)u,$$

which shows that v' and u have opposite sign. We then obtain the flow suggested by the arrows in Fig. 1.

We conclude that if

$$\frac{\gamma^2}{4} > 1 + 2L > 2,$$

the solutions will enter the shaded area and converge to the origin, as in the example of the orbit shown in the figure. That is, $x_1 - x_2$ and $v_1 - v_2$ both tend to zero and the solution $(x_1, v_1, x_2, v_2)^T$ converges to S. Asymptotically we have $x_1 = x_2$ and $v_1 = v_2$.

The last equation suggests that L must be large enough for the system to synchronise, which is normal, given that this parameter measures the strength of the coupling. In addition, the coefficient of friction, γ, also needs to be made large enough, which is also expected, a stronger dissipation is more likely to create attractors that "fit" in a subspace of smaller dimension.

The situation shown in the previous example is what we can call identical or complete synchronization. This type of synchronization appears more typically in systems where the two oscillators are identical, as in the previous case (see [3, 16, 24]). However, in practical applications, it is common to have different oscillators. Consider for example the case of the synchronization of the circadian cycle with the solar cycle, in which we have two oscillators with a completely different nature. On the other hand, we often have similar but not identical oscillators and it is important to take into account these differences. An example would be the case of the previous system if, for example, the friction coefficient or the natural frequency were distinct in the two pendula. In these cases, the symmetry is lost and the problem becomes considerably more complicated. In particular, it is not normal to expect to have a linear invariant subspace, as in the previous case. Therefore we will look for a generic synchronization manifold S that will assume the role taken by the synchronization subspace of the previous example. Although it is the kind of synchronization that is often not identifiable by the "naked eye", technically the orbits converge to a lower dimension manifold. Asymptotically, the behaviour of one oscillator depends of the behaviour of the other. This is the so-called generalised synchronization. In Sect. 3 we will develop a theory on the existence of invariant manifolds based on the work of R. Smith and the Wazewsky principle, which will be applied to several examples.

In fact, one of the simplest examples of synchronization, in which the two oscillators are definitely different, is the situation in which we have a system driven by a periodic force. In this case, we can try to determine when this force leads the system to, asymptotically, have a periodic behaviour. This is the theme of Sect. 2. In Sect. 4 the case of several coupled oscillators will be studied and in Sect. 5 the case of a generic equation of order n periodically driven.

Recently, it was discovered that it is possible to synchronise chaotic oscillators, which brought a wave of possible applications to information encryption schemes. In Sect. 7 we will see the case where two Lorenz oscillators synchronise.

In many applications, oscillators are not directly connected. Imagine a situation in which there are several cells interacting with a common medium where they are

immersed. In this case, each oscillator interacts directly with its environment and the other oscillators are perturbed through this environment. This is, actually, the case of Huygens, each clocks interfere with a base, which has its own dynamics. This special case is treated in Sect. 8.

2 Forced Oscillations

Unlike the example presented in the introduction, the coupling does not have to be bidirectional. We can consider a situation where we have two oscillators in which one of them follows its own rhythm without interference. One of the simplest situations occurs when we have an oscillator driven by a periodic force, for example:

$$x'' + h(x)x' + g(x) = f(t), \tag{4}$$

where $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$ are continuous functions, which verify the following periodicity conditions $h(x + 1) = h(x)$ and $g(x) = g(x + 1)$, for all $x \in \mathbb{R}$, and $f(t) = f(t + T)$ for some positive constant T. We can denote this fact by $h, g \in C(\mathbb{R}/\mathbb{Z})$ and $f \in C(\mathbb{R}/T\mathbb{Z})$. Suppose h is strictly positive, that is, $0 < \gamma = \min_{\mathbb{R}} h(x)$. We will also assume that the above equation has existence and uniqueness for each set of initial conditions.

Although we are writing rather general equation, we can think in the model case $g(x) = \sin(x)$, where we obtain the equation of a forced mathematical pendulum.

Considering, $H(x) = \int_0^x h(s)ds$, we see that x is a solution of (4) if and only if $(x, x' + H(x))$ is a solution of

$$\begin{cases} y_1' = y_2 - H(y_1) \\ y_2' = -g(y_1) + f(t) \end{cases}. \tag{5}$$

The above system has the form $y' = F(y, t)$ with $y = (y_1, y_2)^T$. Writing $h = \tilde{h} + \hat{h}$, where $\hat{h} = \int_0^1 h(s)ds$ is constant and $\int_0^1 \tilde{h}(s)ds = 0$ we get

$$H(x + 1) = \int_0^x h(s)ds + \int_x^{x+1} [\tilde{h}(s) + \hat{h}]ds = H(x) + \hat{h}.$$

So, considering $R = (1, \hat{h})^T$ we obtain $F(y + R, t) = F(y, t)$, for all $(y, t) \in \mathbb{R}^3$, that is, the natural phase space for the system (5) is the quotient space $\mathscr{C} = \mathbb{R}^2/R\mathbb{Z}$. In practice, we are identifying each solution y with the solution $y + kR$, $k \in \mathbb{N}$. In this quotient space, which is a cylinder, we will denote by $\overline{y}$ the class of $y \in \mathbb{R}^2$.

Since we are in the presence of a periodic non-autonomous system, a natural thing to do is to consider the associated Poincaré map, also known as stroboscopic map, as it only records the state of the system for $t = kT$, $k \in \mathbb{Z}$. If $y(t; t_0, y_0)$

is the solution of (5) that verifies $y(t_0) = y_0$ then we consider the Poincaré map $P(y) = y(T; 0, y)$. Since $y(t; 0, y_0 + R) = y(t; 0, y_0) + R$, we can consider the corresponding Poincaré map in the cylinder $\mathscr{C}$ defined by

$$\overline{P} : \mathscr{C} \to \mathscr{C}$$

$$\overline{y} \to \overline{y(T; 0, y)}.$$

which is a homeomorphism.

In [10] we proved that the system (5) is dissipative, that is, there is a compact and non-empty set, called a window, $\overline{B} \subset \mathscr{C}$ such that $\overline{P}(\overline{B}) \subset \text{int}\overline{B}$. We then obtain

$$\ldots \overline{P}^n(\overline{B}) \subset \overline{P}^{n-1}(\overline{B}) \subset \ldots \subset \overline{P}^2(\overline{B}) \subset \overline{P}(\overline{B}) \subset \overline{B}.$$

We can consider the intersection of all these compacts

$$\mathscr{A} = \bigcap_{n=0}^{\infty} \overline{P}^n(\overline{B}).$$

This type of construction is classic, and it is not difficult to prove that $\mathscr{A}$ is non-empty, compact, invariant for the Poincaré map and does not depend on $\overline{B}$. In addition, $\mathscr{A}$ projects over $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$. Finally, if g is continuously differentiable then $\mathscr{A}$ has measure zero (see [10]).

The set $\mathscr{A}$ is called an attractor for the Poincaré map in $\mathscr{C}$. In fact, it can be proved that for each $\overline{x} \in \mathscr{C}$, $d(\overline{P}^n(\overline{x}), \mathscr{A}) \to 0$, when $n \to \infty$ and convergence is uniform across each compact set $S \subset \mathscr{C}$ (see [10]).

The set $\mathscr{A}$ can be quite complex, this is what typically happens for small friction coefficients. However, as we will see, if γ is large enough then $\mathscr{A}$ is a curve homeomorphic to the circle. In this case, the dynamics become much simpler, all the orbits given by iterates of the Poincaré map converge to an invariant curve homeomorphic to a circle. The dynamics then can be described by the theory of the homeomorphisms of the circle, in particular we can consider an associated rotation number, which essentially determines the dynamics in the limit set. If the rotation number is integer or rational, we have situations where the orbits converge to a periodic orbit of period T or orbits whose period is a multiple of T. In some sense, we can say that the system synchronises with the periodic external force. We have the following:

Theorem 1 *If there is a constant γ_1 such that*

$$\gamma_1 < \frac{g(x) - g(y)}{x - y} < \frac{\gamma^2}{4}, \tag{6}$$

for all $(t, x, y) \in \mathbb{R}^3$, with $x \neq y$, then $\mathscr{A}$ is homeomorphic to $\mathbb{T}^1$.

The proof of this theorem is essentially technical and can be seen in [10]. It involves the fact that $\mathscr{A}$ is formed by the initial conditions of solutions that are bounded in the cylinder $\mathscr{C}$ along with an analysis of the phase space of (5).

Similar estimates were obtained in [8, 14] for particular cases of equations of the type of (4), but with stronger regularity conditions on g.

It can also be shown that the previous theorem is optimal, in the sense that the second inequality in (6) cannot be improved. More specifically, we have the following theorem whose proof can be seen in [9].

Theorem 2 *Given $\mathscr{H} > \gamma^2/4$, there is $g \in C^\infty(\mathbb{R}/\mathbb{Z})$ with $g' < \mathscr{H}$ and there is $T \in \mathbb{R}$ and $f \in C(\mathbb{R}/T\mathbb{Z})$ such that attractor associated with the Eq. (5) (with $h \equiv \gamma$) is not homeomorphic to $\mathbb{T}^1$.*

Behind this result there is a topological concept: the idea of an inversely unstable solution. We say that y, a solution of (5), is (a, b)-periodic, for $a, b \in \mathbb{Z}$, $b \geq 1$, iff $y(t + bT) = y(t) + aR$, for all $t \in \mathbb{R}$. Intuitively, an (a, b)-periodic solution is a bT-periodic solution in $\mathscr{C}$ that goes around the cylinder a times in each period.

If y is a periodic (a, b)-solution of (5), then $y(0)$ is a fixed point of

$$P^b - aR : \mathbb{R}^2 \to \mathbb{R}^2,$$

where P is the Poincaré map in the plane. Assuming that $y(0)$ is an isolated fixed point, we can define the index of y as

$$\gamma_b(y) = \deg(I - [P^b - aR], B),$$

where deg designates the Brouwer degree and $B \subset \mathbb{R}^n$ is a sufficiently small disk so that $y(0)$ is the only fixed point of $P^b - aR$ in B.

If y is an (a, b)-periodic solution of (5) then this same solution is also $(2a, 2b)$-periodic. We will say that y is inversely unstable if $y(0)$ is an isolated fixed point of $P^b - aR$ and $P^{2b} - 2aR$ and if

$$\gamma_b(y) = 1 \quad \text{and} \quad \gamma_{2b}(y) = -1.$$

The presence of inversely unstable periodic (a, b)-solutions and the flow characteristics associated with (4) requires the existence of a periodic $(2a, 2b)$-solution that is not (a, b)-periodic which implies that the rotation number is not well defined and therefore $\mathscr{A}$ cannot topologically be a circle. The complete proof of the next theorem can be seen in [9].

Theorem 3 *Suppose for some $(a, b) \in \mathbb{Z} \times \mathbb{N}$, $b \geq 1$, the set of (a, b)-periodic solutions of (4) is finite and given by*

$$y_1, \ y_2, \ldots, y_p$$

(where we are assuming that y_i and $y_i + kR$, $k \in \mathbb{Z}$, is the same solution). If there is an inversely unstable (a, b)-periodic solution then $\mathscr{A}$ is not homeomorphic to $\mathbb{T}^1$.

Later on, the conditions of Theorem 1 were shown to be optimal using a class of equations studied by F. Tricomi in the 1930s (see [11]). It is essentially a class of pendulum-type oscillators forced by a constant torque. Considering a type of nonlinearity formed by a sectionally linear function, it was possible to find an alternative proof that the second inequality in Theorem 1 cannot be improved.

3 Invariant Manifolds

As we have seen, the study of synchronization depends on the existence of invariant manifolds, which somehow attract the system's orbits. This was the case of the previous section. However, in the previous section we were essentially working in the plane, which allows the use of geometric and topological techniques that are not available in higher dimensions. In the next sections we will study systems of several coupled oscillators and we would like to prove the existence of synchronization manifolds, with this idea in mind we need a theory that is applicable in a higher dimension. In this section we will present some general results of invariant manifolds inspired by the work of R. Smith while we was studying periodic solutions of systems of differential equations (see [18, 19]). The proof of these results was presented in [12] and the application to several scenarios can be seen in [10, 12, 13].

In general we will work with a system of the form

$$x' = F(x,t) + Cx \tag{7}$$

where $F : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}^p$ is T-periodic and continuous function in t and locally Lipchitz in x. On the other hand, C is a real matrix. In general, in applications, F will contain the nonlinear part of each oscillator and C will include the coupling.

Let's start by briefly summarise the Wazewski topological principle [21, 23]. Given an equation of the type of (7) we will denote by $(\alpha(t_0, x_0), \omega(t_0, x_0)) \subset \mathbb{R}$ the maximal interval of definition of $x(t; t_0, x_0)$, the solution that verifies the initial condition $x(t_0) = x_0$. Given an open set $\Omega \subset \mathbb{R}^p \times \mathbb{R}$, let's say that the point $(x_0, t_0) \in \partial\Omega$ is an ingress point if there is a $\varepsilon > 0$ such that $(x(t; t_0, x_0), t) \in \Omega$ for each $t \in (t_0, t_0 + \varepsilon)$. If in addition $(t, x(t; t_0, x_0)) \notin \bar{\Omega}$ for each $t \in (t_0 - \varepsilon, t_0)$ and some $\varepsilon > 0$, we say that (x_0, t_0) is a strict ingress point. We will denote by Ω_i and Ω_{si}, the set of ingress points and the set of strict ingress points, respectively (see Fig. 2). Clearly $\Omega_{si} \subset \Omega_i \subset \partial\Omega$. Finally, if X is a topological space and $A \subset X$ is a subspace, we say that A is a retract of X if there is a continuous function $r : X \to A$ such that $r(x) = x$ for each $x \in A$. In that case, we say that r is a retraction.

We are now able to present the topological principle behind the results of this section. The Fig. 2 illustrate a typical example of the associated setup.

Theorem 4 (Wazewski's principle) *Let's assume that $\Omega_i = \Omega_{si}$ and the existence of a set $S \subset \Omega \cup \Omega_i$ such that $S \cap \Omega_i$ is a retraction of Ω_i but $S \cap \Omega_i$ is not a retraction*

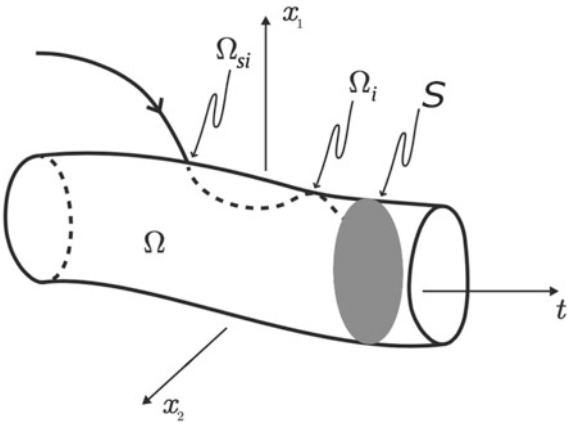


Fig. 2 Example of Wazewski topological configuration

of S. In this case, there is a point $(x_0, t_0) \in S \cap \Omega$ *such that* $(x(t; t_0, x_0), t) \in \Omega$ *for all* $t \in (\alpha(t_0, x_0), t_0]$.

In what follows we will outline the proof of the existence of an invariant manifold for (7), the details of the proof can be seen in [12]. The next hypothesis will be central.

(H) There are constants $\lambda > 0$, $\varepsilon > 0$ and a symmetric real matrix P with precisely n negative eigenvalues, such what $(x - y)^T P[F(x, t) - F(y, t) + (C + \lambda I)(x - y)] \leq -\varepsilon \|x - y\|^2$, for all $x, y \in \mathbb{R}^p$ and $t \in \mathbb{R}$.

If $V(x) := x^T Px$, then it is easy to see that (H) is equivalent to

$$\frac{d}{dt}\left\{e^{2\lambda t} V(x(t) - y(t))\right\} \leq -e^{2\lambda t}\varepsilon ||x(t) - y(t)||^2 \tag{8}$$

for each pair $x(t), y(t)$ of solutions of (7) and for each $t \in \mathbb{R}$. In particular, the function $t \to e^{2\lambda t} V(x(t) - y(t))$ is strictly decreasing. This means that if we consider the cone

$$B := \{x \in \mathbb{R}^p : V(x) < 0\},$$

then the set

$$x(t) + B,$$

(which depends on t) attracts in the future all other orbits of (7). In other words, (H) can be seen as a hypothesis of dissipation. In particular we note that if $x(t)$ and $y(t)$ are solutions of (7) such that $V(x(t_0) - y(t_0)) = 0$ for some $t_0 \in \mathbb{R}$, then $V(x(t) - y(t)) < 0$, $\forall t \in (t_0, +\infty)$ and $V(x(t) - y(t)) > 0$, $\forall t \in (-\infty, t_0)$. From

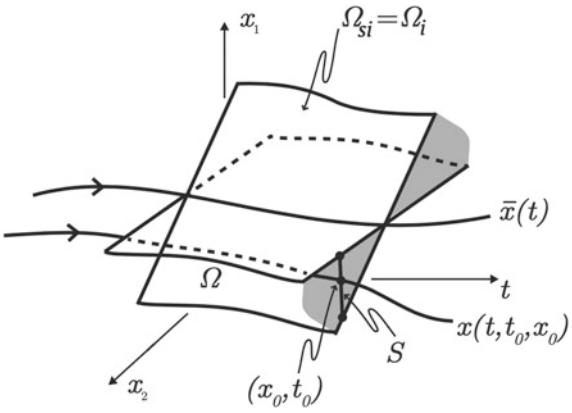


Fig. 3 Dynamics in the extended phase space, given (H)

the point of view of $x(t) + B$ this means that if $y(t_0) \in x(t_0) + \partial B$ then $y(t) \in x(t) + B,\ \forall t \in (t_0, +\infty)$ e $y(t) \notin x(t) + B, \forall t \in (-\infty, t_0)$ (see Fig. 3).

Let us now consider a special class of solutions. We will say that a solution $x(t)$ of (7) is amenable if the integral

$$\int_{-\infty}^{t_0} e^{2\lambda t} \|x(t)\|^2 dt \tag{9}$$

converges. For each $t \in \mathbb{R}$ we consider the amenable set

$$\mathscr{A}_t = \{x(t) \,:\, x(\cdot) \text{ is a amenable solution of } (7)\}.$$

Let's suppose as a hypothesis that we have an amenable solution $\bar{x}(t)$, it was proved in [19] that another solution $y(t)$ de (7) is also amenable if $V(\bar{x}(t) - y(t)) < 0$, for all $t \in \mathbb{R}$. Which implies, from the geometric point of view that for each $t \in \mathbb{R}$

$$\mathscr{A}_t \setminus \{\bar{x}(t)\} \subset \bar{x}(t) + B.$$

Let us now consider the subspaces V_- and V_+ of $\mathbb{R}^p$ generated respectively by the eigenvectors of P corresponding to the negative and positive eigenvalues. These subspaces of dimension n and $p - n$, are orthogonal and complementary, that is $\mathbb{R}^p = V_- \perp V_+$. Let $\mathscr{P}_-$ an orthogonal projection of $\mathbb{R}^p$ on V_-.

We are finally in a position to state the main theorem of this section.

Theorem 5 *If the system (7) verifies the hypothesis (H) and there is at least one amenable solution $\bar{x}(t)$, then for each $t \in \mathbb{R}$, $\mathscr{P}_-$ is a homeomorphism between $\mathscr{A}_t$ and V_-. Moreover, $\mathscr{A}_t$ is a graph of a globally Lipschitz function. In particular the amenable set $\mathscr{A}_t$ is a manifold of dimension n. Finally, the bounded solutions*

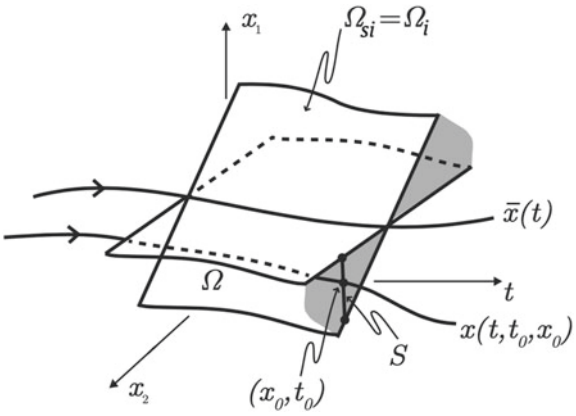


Fig. 4 Theorem proof scheme Theorem 5

in the future converge to $\mathscr{A}_t$*, that is, given a bounded* $x(t)$ *solution in the future,* $dist(\mathscr{A}_t, x(t)) \to 0$ *when* $t \to +\infty$.

The basis of the proof is to consider $\Omega := \{(\bar{x}(t) + B, t),\ t \in \mathbb{R}\}$, in this case $\Omega_i = \Omega_{si} = \partial\{\bar{x}(t) + B, t \in \mathbb{R}\}\backslash\{\bar{x}(t),\ t \in \mathbb{R}\}$. For each $\xi \in V_-$ we will consider $\Omega_r = \bar{x}(t_0) + B$ and

$$S := \mathscr{P}_-^{-1}\xi \cap \bar{\Omega}_r$$

(see Fig. 4). It can be proved rigorously what is somehow intuitive in a geometric way, that $S \cap \Omega_i$ is a retract of Ω_i but $S \cap \Omega_i$ is not a retract of S. So, by the Wazewski, principle and assuming that the solutions are defined up to $-\infty$, there is a point $(x_0, t_0) \in S \cap \{(\bar{x}(t) + B, t),\ t \in \mathbb{R}\}$ such that $(x(t; t_0, x_0), t) \in \{\bar{x}(t) + B\}$ for all $t \in (-\infty, t_0]$. That is, we have $\mathscr{P}_-(x_0) = \xi$ and $x(t, t_0, x_0)$ is amenable, so $\mathscr{P}_-$ is onto. The convergence for $\mathscr{A}_t$, is a standard argument. For a complete proof see [12].

We will end this section with a criterion for equation (7) to verify the condition (H). Let's suppose that there is a $\lambda \geq 0$ such that C has no eigenvalues with real part equal to $-\lambda$ and in such a way that C has precisely n eigenvalues with strictly real part greater than $-\lambda$. In this case $C + \lambda I$ will have precisely n eigenvalues with positive real part and the Lyapunov equation

$$(C + \lambda I)^T P + P(C + \lambda I) = -I \tag{10}$$

has a single P solution if and only if (see [5])

$$\sigma(C + \lambda I) \cap \overline{\sigma(-C - \lambda I)} = \emptyset. \tag{11}$$

Since the eigenvalues are in finite number, we can easily choose λ such that (11) is verified, let P be the solution of the Lyapunov equation for this λ. From (10) we get

$$(C + \lambda I)^T P^T + P^T(C + \lambda I) = -I^T = -I, \tag{12}$$

and from the uniqueness of the solution to this equation we conclude that P is symmetrical. Finally the General Inertia Theorem (see [5]) shows that P has n negative eigenvalues and $p - n$ positive.

The next theorem shows that under certain conditions (7) verifies (H) with this matrix P.

Theorem 6 *Given λ in the above conditions and P the solution to the Lyapunov equation (10), if there is an $\varepsilon > 0$ such that*

$$(x - y)^T P[F(x, t) - F(y, t)] \leq (1/2 - \varepsilon)\|x - y\|^2, \tag{13}$$

then the Eq. (7) verifies (H) for that λ, ε and P.

For the proof just observe that

$$(x - y)^T P[F(x, t) - F(y, t) + (C + \lambda I)(x - y)]$$

$$= \frac{1}{2}(x, y)^T[(C + \lambda I)^T P + P(C + \lambda I)](x, y) + (x, y)^T P[F(x, t) - F(y, t)]$$

$$\leq -\varepsilon\|x - y\|^2.$$

Often in applications the non-linearity F is globally K-Lipschitz in x, that is, there is a constant $K > 0$ such that

$$\|F(x_1, t) - F(x_2, t)\| \leq K\|x_1 - x_2\|,$$

for each $x_1,\ x_2 \in \mathbb{R}^p$ and $t \in \mathbb{R}$. We finally observe that the inequality in the previous theorem is obviously verified if F is globally K-Lipschitz and

$$K < \frac{1}{2\|P\|}.$$

4 Synchronization of Several Coupled Forced Oscillators

In this section we will present results similar to those in Sect. 2 but for a system of coupled differential equations. This systems also generalize the example given in the introduction. We consider a system in $\mathbb{R}^{p/2}$, for an even $p \geq 4$, as follows

$$u'' + \gamma u' + Au + g(u) = f(t), \tag{14}$$

where $u \in \mathbb{R}^{p/2}$, γ is a positive constant and A is a matrix in $M_{p/2\times p/2}(\mathbb{R})$, symmetric, with an eigenvalue $\alpha_1 = 0$ and all other eigenvalues positive $\alpha_2 \leq \ldots \leq \alpha_{p/2}$ (writ-

ten in ascending order and according to its multiplicity). Let's assume that $\eta \in \mathbb{R}^{p/2}$ is such that $\mathrm{Ker} A = \mathrm{span}\{\eta\}$. For simplicity, let's consider $\|\eta\| = 1$. In addition, $g : \mathbb{R}^{p/2} \to \mathbb{R}^{p/2}$ is locally Lipchitz and periodic in the direction of η, more precisely $g(u + \eta) = g(u)$ for all $u \in \mathbb{R}^{p/2}$. Finally $f : \mathbb{R} \to \mathbb{R}^{p/2}$ is continuous and T-periodic.

One of the most natural concretisations for this system of equations is the case

$$g(u_1, u_2, \ldots, u_{p/2}) = (\, sin(u_1),\ sin(u_2), \ldots,\ sin(u_{p/2}))^T$$

with the matrix A

$$A = - \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -2 & 1 \\ 0 & \ldots & 0 & 1 & -1 \end{pmatrix}.$$

In this special case, the eigenvalues of A are:

$$4\, sin^2 \left(\frac{(j-1)\ pi}{p} \right),\ j = 1,\ \ldots,\ p/2.$$

This case is a model of a set of pendula arranged in a line, each of them connected by a spring to the pendula that are beside it. It is also the type of equation that arises naturally when the sine-Gordon partial differential equation is discretised. Furthermore, it is also a model for a set of Josephson junctions (see [4]). Although this concretisation is very natural, presenting the results for a generic equation, of the type (14), does not complicate the presentation at all, in fact it even makes it simpler, which is why we chose the more general formulation.

If we consider $u' = v$, the system (14) can be rewritten as a first order system of the type of (7),

$$x' = F(x, t) + Cx \tag{15}$$

where $x = (u, v)^T \in \mathbb{R}^p$,

$$C = \begin{pmatrix} 0 & I \\ -A & -\gamma I \end{pmatrix},\quad F(x, t) = \begin{pmatrix} 0 \\ -g(u) + f(t) \end{pmatrix}.$$

Note that for $R = (\eta, 0)^T \in \mathbb{R}^p$ we have $F(y + R, t) = F(y, t)$ for all $(y, t) \in \mathbb{R}^p \times \mathbb{R}$. That is, as in Sect. 2, the natural phase space for this system is a cylinder.

Since A is a symmetric matrix, we can take an orthonormal basis $\{\eta, \eta_2, \eta_3, \ldots, \eta_{p/2}\}$ of eigenvectors with associated eigenvalues $\alpha_1 = 0, \alpha_2, \alpha_3, \ldots, \alpha_{p/2}$ respectively. The matrix whose columns are made up of

the base vectors above

$$P_1 = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \eta & \eta_2 & \dots & \eta_{p/2} \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

is orthogonal, that is $P_1^{-1} = P_1^T$, and $P_1^T A P_1 = \text{diag}(0, \alpha_2, \dots, \alpha_{p/2})$. Considering

$$P_2 = \begin{pmatrix} P_1^T & 0 \\ 0 & P_1^T \end{pmatrix}$$

we get

$$P_2 C P_2^{-1} = \begin{pmatrix} 0 & I \\ -\text{diag}(0, \alpha_2, \dots, \alpha_n) & -\gamma I \end{pmatrix}.$$

On the other hand, the matrix P_3 corresponding to the linear map

$$P_3 : \mathbb{R}^p \to \mathbb{R}^p$$

$$(u_1, u_2, \dots, u_p, v_1, v_2, \dots, v_p)^T \to (u_1, v_1, u_2, v_2, \dots, u_p, v_p)^T$$

it is such that

$$P_3 P_2 C P_2^{-1} P_3^{-1} = \begin{pmatrix} 0 & 1 & & & \dots & 0 \\ 0 & -\gamma & & & & \\ & & 0 & 1 & & \\ & & -\alpha_2 & -\gamma & & \\ \vdots & & & & \ddots & \vdots \\ & & & & 0 & 1 \\ 0 & & \dots & & -\alpha_p & -\gamma \end{pmatrix}.$$

For each $i = 1, \dots, p/2$ consider the block

$$A_i = \begin{pmatrix} 0 & 1 \\ -\alpha_i & -\gamma \end{pmatrix}.$$

Let's now assume that $\gamma^2 - 4\alpha_i > 0$, for all $i = 1, \dots, p/2$. In this case the matrix A_i has two real eigenvalues, $-\frac{\gamma}{2} + \frac{\sqrt{\gamma^2 - 4\alpha_i}}{2}$ e $-\frac{\gamma}{2} - \frac{\sqrt{\gamma^2 - 4\alpha_i}}{2}$ associated to the eigenvectors

$$\begin{pmatrix} \frac{1}{\sqrt{2}\sqrt{\gamma^2 - 4\alpha_i}} \\ -\frac{\gamma}{2\sqrt{2}\sqrt{\gamma^2 - 4\alpha_i}} + \frac{1}{2\sqrt{2}} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{1}{\sqrt{2}\sqrt{\gamma^2 - 4\alpha_i}} \\ -\frac{\gamma}{2\sqrt{2}\sqrt{\gamma^2 - 4\alpha_i}} - \frac{1}{2\sqrt{2}} \end{pmatrix}$$

respectively (these vectors were chosen in order to facilitate the computations later on). We conclude that the matrix

$$M_i = \begin{pmatrix} \frac{1}{\sqrt{2}\sqrt{\gamma^2-4\alpha_i}} & -\frac{1}{\sqrt{2}\sqrt{\gamma^2-4\alpha_i}} \\ -\frac{\gamma}{2\sqrt{2}\sqrt{\gamma^2-4\alpha_i}} + \frac{1}{2\sqrt{2}} & -\frac{\gamma}{2\sqrt{2}\sqrt{\gamma^2-4\alpha_i}} - \frac{1}{2\sqrt{2}} \end{pmatrix},$$

with inverse

$$M_i^{-1} = \begin{pmatrix} \frac{\gamma}{\sqrt{2}} + \frac{\sqrt{\gamma^2-4\alpha_i}}{\sqrt{2}} & \sqrt{2} \\ -\frac{\gamma}{\sqrt{2}} + \frac{\sqrt{\gamma^2-4\alpha_i}}{\sqrt{2}} & -\sqrt{2} \end{pmatrix},$$

is such that

$$M_i^{-1} A_i M_i = \operatorname{diag}\left(-\frac{\gamma}{2} + \frac{\sqrt{\gamma^2-4\alpha_i}}{2},\ -\frac{\gamma}{2} - \frac{\sqrt{\gamma^2-4\alpha_i}}{2} \right).$$

We can then consider the matrix

$$P_4 = \begin{pmatrix} M_1^{-1} & & 0 \\ & \ddots & \\ 0 & & M_{p/2}^{-1} \end{pmatrix}$$

such that $P_4 P_3 P_2 C P_2^{-1} P_3^{-1} P_4^{-1} = D$ is a diagonal matrix, with the eigenvalues of C in the diagonal. In the next two lemmas, we will try to estimate the Lipschitz constant of

$$G(y,t) = P_4 P_3 P_2 F(P_2^{-1} P_3^{-1} P_4^{-1} y, t)$$

in y.

Lemma 1 *If $y = (0, v)^T \in \mathbb{R}^p$ then $\|P_4 P_3 P_2 y\| = 2\|v\|$.*

Proof Given $y = (0, v)^T \in \mathbb{R}^p$, we have

$$\begin{aligned} \|P_4 P_3 P_2 y\| &= \|P_4(0, \eta^T v, 0, \eta_2^T v, 0, \ldots, 0, \eta_n^T v)\| \\ &= \|(\sqrt{2}\eta^T v, -\sqrt{2}\eta^T v, \sqrt{2}\eta_2^T v, \ldots, -\sqrt{2}\eta_s^T v\| \\ &= 2\|(\eta^T v, \eta_2^T v, \ldots, \eta_n^T v)\| = 2\|P_1^T v\| = 2\|v\|. \end{aligned}$$

□

Lemma 2 *For all $x = (u, v)^T \in \mathbb{R}^p$ we have*

$$\|P_4 P_3 P_2 x\| \geq \sqrt{\gamma^2 - 4\alpha_{p/2}}\|u\|.$$

Proof If $x = (u, v)^T \in \mathbb{R}^p$ then

$$\|P_4P_3P_2x\| = \|P_4(\eta^T u, \eta^T v, \eta_2^T u, \eta_2^T v, \ldots, \eta_{P/2}^T u, \eta_{P/2}^T v)\| =$$

$$\left\| \left(\left(\frac{\gamma}{\sqrt{2}} + \frac{\sqrt{\gamma^2 - 4\alpha_1}}{\sqrt{2}} \right) \eta^T u + \sqrt{2}\eta^T v, \left(-\frac{\gamma}{\sqrt{2}} + \frac{\sqrt{\gamma^2 - 4\alpha_1}}{\sqrt{2}} \right) \eta^T u - \sqrt{2}\eta^T v, \right. \right.$$

$$\left. \left. \ldots, \left(-\frac{\gamma}{\sqrt{2}} + \frac{\sqrt{\gamma^2 - 4\alpha_{P/2}}}{\sqrt{2}} \right) \eta_{P/2}^T u - \sqrt{2}\eta_{P/2}^T v \right) \right\|$$

$$= \sqrt{2\left(\frac{\sqrt{\gamma^2 - 4\alpha_1}}{\sqrt{2}} \eta^T u \right)^2 + 4\left(\frac{\gamma}{2}\eta^T u + \eta^T v \right)^2 + \ldots + 4\left(\frac{\gamma}{2}\eta_{P/2}^T u + \eta_{P/2}^T v \right)^2}$$

$$\geq \sqrt{(\sqrt{\gamma^2 - 4\alpha_1}\eta^T u)^2 + \ldots + (\sqrt{\gamma^2 - 4\alpha_{p/2}}\eta_{p/2}^T u)^2}$$

$$\geq \sqrt{\gamma^2 - 4\alpha_{p/2}}\|P_1^T u\| = \sqrt{\gamma^2 - 4\alpha_{p/2}}\|u\|.$$

Remember that $\alpha_{p/2}$ is the largest eigenvalue of A. □

We are then in a position to state the main result of this section:

Theorem 7 *Suppose $\gamma^2 - 4\alpha_i > 0$, for all $i = 0, \ldots, p/2$. We will also assume that there is at least one amenable solution for the corresponding first order Eq. (15). If g is K-Lipschitzian and*

$$K < \frac{1}{8}\left(\gamma - \sqrt{\gamma^2 - 4\alpha_2} \right) \sqrt{\gamma^2 - 4\alpha_{p/2}} \tag{16}$$

then the Poincaré map associated with this equation has an attractor homeomorphic to $\mathbb{T}^1$. Moreover, this manifold can be seen as a graph of a function with domain in the subspace spanned by $(\eta, 0)$.

Proof The change of variables $y = P_4P_3P_2x$ turns the Eq. (15) into

$$y' = G(y, t) + Dy.$$

On the other hand, for each $y_1 = P_4P_3P_2(u_1, v_1)^T$, $y_2 = P_4P_3P_2(u_2, v_2)^T$ and $t \in \mathbb{R}$, we get from the previous two lemmas

$$\sup_{y_1 \neq y_2} \frac{\|G(y_1,t) - G(y_2,t)\|}{\|y_1 - y_2\|} = \sup_{y_1 \neq y_2} \frac{\|P_4 P_3 P_2 (0, -g(u_1) + g(u_2))^T\|}{\|y_1 - y_2\|}$$

$$= \sup_{y_1 \neq y_2} \frac{2\|g(u_1) - g(u_2)\|}{\|y_1 - y_2\|} \leq \sup_{y_1 \neq y_2} \frac{2K\|u_1 - u_2\|}{\|y_1 - y_2\|}$$

$$\leq \frac{2K}{\sqrt{\gamma^2 - 4\alpha_{p/2}}}$$

We now consider $\lambda \in]0, \frac{\gamma}{2} - \frac{\sqrt{\gamma^2 - 4\alpha_2}}{2}[$. Note that $-(\frac{\gamma}{2} - \frac{\sqrt{\gamma^2 - 4\alpha_2}}{2})$ is the largest of the non-null eigenvalues of C. Since D is diagonal, the Eq. (10), for this λ, can be easily solved, obtaining

$$P = \operatorname{diag}\left(\frac{1}{-2\lambda + \gamma \pm \sqrt{\gamma^2 - 4\alpha_i}},\ i = 1, \ldots, p/2 \right).$$

So for $\lambda = \frac{\gamma}{4} - \frac{\sqrt{\gamma^2 - 4\alpha_2}}{4}$ we get

$$\|P\| = \frac{2}{\gamma - \sqrt{\gamma^2 - 4\alpha_2}}.$$

Finally, by the remark following Theorem 6, the equation verifies (H) if

$$\frac{2K}{\sqrt{\gamma^2 - 4\alpha_{p/2}}} < \frac{1}{2\|P\|},$$

which is equivalent to (16). By Theorem 5, we conclude that the system has an invariant manifold of dimension one. It is not difficult to see, reverting the change of variable, that the invariant manifold projects over the subspace generated by $R = (\eta, 0)^T$. We conclude, by the periodicity of the equation in the direction of R, that if we consider this system in the cylinder, then this manifold is homeomorphic to $\mathbb{T}^1$. □

5 An Equation of Order N

Something similar to what was done in the last sections can also be done with an ordinary equation of order n, periodically forced. So let's consider a n'th order equation with the form

$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_2 x'' + a_1 x' + g(x, x', \ldots, x^{(n-1)}) = f(t) \qquad (17)$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is locally Lipchitz and $f : \mathbb{R} \to \mathbb{R}$ is a continuous function. Let's also assume that

$$g(x, x', \ldots, x^{(n-1)}) = g(x, x', \ldots, x^{(n-1)} + (1, 0, \ldots, 0))$$

and that f is T-periodic for some $T > 0$. A function x is a solution of (17) if and only if $y = (x, x', x'', \ldots, x^{(n-1)})^T \in \mathbb{R}^n$ is a solution of the system

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \vdots \\ y_{n-1}' = y_n \\ y_n' = -a_{n-1}y_n - \ldots - a_1 y_2 - g(y) + f(t) \end{cases},$$

that is, solution of

$$y' = Cy + J(y, t), \tag{18}$$

where

$$C = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix} \quad \text{and} \quad J(y, t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ -g(y) + f(t) \end{pmatrix}$$

The C matrix has an eigenvalue $\lambda_1 = 0$ associated to the eigenvector $(1, 0, \ldots, 0)^T$. It is not difficult to prove by induction that the characteristic polynomial of C is

$$(-1)^n(x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x).$$

Let's suppose that all the roots of this polynomial are real, negative and with multiplicity one. That is, we will additionally assume that the remaining eigenvalues of C are distinct, real and negative. We will list them in order $\lambda_n < \lambda_{n-1} < \ldots < \lambda_3 < \lambda_2 < 0$. Each of these eigenvalues λ_i has its own eigenvector of the form $(1, \lambda_i, \lambda_i^2, \ldots, \lambda_i^{n-1})^T$. Therefore, the matrix

$$B = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \lambda_2 & \lambda_3 & \cdots & \lambda_n \\ 0 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \lambda_2^{n-1} & \lambda_3^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

is a diagonalizing matrix, that is,

$$B^{-1}CB = \text{diag}(0, \lambda_2, \ldots, \lambda_n) = D.$$

The matrix B is well known in polynomial interpolation problems. In fact, if $y, z \in \mathbb{R}^n$, the equation $B^T y = z$ is equivalent to finding the coefficients $y_1, y_2, \ldots, y_n$ of a polynomial L of degree $n-1$ that verifies

$$L(0) = z_1, \; L(\lambda_2) = z_2, \ldots, \; L(\lambda_n) = z_n. \tag{19}$$

L is the so-called Lagrange interpolator polynomial. Obviously, the uniqueness of this polynomial depends of the determinant of B, which is well known (see [7], p. 221) to be

$$\prod_{\substack{i,j=1 \\ i>j}}^{n} (\lambda_i - \lambda_j).$$

In this case the determinant of B is different from 0 and therefore the polynomial L is well defined. We can now follow the numerical analysis texts and, for each $i = 1, \ldots, n$, consider the polynomial

$$L_i(t) = \frac{(t-\lambda_1)(t-\lambda_2)\ldots(t-\lambda_{i-1})(t-\lambda_{i+1})\ldots(t-\lambda_n)}{(\lambda_i-\lambda_1)(\lambda_i-\lambda_2)\ldots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_n)}$$

$$= \frac{\prod_{\substack{j=1 \\ j\neq i}}^{n}(t-\lambda_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{n}(\lambda_i-\lambda_j)}.$$

This polynomial has degree $n-1$ and

$$L_i(\lambda_j) = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ if } j \neq i \end{cases},$$

so

$$L(t) = z_1 L_1(t) + z_2 L_2(t) + \ldots + z_n L_n(t)$$

is a polynomial of degree $n-1$ and verifies (19). As there is only one polynomial in these conditions, this is the Lagrange interpolation polynomial.

We will now compute the Lipschitz constant of $G(y,t) = B^{-1}J(By,t)$. Consider the vector

$$\Omega = \left(\frac{1}{\prod_{\substack{j=1 \\ j\neq 1}}^{n}(\lambda_1-\lambda_j)}, \frac{1}{\prod_{\substack{j=1 \\ j\neq 2}}^{n}(\lambda_2-\lambda_j)}, \ldots, \frac{1}{\prod_{\substack{j=1 \\ j\neq n}}^{n}(\lambda_n-\lambda_j)} \right)^T.$$

Lemma 3 *Given* $y = (0, 0, \ldots, 0, y_n)^T \in \mathbb{R}^n$, *we have* $B^{-1}y = y_n\Omega$.

Proof If $y = (0, 0, \ldots, 0, y_n)^T \in \mathbb{R}^n$, then

$$B^{-1}y = y_n\Omega \quad \text{iff} \quad (B^{-1}y)^T z = (y_n\Omega)^T z, \ \forall z \in \mathbb{R}^n.$$

But

$$(B^{-1}y)^T z = y^T (B^{-1})^T z = y^T (B^T)^{-1} z$$

is the product of y_n by the coefficient of t^{n-1} in the polynomial $L(t)$ and this coefficient (given the form of L) is

$$\sum_{i=1}^{n} \frac{z_i}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (\lambda_i - \lambda_j)}.$$

So

$$(B^{-1}y)^T z = y_n \sum_{i=1}^{n} \frac{z_i}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (\lambda_i - \lambda_j)} = y_n \Omega^T z = (y_n\Omega)^T z.$$

□

We can finally state a sufficient condition for the existence of an attractor, that will be a manifold of dimension one.

Theorem 8 *Suppose the Eq.(18) has at least one amenable solution. If* g *is* K*-Lipschitz and*

$$K < -\frac{\lambda_2}{2\|\Omega\|\|B\|},$$

then the Poincaré map associated with the system (18) has an attractor $\mathscr{A}$ *homeomorph to* $\mathbb{T}^1$.

Proof Since B diagonalizes C, we consider the change of variables $y = Bx$ that turns the Eq. (18) into

$$y' = G(y, t) + Dy.$$

Where $G(y, t) = B^{-1}J(By, t)$. Let's start by showing that the Lipschitz constant of $G(y, t) = B^{-1}J(By, t)$ in the second variable is less than $K\|\Omega\|\|B\|$. Given $z' = Bz, \ y' = By \in \mathbb{R}^n$, from the last lemma we have

$$\|G(z, t) - G(y, t)\| = \|B^{-1}(0, 0, \ldots, 0, -g(z') + g(y'))^T\|$$

$$= \|\Omega\||g(z') - g(y')| \leq K\|\Omega\|\|z' - y'\|$$

$$= K\|\Omega\|\|BB^{-1}(z' - y')\| \leq K\|\Omega\|\|B\|\|z - y\|.$$

On the other hand, given $\lambda \in]0, -\lambda_2[$, the solution of the equation (12), for this λ is

$$P = \text{diag}\left(\frac{-1}{2(\lambda + \lambda_i)},\ i = 1, \ldots, n\right).$$

So for $\lambda = -\lambda_2/2$ we get $\|P\| = -1/\lambda_2$. Finally, by the observation that follows the Theorem 6, the Eq. (18) verifies (H) if

$$K\|\Omega\|\|B\| < \frac{1}{2\|P\|}.$$

By Theorem 5 we conclude that there is an invariant manifold of dimension one that attracts the orbits of the system. This manifold can be seen as a graph with domain in the subspace generated by $(1, 0, \ldots, 0)^T$. Therefore, due to the periodicity of the equation (18) in this direction, this manifold is homeomorphic to $\mathbb{T}^1$. □

A concrete example of application of the results of this section would be to consider the system

$$\begin{cases} x'' + c_1 x' + \ sin(x) = p(t) \\ y'' + c_2 y' + y = x \end{cases},$$

which can seen as an harmonic damped oscillator forced by the movement of a pendulum, that is driven by a periodic function $p(t)$. In this case, y verifies the fourth order equation

$$y^{(4)} + (c_1 + c_2)y''' + (1 + c_1 c_2)y'' + c_1 y' = -\ sin(y'' + c_2 y' + y) + p(t),$$

which is of the type of (17). If $c_2 > 2,\ c_1 > 0$ then the roots of

$$\lambda^3 + (c_1 + c_2)\lambda^2 + (1 + c_1 c_2)\lambda + c_1$$

are $-c_1$ and $\frac{-c_2 \pm \sqrt{c_2^2 - 4}}{2}$ (real and negative), so we are in the conditions of the last theorem.

6 Two n Dimensional Coupled Systems

Let's consider a system of two n dimension systems coupled in a bidirectional way

$$\begin{cases} x_1' = f_1(x_1, t) + L(x_2 - x_1) \\ x_2' = f_2(x_2, t) + L(x_1 - x_2) \end{cases}, \tag{20}$$

where $L > 0$ is a parameter, called the coupling coefficient, and measures the coupling force. Let's assume that f_1 and f_2 are locally Lipchitz in the first variable and continuous T-periodic in t for some positive T.

In the example we saw in the introduction, the oscillators were identical, which allows us to consider a synchronization manifold that is the diagonal subspace. This type of synchronization is called identical synchronization. In this section the goal is to consider different oscillators, $f_1 \neq f_2$, even if the coupling is symmetric it is not expected to have an attractor contained in a subspace. We are going to prove the existence of invariant manifolds of inferior dimension that attract the orbits of the system, the so-called generalised synchronization. Anyway, as an exercise we start by considering the case where the oscillators are equal, $f_1 = f_2 := f$, later on we can compare the results with the general case. In this particular case the problem comes down to finding a suitable Lyapunov function. Note that the previous system can be written in the form of (7) with

$$C = \begin{pmatrix} -LI & LI \\ LI & -LI \end{pmatrix}.$$

Clearly, the diagonal $S = \{x_1 = x_2\}$ is invariant for the solutions of this system. Given a $(x_1, x_2)^T$ solution, we consider $u = x_1 - x_2$ which solves

$$u' = f(x_1, t) - f(x_2, t) - 2\,Lu.$$

Assuming that f is globally K-Lipschitz and $K < 2L$, then $E(u) = \|u\|^2$ is a Lyapunov function for the last equation. In fact, the derivative over a solution verifies

$$\dot{E}(u) = 2uu' = 2u(f(x_1, t) - f(x_2, t)) - 4\,L\|u\|^2 \leq 2(K - 2\,L)\|u\|^2 < 0.$$

We conclude that $\|u\| = \|x_1(t) - x_2(t)\| \to 0$ when $t \to +\infty$. In other words, we have identical synchronisation.

Let's now consider the case where f_1 and f_2 are not necessarily identical. Since the eigenvalues of C are 0 and $-2\,L$, each with multiplicity n, we will choose $\lambda \in]0, 2L[$. We can now solve the equation (12) by blocks and get

$$P = \begin{pmatrix} -\frac{L-\lambda}{2(2\,L-\lambda)\lambda} I & -\frac{L}{2(2\,L-\lambda)\lambda} I \\ -\frac{L}{2(2\,L-\lambda)\lambda} & -\frac{L-\lambda}{2(2\,L-\lambda)\lambda} I \end{pmatrix}.$$

Since the eigenvalues of P are $\frac{1}{2(2\,L-\lambda)}$ and $-\frac{1}{2\lambda}$, we have $\|P\| = max\{\frac{1}{2(2L-\lambda)}, \frac{1}{2\lambda}\}$.

In view of the observations following the Theorem 6, (H) is verified if $F = (f_1, f_2)$ is K-Lipschitz in x and

$$K < \max_{\lambda \in]0,2\,L[} \frac{1}{2\|P\|} = \max_{\lambda \in]0,2\,L[} \min\{2\,L - \lambda, \lambda\} = L.$$

We then obtain the following result as a consequence of Theorem 5:

Theorem 9 *If $K < L$ and there is at least one amenable solution, then the system (20) has a synchronization manifold of dimension n, which can be seen as a graph with domain in the subspace generated by the eigenvectors associated to the negative eigenvalues of P.*

Note that the condition for the existence of generalised synchronization is stronger than the corresponding condition for obtaining identical synchronization. Not surprisingly, it is intuitively more difficult to have synchronization when the oscillators are not equal. In [12] there is a proof that the manifold given by the last theorem can be seen as a graph over the subspace spanned by $(x_1, 0)^T$ or the subspace spanned by $(0, x_2)^T$.

7 Synchronization of Two Chaotic Oscillators

In this section we will see an example of two chaotic oscillators that synchronise. This example is also useful to illustrate what can be done when we have a non-linearity that is not globally Lipschitz. We will consider two chaotic oscillators, more specifically two Lorenz systems. We will also consider a unidirectional coupling. Thus, choosing parameters in an interval in which we have chaotic behaviour and leaving one of the systems free, we guarantee that when the coupled system synchronises it follows a chaotic orbit.

More precisely, let's consider the system

$$\begin{cases} x_1' = \sigma_1(y_1 - x_1) \\ y_1' = -y_1 - x_1 z_1 + \rho_1 x_1 \\ z_1' = -\beta_1 z_1 + x_1 y_1 \\ x_2' = \sigma_2(y_2 - x_2) + L(x_1 - x_2) \\ y_2' = -y_2 - x_2 z_2 + \rho_2 x_2 + L(y_1 - y_2) \\ z_2' = -\beta_2 z_2 + x_2 y_2 + L(z_1 - z_2) \end{cases},$$

where $\sigma_1,\ \sigma_2,\ \rho_1,\ \rho_2,\ \beta_1,\ \beta_2$ are positive parameters of the Lorenz system and L is the coupling parameter. As the origin is an amenable solution, according to Theorem 5 the system synchronises if (H) is verified. We will also see that this system is dissipative, that is, there will be a positively invariant set.

Similar to the last example, let's consider

$$C = \begin{pmatrix} 0 & 0 \\ LI & -LI \end{pmatrix},$$

with eigenvalues $0, -L$ and choose $\lambda = L/4$. With these values the Eq. (12) can be solved, obtaining

$$P = \frac{2}{3L}\begin{pmatrix} -11I & 2I \\ 2I & I \end{pmatrix}.$$

This matrix has eigenvalues $\frac{2(-5\pm 2\sqrt{10})}{3L}$, and so $\|P\| = \frac{2(5+2\sqrt{10})}{3L}$.

In this case F is not globally Lipschitz. However, we can prove that there is a compact set so that all orbits enter and do not leave it, this way we can truncate F out of this compact and apply the results of Sect. 3 to the truncated system.

Since the first three variables are decoupled from the remaining three, let's consider the Lyapunov function

$$E_1(x_1, y_1, z_1) = x_1^2 + y_1^2 + (z_1 - \sigma_1 - \rho_1)^2.$$

The derivative along a solution of the first three equations is

$$\dot{E}_1 = -2\left(\sigma_1 x_1^2 + y_1^2 + \beta_1\left(z_1 - \frac{\sigma_1 + \rho_1}{2}\right)^2 - \beta_1\frac{(\sigma_1 + \rho_1)^2}{4}\right),$$

so we conclude that there is a compact set (dependent on $\sigma_1,\ \rho_1,\ \beta_1$) that attracts the solutions and is positively invariant for $x_1,\ y_1,\ z_1$.

Now let's consider the last three equations as a system forced by (x_1, y_1, z_1) and consider a second Lyapunov function

$$E_2(x_2, y_2, z_2) = x_2^2 + y_2^2 + (z_2 - \sigma_2 - \rho_2)^2.$$

The derivative along the solutions is

$$\begin{aligned}\dot{E}_2 = -2\Bigg[&\left(\sqrt{\sigma_2 + L}\, x_2 - \frac{Lx_1}{2\sqrt{\sigma_2 + L}}\right)^2 - \frac{L^2x_1^2}{4(\sigma_2 + L)} + \left(\sqrt{1 + L}\, y_2 - \frac{Ly_1}{2\sqrt{1 + L}}\right)^2 \\ &- \frac{L^2y_1^2}{4(1 + L)} + \left(\sqrt{\beta_2 + L}\, z_2 - \frac{(\sigma_2 + \rho_2)(\beta_2 + L) + Lz_1}{2\sqrt{\beta_2 + L}}\right)^2 \\ &- \left(\frac{(\sigma_2 + \rho_2)(\beta_2 + L) + Lz_1}{2\sqrt{\beta_2 + L}}\right)^2 + L(\sigma_2 + \rho_2)z_1\Bigg],\end{aligned}$$

in particular,

$$\begin{aligned}\dot{E}_2 < -2L\Bigg[&\left(x_2 - \frac{L}{2(\sigma_2 + L)}x_1\right)^2 - \frac{L}{4(\sigma_2 + L)}x_1^2 + \left(y_2 - \frac{L}{2(1 + L)}y_1\right)^2 \\ &- \frac{L}{4(1 + L)}y_1^2 + \left(z_2 - \frac{\sigma_2 + \rho_2}{2} - \frac{L}{2(\beta_2 + L)}z_1\right)^2\end{aligned}$$

$$-\left(\frac{\sigma_2+\rho_2}{2}\sqrt{\frac{\beta_2+L}{L}}+\frac{1}{2}\sqrt{\frac{L}{\beta_2+L}}z_1\right)^2+(\sigma_2+\rho_2)z_1\Bigg].$$

We conclude that there is a convex set B in $\mathbb{R}^6$, that contains the origin, is positively invariant, and attracts the orbits of the system. This set depends on $\sigma_1,\ \rho_1,\ \beta_1,\ \sigma_2,\ \rho_2,\ \beta_2$. We note however that given a positively invariant set for $L = L_0$, then the same set it is also positively invariant for each $L > L_0$.

If $K = \sup_{x\in\tilde{B}} \|D_x F\|$, then F is K-Lipschitz in x in B. We now consider the truncated function

$$\tilde{F}(x,t) = \begin{cases} F(x,t), \text{ if } x \in B \\ F(g(x),t), \text{ if } x \notin B \end{cases},$$

where $g(x)$ is the only point in the boundary of B and in the line connecting the origin to x. Clearly that $\tilde{F}$ is also K -Lipschitz in x in all $\mathbb{R}^6$.

Now using the observation that follows the Theorem 6, we conclude that we have a synchronization manifold for the equation $x' = \tilde{F}(x,t) + Dx$ if

$$K \leq \frac{1}{2\|P\|} = \frac{3\,L}{4(5+2\sqrt{10})}.$$

Each orbit of the original system enters and never leaves B, on the other hand, within B coincides with one of the solutions of the truncated equation, so converges to the synchronization manifold. We conclude that we have generalised synchronization whenever the last inequality is verified for the original equation.

8 Oscillators Coupled by Medium

In the previous section, we considered situations where the oscillators are somehow directly connected, even though the coupling could be unidirectional or bidirectional. In this section we will consider the situation in which we have several oscillators, all connected to a medium that will have its own dynamics. This situation is very natural, imagine for example the case of a group of cells immersed in a common environment, each of them exchanges a certain chemical substance with the environment where it is immersed and in this way interacts indirectly with all the others. In the examples we have seen, typically each oscillator interacts with its neighbour oscillators, in this case, with a coupled through a medium, each oscillator interfere with all the others simultaneously (see [6, 15] for more examples).

Let's start with a very simple example, a linear model, yet with an interesting interpretation. Imagine two reservoirs with a chemical substance. Each of these reservoirs is connected by a semi-permeable membrane to a third common reservoir. If the concentration of this substance is measured by the variables x_1, x_2 and y respectively, then the evolution of the concentrations can be described by the following linear

model

$$\begin{cases} x_1' = L\,(y - x_1) \\ x_2' = L\,(y - x_2) \\ y' = L\,(x_1 - y) + L\,(x_2 - y)\,, \end{cases} \tag{21}$$

where L is a constant that depends on the permeability of the membrane.

This system can also be written in the matrix form

$$\begin{pmatrix} x_1' \\ x_2' \\ y' \end{pmatrix} = LA \begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix}$$

where

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

The matrix A has eigenvalues -3, -1 and 0 with corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

In this way the structure of the phase space is very clear, there is a stable central manifold which is the subspace generated by $(1, 1, 1)$. Asymptotically, the orbits converge to a state where $x_1 = x_2 = y$, we can say that the system synchronises, in fact it is a case of identical synchronization.

Now let's see what happens when we perturb this model, for that we consider the following system

$$\begin{cases} x_1' = L\,(y - x_1) + f_1(x_1, t) \\ x_2' = L\,(y - x_2) + f_2(x_2, t) \\ y' = L\,(x_1 - y) + L\,(x_2 - y) + h(y, t). \end{cases} \tag{22}$$

Let's assume that f_1, f_2 and h are locally Lipchitz in the first variable and continuous and T-periodic in t.

Let's start by looking at what happens when $f_1 = f_2 = f$, that is, when the perturbation is the same in each oscillator. This is the case where the oscillators are equal. This case is relatively simple because we are able to write a Lyapunov function. Let us start by noting that in this case the subspace generated by $(1, 1, 1)^T$ is no longer invariant, however it is contained in an invariant two-dimensional subspace: the subspace $\{x_1 = x_2\}$. In addition, we can find conditions for this subspace to attract all the other solutions.

Let $z = x_1 - x_2$, if $x_1 \neq x_2$ then

$$\begin{aligned}\dot{z} &= -L(x_1 - x_2) + \frac{f(x_1,t) - f(x_2,t)}{x_1 - x_2}(x_1 - x_2)\\ &= -(L - a(x_1, x_2, t))z_1,\end{aligned}$$

where $a(x_1, x_2, t) = (f(x_1, t) - f(x_2, t))/(x_1 - x_2)$. So, if $a(x_1, x_2, t) < L_1 < L$, for some $L_1 \in \mathbb{R}$ and for all x_1, x_2, t, $x_1 \neq x_2$, then $z(t) \to 0$ when $t \to +\infty$. This shows that in this case $\{x_1 = x_2\}$ is a synchronization manifold. We then have the following result.

Theorem 10 *If $f_1 = f_2 = f$ e for some $L_1 \in \mathbb{R}$ we have*

$$\frac{f(x_1,t) - f(x_2,t)}{x_1 - x_2} < L_1 < L,$$

for all x_1, x_2, t, $x_1 \neq x_2$, then the system (22) synchronizes with $\{x_1 = x_2\}$ as the synchronization manifold.

What we have just seen is a case of identical synchronization. Now let's see what happens when the oscillators are not necessarily equal. Note that we can also write the system (22) in a matrix form similar to (7)

$$\begin{pmatrix} x_1' \\ x_2' \\ y' \end{pmatrix} = \begin{pmatrix} f_1(x_1,t) \\ f_2(x_2,t) \\ h(y,t) \end{pmatrix} + L \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix}$$

In order to apply the results of Sect. 3, we will solve (12) for $\lambda \in]0, L[$ and $D = LA + \lambda I$ obtaining

$$P = \begin{pmatrix} -\frac{L^2 - 3L\lambda + \lambda^2}{2\lambda(\lambda - 3L)(\lambda - L)} & -\frac{L^2}{2\lambda(\lambda - 3L)(\lambda - L)} & \frac{L}{2\lambda(\lambda - 3L)} \\ -\frac{L^2}{2\lambda(\lambda - 3L)(\lambda - L)} & -\frac{L^2 - 3L\lambda + \lambda^2}{2\lambda(\lambda - 3L)(\lambda - L)} & \frac{L}{2\lambda(\lambda - 3L)} \\ \frac{L}{2\lambda(\lambda - 3L)} & \frac{L}{2\lambda(\lambda - 3L)} & \frac{L - \lambda}{2\lambda(\lambda - 3L)} \end{pmatrix}.$$

Although this matrix does not look nice, the eigenvalues of P are

$$-\frac{1}{2(\lambda - 3L)}, -\frac{1}{2(\lambda - L)}, -\frac{1}{2\lambda},$$

with corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

On the other hand, for $x_1 \neq q_1$, $x_2 \neq q_2$, and $y \neq w$, defining

$$\alpha = \alpha(x_1, q_2, t) = \frac{f_1(x_1, t) - f_1(q_1, t)}{x_1 - q_1},$$

$$\beta = \beta(x_2, q_2, t) = \frac{f_2(x_2, t) - f_2(q_2, t)}{x_2 - q_2},$$

$$\delta = \delta(y, w, t) = \frac{h(y, t) - h(w, t)}{y - w},$$

we can rewrite inequality (13) as

$$\begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ y - w \end{pmatrix}^T \left[\left(\frac{1}{2} - \varepsilon \right) I - P \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \delta \end{pmatrix} \right] \begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ y - w \end{pmatrix} \geq 0.$$

Let us now consider the symmetric matrix of the associated quadratic form

$$\Omega = \frac{1}{2} \left(\left[\left(\frac{1}{2} - \varepsilon \right) I - P \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \delta \end{pmatrix} \right]^T + \left(\frac{1}{2} - \varepsilon \right) I - P \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \delta \end{pmatrix} \right).$$

Which can be written explicitly as

$$\Omega = \begin{pmatrix} \frac{1}{2} - \varepsilon + \frac{\alpha(L^2 - 3L\lambda + \lambda^2)}{2\lambda(\lambda - 3L)(\lambda - L)} & \frac{(\alpha+\beta)L^2}{4\lambda(\lambda-3L)(\lambda-L)} & \frac{(\alpha+\delta)L}{4\lambda(3L-\lambda)} \\ \frac{(\alpha+\beta)L^2}{4\lambda(\lambda-3L)(\lambda-L)} & \frac{1}{2} - \varepsilon + \frac{\beta(L^2 - 3L\lambda + \lambda^2)}{2\lambda(\lambda-3L)(\lambda-L)} & \frac{(\beta+\delta)L}{4\lambda(3L-\lambda)} \\ \frac{(\alpha+\delta)L}{4\lambda(3L-\lambda)} & \frac{(\beta+\delta)L}{4\lambda(3L-\lambda)} & \frac{1}{2} - \varepsilon + \frac{\delta(L-\lambda)}{2\lambda(3L-\lambda)} \end{pmatrix}.$$

In this way, the inequality (13) in Theorem 6 is verified if the matrix Ω is positive defined. We then have the following result.

Theorem 11 *Suppose there is a* $\lambda \in (0, L)$ *and* ε *such that* Ω *is positive definite for all* x_1, x_2, y, q_1, q_2, w, $x_1 \neq x_2$, $q_1 \neq q_2$, $y \neq w$. *Assuming that there is at least one amenable solution, then there is a synchronization manifold that can be seen as a graph of the subspace generated by* $(1, 1, 1)^T$.

We can try to see under what conditions Ω is positive definite, which can be done in two ways, by calculating eigenvalues or studying the smallest ones. We will follow the second option.

Our intuition says that if the α, β and δ coefficients are bounded and if L is large empty, then the system synchronise, that's what the next theorem says.

Theorem 12 *Let's assume that the system (22) has at least one amenable solution and the* α, β *and* δ *coefficients are bounded. So, if* L *is large enough, there is a synchronization manifold that can be seen as a graph with domain in the subpace spanned by* $(1, 1, 1)^T$.

Proof If we choose a concrete value for λ, the previous expressions are much simpler. Let's choose $\lambda = L/2$. For this choice of λ, the minors of Ω are

$$\begin{aligned} m_1(\alpha) =& \frac{1}{2} - \varepsilon + \alpha \frac{2}{5L}; \\ m_2(\alpha, \beta) =& \left(\frac{1}{2} - \varepsilon + \alpha \frac{2}{5L}\right)\left(\frac{1}{2} - \varepsilon + \beta \frac{2}{5L}\right) - \frac{4}{25L^2}(\alpha + \beta)^2; \\ m_3(\alpha, \beta, \delta) =& \left(\frac{1}{2} - \varepsilon + \frac{\delta}{5L}\right) m_2(\alpha, \beta) + \frac{4}{125L^2}(\alpha + \beta)(\alpha + \delta)(\beta + \delta) \\ & - \frac{1}{25L^2}\left(m_1(\alpha)(\beta + \delta)^2 + m_1(\beta)(\alpha + \delta)^2\right). \end{aligned}$$

As we can see, if α, β and δ are bounded, we can choose a sufficiently large L and an ε small enough so that the three minors are all positive. The result is then a consequence of the previous theorem. □

In [13] we can see a study on the geometry of the values of α, β and δ that make the matrix Ω positive definite, and a numerical study on this same set as well.

9 Conclusion

We studied several examples of oscillators and coupling schemes that lead to the so-called generalised synchronisation. The results were mainly obtained through the use of a general result on the existence of invariant manifolds that attract the bounded solutions of the system. Even if we considered, as an example, some cases where the oscillators are equal, and where the synchronization manifold is a diagonal, the full potential of our methods is shown when we consider systems of oscillators which are not equal, or coupling schemes without symmetry, where the synchronization manifold is no longer a diagonal.

It would be interesting to consider similar but not equal oscillators and to obtain convergence results of the synchronization manifold to the diagonal when a parameter, which measures the difference between the oscillators, tends to zero.

It would be interesting to obtain results like those obtained in Sect. 8 but for more natural oscillators. For example second order equations modelling mechanical oscillators. Huygens' experiments were done with relatively heavy clocks on top of a wooden shelf. There are some works trying to reproduce Huygens' conditions in some way, building replicas of the clocks developed in the 17th century [17] or experimenting with two metronomes on top of a board, suspended in two empty cans [15]. The model of two coupled pendula that we presented in the introduction does not capture the dynamics of the situation described by Huygens, the case we considered is a simple model of two pendula connected directly by a spring. Huygens' setup is much more the type of systems studied in Sect. 8, each pendulum interferes with a medium, which in this case is the common support. When we leave just

one metronome on the board it is very clear that the moment of the pendulum in motion creates an oscillation in the board, it is this oscillation that interferes with the movement of the second pendulum. It is not difficult to write equations for this situation, we can see this deduction for example in [15]. Where its concluded that two pendula, whose position is described by the variables x_1, x_2, on top of a base that moves in a direction parallel to the pendula, are described by the following system of equations,

$$\begin{cases} x_1'' + \gamma \left(\left(\frac{x_1}{x_0} \right)^2 - 1 \right) x_1' + sin(x_1) - L\, cos(x_1)(\, sin(x_1) + \, sin(x_2))'' = 0 \\ x_2'' + \gamma \left(\left(\frac{x_2}{x_0} \right)^2 - 1 \right) x_2' + sin(x_2) - L\, cos(x_2)(\, sin(x_1) + \, sin(x_2))'' = 0 \end{cases}$$

It would be nice to gain a deeper understanding of this system from an analytical point of view and eventually present intervals for the parameters γ and L so that we have synchronization.

References

1. Birkhoff, G., Rota, G.-C.: Ordinary Differential Equations, 4th edn. Wiley, New York (1989)
2. Baker, G.L., Blackburn, J.A.: The Pendulum, a Case Study in Physics. Press, Oxford Uni (2005)
3. Boccaletti, S., Kurths, J., Osipov, G., Valladares, D.L., Zhou, C.S.: The Syncronization of Chaotic Systems. Physics Reports **366**, 1–10 (2002)
4. Hoppensteadt, F.C., Levi, M., Miranker, W.L.: Dynamics of the Josephson Junction. Quart. Appl. Math. **36**, 167–198 (1978)
5. Horn, R., Johnson, C.: Topics in Matrix Analysis. Press, Cambridge Uni (1991)
6. Katriel, G.: Synchronization of Oscillators Coupled Through an Environment. Physica D **237**, 2933–2944 (2008)
7. Lax, P.: Linear Algebra. Wiley, New York (1997)
8. Levi, M.: Nonchaotic Behavior in the Josephson Junction. Physical Review A **37**, 927–931 (1988)
9. Martins, R.: The Effect of Inversely Unstable Solutions on the Attractor of the Forced Pendulum Equation With Friction. J. Differential Equations **212**, 351–365 (2005)
10. Martins, R.: One-dimensional Attractor for a Dissipative System With a Cylindrical Phase Space. Discrete and Continuous Dynamical Systems-Series A **14**, 533–547 (2006)
11. Martins, R.: The Attractor of an Equation of Tricomi's Type. J. Math. Anal. Appl. **342**, 1265–1270 (2008)
12. Margheri, A., Martins, R.: Generalized Synchronization in Linearly Coupled Time Periodic Systems. Journal of Differential Equations **249**, 3215–3232 (2010)
13. Martins, R., Morais, G.: Generalized Synchronization in a System of Several Non-autonomous Oscillators Coupled by a Médium. Kybernetika **51**, 347–373 (2015)
14. Min, Q., Xian, S., Jinyan, Z.: Global Behavior in the Dynamical Equation of J-J type. J. Differential Equations **71**, 315–333 (1988)
15. Pantaleone, J.: Synchronization of Metronomes. Am. J. Phys. **70**, 992–1000 (2002)
16. Pecora, L., Carroll, T., Johnson, G., Mar, D., Heagy, J.: Fundamentals of Synchronization in Chaotic Systems, Concepts, and Applications. Chaos **7**, 520–543 (1997)
17. Ramirez, J., Olvera, L., Nijmeijer, H., Alvarez, J.: The Sympathy of Two Pendulum Clocks: Beyond Huygens Observations. Scientific Reports **6**, 1–16 (2016)

18. Smith, R.A.: Absolute Stability of Certain Differential Equations. J. London Math. Soc. **7**, 203–210 (1973)
19. Smith, R.A.: Massera's Convergence Theorem for Periodic Nonlinear Differential Equations. J. Math. Anal. Appl. **120**, 679–708 (1986)
20. Sobel D.: Longitude: the true story of a lone genius who solved the greatest scientific problem of his time. Walker & Company (1995)
21. Srzednicki, R.: Wazewski method and Conley index. In: Handbook of Differential Equations, pp. 591–684. Elsevier/North-Holland, Amsterdam (2004)
22. Strogatz, S.: Sync. The Emerging Science Of Spontaneous Order. Penguin (2004)
23. Wazewski, T.: Sur une Méthode Topologique de l'Examen de l'Allure Asymptotique des Intégrales des Équations Différentielles. In: Proceedings of the International Congress of Mathematicians, 1954, Amsterdam, vol. III, pp. 132–139. Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co., Amsterdam (1956)
24. Wu, C.W., Chua, L.O.: Synchronization in an array of linearly coupled dynamical systems. IEEE Trans. Circuits Syst.-I: Fundam. Theory Appl. **42**, 430–447 (1995)

Demand Forecasting with Clustering and Artificial Neural Networks Methods: An Application for Stock Keeping Units

Zehra Kamisli Ozturk, Yesim Cetin, Yesim Isik, and Zeynep Ídil Erzurum Cicek

Abstract **Introduction**: Firms' production strategy consists of interrelated strategic decisions, including pricing, demand forecasting and demand response planning, capacity planning, capital and cost structure. Production strategy is also the most important factor affecting the overall return of companies. Companies must estimate future situations in order to maintain their current position. Nowadays, with the developing technology, companies contain a wide variety of data. With various data analytics and optimization methods and tools, the data that can help the decision making process of companies can be made meaningful and usable. **Objective**: For companies like sanitary wares with large number of product variants, product groups based on estimated requirements will emerge. The ceramic sanitary ware sector, where the product variety is very high, shows seasonal effects in itself and seasonal and trend effects in its products. This situation makes it difficult to make estimations in the ceramic sector. The aim of this study is to increase the accuracy rate of demand estimation ratio by more than 70%. **Methods**: First of all, k-means clustering algorithm is used to obtain product groups. Then, an artificial Neural Networks model is used to estimate demands of product groups. **Results**: The obtained estimation error ratios are compared with those which are obtained by exponential smoothing and moving average methods in time series methods. It is observed that the most suitable method is Artificial Neural Networks (ANNs) for obtaining best results. **Conclusion**: The results prove the efficiency of the applying ANNs to clustered products as a nature-inspired method for demand estimation problem.

Keywords Clustering · k-means algorithm · Forecasting · Artificial Neural Networks (ANN)

Z. K. Ozturk (✉) · Y. Cetin · Y. Isik · Z. Í. E. Cicek
Department of Industrial Engineering, Eskisehir Technical University, Eskisehir, Turkey
e-mail: zkamisli@eskisehir.edu.tr

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_15

1 Introduction

Accurate forecasting for manufacturing companies helps to respond to orders faster and indirectly increases customer satisfaction. Therefore, companies have to work on demand forecasting in order to respond to their customers in a timely manner and increase their profits. Thus, they retain their competitive advantages.

Demand forecasts form the basis for many managerial decisions in demand planning, ordering, production planning and stock control. It is usually difficult to carry out forecasting with a desired level of precision because of the volatility and varying uncertainties involved [1]. Since inventory policies are directly influenced by the forecast results, demand forecasts that are above or below the actual values bring additional costs for firms. On the other hand, if we try to customize the demand forecasting problem of companies, product groups will emerge based on the estimated requirements for companies with a wide range of products. As Hofer and Halman [2] noted, although the variety of products promotes sales, companies should control the cost of inventory of these products and ensure the best quality and timely delivery to their customers.

The increase in product diversity and the factors affecting demand requires the analysis of big data. Ren et al. [3] noted that the advent of the big data age leads to a shift in demand forecasting for fashionable products and is also a major challenge for traditional forecasting methods and inventory planning.

Cluster analysis has been used frequently in product positioning and market segmentation studies [4]. Based on Simpson et al. [5], Moon et al. [6] defined a product family as a group of related products based on a product platform, facilitating mass customization by providing a variety of products for different market segments cost-effectively. We can obtain product groups for existing products by clustering as a data mining technique.

In this study, we focus on a ceramic sanitary ware company's demand forecasting problem. The ceramic sanitary ware sector, where the product variety is very high, shows seasonal effects in itself and seasonal and trend effects in its products. This situation makes it difficult to make estimations in the ceramic sector. As a matter of fact, the accuracy rate of demand estimation ratio which is generally accepted by the company is around 70%. In this study, we aim to forecast the demands for the company by using a clustering technique and demand forecasting methods and also by increasing the current accuracy rate. To reach these goals, first of all we determine product groups of the company by k-means algorithm that is one of the clustering algorithms. After then, to forecast the demands of these groups the artificial neural networks (ANN) is used. Finally, moving averages and exponential smoothing methods, which are the classical time series forecasting methods are used to evaluate the performance of ANN.

The remainder of this study is as follows. Section 2 gives a literature review about demand forecasting and also clustering. Section 3 introduces the demand forecasting

method developed for companies with a wide range of products and discusses a real-world case study results. The computational results are given in Sect. 4. Finally, conclusions and remarks are given in Sect. 5.

2 Literature Review

The literature on forecasting contains a variety of methods. The range of different approaches includes state-space methods with the Kalman filter, general exponential smoothing, artificial neural networks, spectral methods and seasonal ARIMA (autoregressive integrated moving average) models [7]. As Castillo et al. [8] have stated, time-series forecasting methods is one of the most widely used techniques to solve problems of sales forecasts. However, the effectiveness of these techniques is highly dependent on the field of application and the accuracy of the problem data. Minimum mean square error, moving average and exponential smoothing techniques are the basic estimation techniques. In the forecasting area, an 'optimal' forecasting model traditionally indicates that the forecasting model has the smallest mean square estimation errors. In general, moving average and exponential smoothing estimation techniques are the most commonly used estimation techniques in practice as a result of their ease of use, flexibility and robustness, although they do not share this optimum feature for a time series process [9].

Sudheer and Suseelatha [10] proposed a hybrid method based on wavelet transform, Triple Exponential Smoothing model and weighted nearest neighbor model for short term load forecasting. Jiang et al. [11] applied ARIMA model to predict coal consumption, price, and investment in China. Aydin [12] used multiple linear regression analysis (MLRA) to predict energy-related CO2 emissions in Turkey in 2013–2020.

To more accurately forecasting energy demand in China and India, Wang et al. [13] developed linear (MGM), hybrid-linear (MGM–ARIMA), and non-linear time-series forecast techniques (NMGM) based on the grey model.

As Taylor [7] mentioned, the most noticeable development in demand forecasting over the last decade has been the increasing interest shown by researchers and practitioners in artificial neural networks (ANNs). One of the leading application areas of ANNs is time series forecasting besides classification, clustering, pattern recognition etc. In recent years, ANNs have been successfully applied as a tool in the prediction of time series, mainly due to the ability of ANNs to capture the linear and nonlinear relationships in the data [14]. ANNs attempt to simulate the structure of the human brain, their thinking capabilities and learning in a machine. This gives an advantage for the modelling of complex and nonlinear data. It is neuroscience which has inspired neural networks and not because they are considered good models of biological phenomenon. These networks are "neural" in the sense that they have been inspired by neuroscience but not necessarily because they are faithful models of biological cognitive phenomena [15]. ANNs are usually used for forecasting in different research areas like parameter estimation [16, 17], forecast of atmospheric radiation

and ozone concentration [18], solution of strictly convex quadratic programming problems [19], pattern recognition [20], etc.

If we look at the studies of ANNs in terms of the content of this work, Jaipuria and Mahapatra [21] indicated that an ANN is preferred as a superior forecasting model because it addresses the limitations of time series models by efficient non-linear mapping between input and output data. Al-Saba and El-Amin [22] showed that, their proposed ANN model provides accurate results using a minimum amount of historical data rather than other forecasting techniques. Geem and Roper [23] better estimated energy demand with an ANN model than a linear regression model or an exponential model in terms of root mean squared error. Al-Saba and El-Amin [22] also used the ANN model to forecast the energy requirements of an electric utility. They compared the obtained results with time series models. The comparison revealed that the ANN produced close to the actual data. Paul and Azeem [24] determined the optimal level of inventory using an ANN, which was supposed to be a function of product demand, material costs, setup and holding cost. In their study Kochak and Sharma [25] observed performance of product demand forecasting for a real company by ANN.

Although ANN is a good estimator, when the number of independent variables to be estimated is too much, estimating with clustered units is usually much easier than estimating each unit individually. As Han and Kamber [26] defined, clustering is a process of grouping a set of physical or abstract objects into classes of similar objects. A cluster is a collection of objects that are similar to one another within the same cluster, yet dissimilar to the objects in other clusters. Clustering is a main task of data mining and is widely used in statistics and science [27]. In the survey of clustering data mining techniques, Berkhin [27] mentioned the application areas of clustering algorithms as information retrieval and text mining, spatial database applications, web applications, sequence and heterogeneous data analysis and DNA analysis in computational biology. Another special example can be given by Thanh et al. [28] about clustering applications for medical diagnosis.

3 Solution Approaches For Demand Forecasting for the Sanitary Company

In this study, demand forecasting issue was investigated on a manufacturing company as a real-world case study. Although the sanitary wares sector is seasonally affected, both seasonal and trend effects are seen on the products. In the company which has more than 700 different models, the demand forecast has been based on the data of information coming from the markets, the determined budget numbers, the sales information of the recent past and the orders in the system. Besides this insufficient data, unfortunately no special method has been used for demand forecasting.

In the considered company, the demand forecasts are obtained by a decision maker who determines the forecast values via past experiences. Because the forecasting

process doesn't have a predefined method, it is impossible to obtain forecast without the decision maker. It is also hard to enter manually the forecasts one by one to the ERP system that company used. This situation is an obstacle to carry out a forecast process every month. In the long term, overstocking and order delay problems occur because of the irregular forecast process. Moreover, a validation of the applied forecast process is not performed and disruptive aspects of the process cannot be reviewed. Due to the production cycle of the firm, it is expected to realize at least 2 months demand estimation with minimum accuracy of 80% for the first month and minimum 70% for the second month due to the application being developed for the stock keeping units depending on the production planning department.

As a result of these issues, it is seen that demand forecasting with a high accuracy is very important for the company. The products are produced by classical and pressure casting methods. In the classical casting process, the lifetime of a mold connected to looms is 90 days, and production continues in the same volume and in a series until it reaches its end of life. In the pressure casting method, it is possible to produce the desired number of pieces with automatic machines. Particularly in periods when the demand falls, production volume is constant in classical casting machines, so production is done for stock. This leads to an increase in the inventory holding costs.

3.1 Data Collection and Analysis

The data gathered to enable this study are (i) 3-year monthly order information of the products, (ii) launch dates of the products, (iii) release dates of the products, (iv) market information about the products they sell, and (v) information about which product group the products belong to. In addition to these existing features, the *coefficient of variation*, which is a coefficient indicating the prevalence of the standard deviation distribution, was calculated. The mean and standard deviations were calculated to determine the variability of the data, and the coefficient of variation was found. Based on only standard deviation information, it is difficult to comment on the distribution of data. For this reason, the coefficient of variation was applied. The coefficient of variation was calculated using Eq. (1):

$$Coefficient\ of\ Variation = \frac{Standard\ Deviation}{Mean} \times 100. \tag{1}$$

Besides, 1-year orders, 3-year orders and the product type (new or not new) were considered as other features of our data. The demand forecasting was performed for each product by using Matlab Neural Network Toolbox 6.0 package. It was observed that error rates of ANN forecasting are very high. This shows us that estimating with clustered units is usually much easier than estimating each unit individually.

3.2 *Determination of Product Clusters*

Nazemi and Nazemi [19] provided a classification of clustering algorithms as hierarchical methods, partitioning relocation methods, density-based partitioning methods, methods based on co-occurance of categorical data, grid-based methods, scalable clustering algorithms and algorithms for high-dimensional data. In this study, to obtain the product groups k-means clustering, which is one of the partitioning relocation clustering method, was used. The main idea behind k-means clustering is to select k number of cluster centers at first and then partition N observations into k clusters (C_j) by the mean (or weighted average) c_j of its points, such that each observation belongs to the closest cluster center. Cluster is a group of similar objects. Given a set of X observations where each observation is d-dimensional real vector, k-means aims to partition these X observations into k sets. Algorithm tries to minimize within the cluster sum of squares that is the sum of the squares of errors between the points and the corresponding centroids [27, 29] as given in Eq. (2).

$$E(C) = \sum_{j=1:k} \sum_{x_i \in C_j} ||x_i - c_j||^2. \tag{2}$$

The obtained data includes different features about 700 products. The features are 3 year order information, product groups to which they belong to, market information to which they are offered for sale, coefficient of variation and whether or not the product is a new type. R programming language was used to perform the clustering analysis of the data in accordance with the determined features.

The obtained data that were explained in Sect. 3.2 were transferred to the R Studio program for the clustering analysis. Using the k-means clustering function in the program, the initial cluster number values were tried as 7, 8, and 9, and the distribution graphs of the clusters were examined and it was observed that the ideal distribution was obtained with $k = 8$ values. The distribution graph of the 8 clusters is given in Fig. 1.

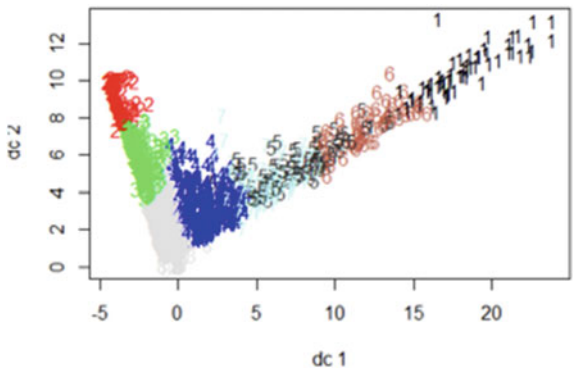


Fig. 1 The cluster distribution graph

After obtaining the product groups, demand forecasts were done for each group in the next sub-section.

3.3 Forecasting with Artificial Neural Networks

3.3.1 Artificial Neural Networks

Artificial neural networks (ANNs), originally developed to mimic basic biological neural systems—the human brain particularly, are composed of a number of interconnected simple processing elements called neurons or nodes [30]. A simple ANN model has three layers: an input, a hidden and an output layer. The input data is supplied by input layer. With linear and nonlinear operations, the input data is processed and the output is generated in the hidden layer. The output layer transforms the output of the hidden layer to the output form which we are interested in. The architecture of a simple ANN model is given in Fig. 2. Back propagation (BP) neural network has been widely used for various data predictions in application [10]. BP algorithm is used in the training phase of supervised ANNs to update the weights between the neurons by observing the error function. The algorithm looks for the minimum of the error function in weight space using the method of gradient descent [31]. To calculate the gradient of the error function, the activation function must be chosen as a differentiable function. Tanh, sigmoid and ReLu functions are known as activation functions in ANNs. In this study, sigmoid function is chosen for the implemented ANN model. The reason for selecting sigmoid function is that sigmoid function is the most preferred activation function in ANN models. The sigmoid function is given in Eq. (3) and Fig. 3 below:

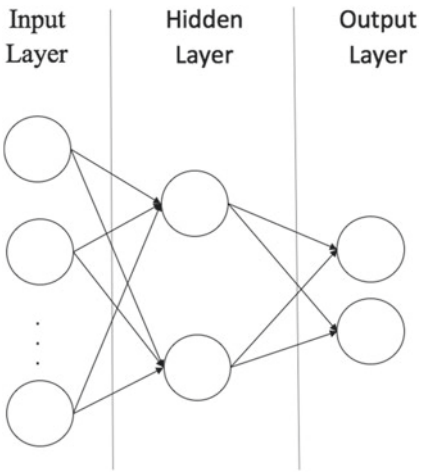


Fig. 2 The architecture of a simple ANN

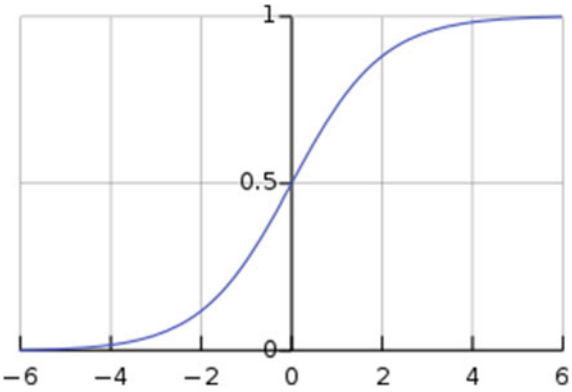


Fig. 3 The graph of the sigmoid function

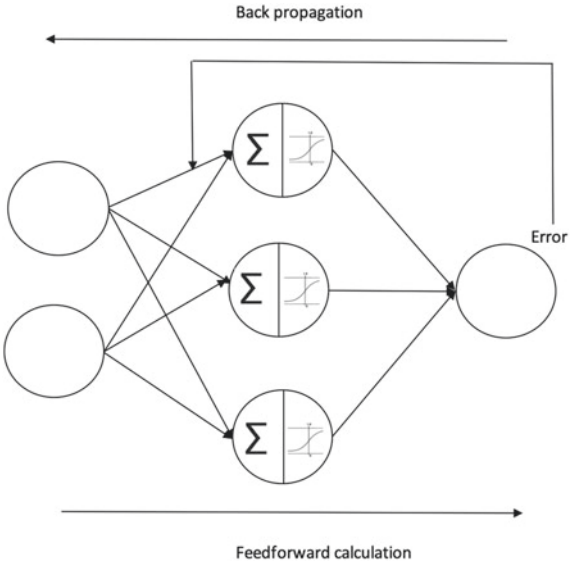


Fig. 4 The working principle of back propagation

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{3}$$

Gradients are computed using chain rule in gradient descent algorithm. A predefined learning rate determines the influence of the gradient. In every iteration, the error values and consequently gradients are calculated for training dataset and thus, the weight values are updated. The main working principle of BP is given in Fig. 4.

3.3.2 The Implementation of the ANN

The necessary arrangements were made for each of the clusters obtained using the R Studio statistical program. After these arrangements, the features that have different scales for each cluster are normalized and the demand estimation is done in the Matlab Neural Network Toolbox 6.0 package.

It is not enough to use directly demand or sales information in forecasting studies. Therefore, global data on the sector of interest should also be taken into account. Also in this study, in order to make predictions for each cluster consisting of data of more than 700 models between the years 2014 and 2016, the inflation rates, internal sales quantity (tonnes), number of competitors in the ceramic sanitary ware sector and annual production quantity (tonnes) are determined as variables. The values of internal ceramic sanitary ware sales quantity (tonnes), number of competitors in the ceramic sanitary ware sector and annual production quantity (tonnes) are obtained from [32]. As the training dataset, the data of the first two years values are considered as input, the number of orders for 2014 and 2015 are used as the target data. The number of orders for 2016 were used as test data. The ANN model is generated by using MATLAB Neural Network Toolbox. TRAINGDX algorithm in the toolbox which is widely used in estimation is used as training algorithm as gradient descent with momentum and adaptive learning rate backpropagation. LEARNGDM is used as learning algorithm. The number of iterations is 1000. An ANN model was created based on the number of 8 hidden neurons as given in Fig. 5.

ANN training has been carried out until the smallest error rate that can be achieved. The Mean Square Error (MSE) ratios are calculated after the last one year of testing.

4 Computational Results

Firstly, statistical analysis was conducted to reveal if trend exists or not and data which was gathered were visualized. After the regression analysis, the trend-lines were generated and only trend detected only in Cluster 5. There was no evidence of any trend existence in other clusters. Moreover, R^2 values of clusters except Cluster 5 is quite low. The three-year sales data graphs with trend-lines are given in Fig. 6.

There are fluctuations in the order numbers of the products in Cluster 1 and there are not very big differences. There is no general trend effect. There are irregular fluctuations in the order numbers of the products in Cluster 2 and also in Cluster 4. There is no general trend effect for both of these clusters. The seasonal effects may be relevant and the reasons for these fluctuations should be determined by the company. In the case of the products in the Cluster 7, the 3-year order quantities are observed to increase at certain times, but in the majority of the relevant time period orders are around 0. In some periods, sales have increased to very high levels due to some new large-scale construction projects. In the first section, we mentioned that there is seasonality and trend in sales of sanitary wares sector. However, due to the

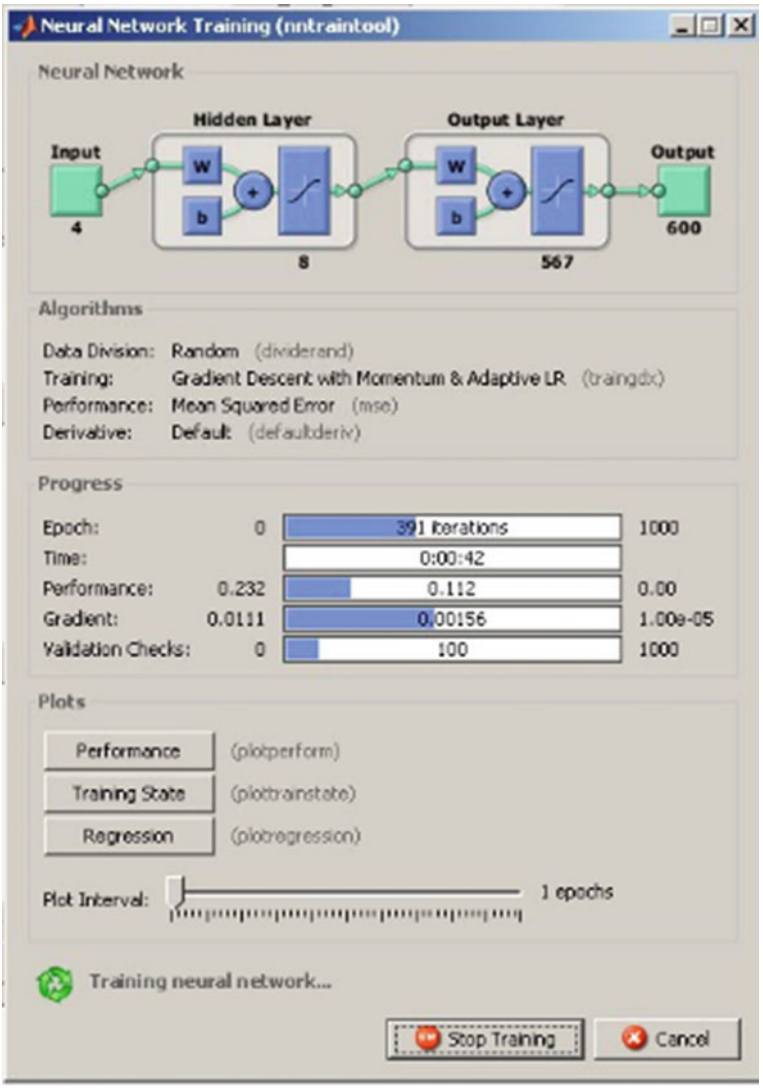


Fig. 5 The ANN model and its traning network

promotional and advertising campaigns applied by the considered company in some periods, amount of sales lost seasonality and trend characteristics.

The most important criterion for choosing a forecasting method is its accuracy. A forecasting method with small forecasting error is generally accepted as good and the forecasting error is calculated based on the difference between the forecast and actual values. Mean Absolute Deviation (MAD), Mean Square Error (MSE), and Mean Absolute Percent Error (MAPE) are utilized as typical measures to determine

Fig. 6 Graphs of the three-year sales data and trend-lines

Table 1 The error rates of each method

	ANN				Moving Averages				Exponential Smoothing			
	1st Month		2nd Month		1st Month		2nd Month		1st Month		2nd Month	
	MSE	MAD	MSE	MAD	MSE	MAD	MSE	MAD	MSE	MAD	MSE	MAD
Cluster1	0.227*	0.35778	0.232*	0.38092	80277	119.7739	314329	198.0368	54.216	134	314329	185
Cluster2	0.208*	0.353881	0.217*	0.37448	79875	274	175357	291	77581.619	207.55582	175357	243.94929
Cluster3	0.195*	0.327763	0.202*	0.334158	32679.767	78.71947	25542.22	74.13119	72969.001	92.516852	25542.224	82.062409
Cluster4	0.237*	0.342882	0.233*	0.341502	35974.538	26.94272	25362.41	27.87595	137404.62	38.501136	25362.408	43.898355
Cluster5	0.127*	0.276106	0.212*	0.364619	127.330	207.0147	244.576	260.7588	157699.09	236.80967	244.576	273.1802
Cluster6	0.195*	0.246718	0.194*	0.245022	2563.2849	7.669132	267.1487	3.986193	2695.2382	9.505701	267.14867	6.0840457
Cluster7	0.137*	0.159821	0.19*	0.212746	266	1.837753	74	1.308712	471.26381	3.0830161	74	3.4468818
Cluster8	0.242*	0.409203	0.228*	0.406529	122474.88	224.5381	208279.3	335.3305	112483.96	245.73457	208279.33	253.88889

accuracy of forecasting methods [33]. Artificial neural networks, moving average and exponential smoothing estimation methods were applied on 8 clusters obtained. MSE (mean square error) and MAD (mean absolute deviation) values were calculated on the obtained results according to Eqs. (4) and (5), respectively. As shown from Table 1, the MSE and MAD values of ANN method is less than the MSE and MAD values of moving averages and exponential smoothing methods for each cluster and for each month. Hence, it can be said that the ANN method is more successful and suitable for this specified forecasting problem.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(D_t - F_t)^2\,, \tag{4}$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|D_t - F_t|\,, \tag{5}$$

where D_t is the demand in period t, F_t is the forecast for period t and n is the total number of periods.

It is obviously seen from Table 1, ANN model outperformed the other forecasting techniques for both of the forecasting periods. The other statistical models performed were rather weak in forecasting the considered sales dataset. This also supports that ANN is more successful in data with high variability.

5 Conclusion

The biggest expense items of the companies are the costs of raw materials and holding costs. In this study, which is based on minimizing holding costs, the company will be able to react more quickly to orders and deliver faster to customers as a result of demand forecasts performed with high accuracy. This situation will bring customer satisfaction and increase sales. The proposed methodology, demand forecasting issue was investigated on a manufacturing company as a real-world case study. The sanitary wares company that is the subject of this study carries out domestic and foreign sales. It has high sales rates in Europe, especially in Germany.

The demand forecasting process requires a systematic approach. In addition to the variables within the system that can be controlled by the enterprise, external factors must also be considered. Factors such as the economic situation of the country, the situation of the construction sector, product trends can be examined. Therefore, these kind of factors are considered in this study.

The effectiveness of the techniques used in demand forecasting is highly dependent on the field of application and the accuracy of the problem data. As the focus of this study is to require large amounts of data to estimate demand in the industry, traditional forecasting methods are not appropriate. For this reason, ANN, one of the artificial intelligence techniques, was used.

If this study is evaluated in terms of sustainability and entrepreneurship; the following comments can be made. The fact that the forecasted values obtained at the end of the study have a great accuracy rate will provide stock optimization and order fulfillment maximization with new forecast values to be obtained. In terms of sustainable development, stock optimization and order fulfillment maximization will increase customer satisfaction. The study will also make a great contribution to the company in terms of entrepreneurship if it goes to the way of increasing the market places and opening up to new markets with the estimated values obtained.

When this study is considered from the innovation standpoint, it is known that the demand estimation ratio of the ceramic industry, which is accepted generally by the firm, is about 70% of the accuracy percentage. As a result of the studies carried out, better value is obtained in estimation, it will be possible to adapt the methods used in the entire ceramic sector.

As a result, stock quantities can be reduced due to the nearest approximation of the minimum 2-month demand forecast values obtained. With the reduction of stock quantities, expansion in warehouse areas in the factory will come into play. In the

normal conditions, due to the safety rules of labor and ceramic sanitary ware in the warehouse layout are heavy products, fifth floor of the shelves should not be placed on the shelf arrangement. Thanks to this expansion in the depot, OSH (Occupational Health and Safety) errors in the product location will be minimized and workforce benefits will be provided. Transport of the products in the warehouse is carried out through forklifts. Expansion of the warehouse will also result in a reduction in forklift movements, which will result in reduced energy consumption and accident risks.

Acknowledgements This study has been partially supported by Eskisehir Technical University Scientific Research Projects Committee (ESTUBAP-19ADP048).

References

1. Abolghasemi, M., Gerlach, R., Tarr, G., Beh, E.: Demand forecasting in supply chain: the impact of demand volatility in the presence of promotion (2019). arXiv preprint arXiv:1909.13084
2. Hofer, A.P., Halman, J.I.M.: The potential of layout platforms for modular complex products and systems. J. Eng. Design **16**, 237–255 (2005)
3. Ren, S., Chan, H.L., Siqin, T.: Ann Oper Res (2019). https://doi.org/10.1007/s10479-019-03148-8
4. Arabie, P., Carroll, D., DeSarbo, W., Wind, J.: Overlapping clustering: a new method for product positioning. J. Market. Res. **18**, 310–317 (1981)
5. Simpson, T.W., Siddique, Z., Jiao, J.: Product Platform And Product Family Design: Methods and Applications, 1st edn. Springer, New York, NY (2006)
6. Moon, S.K., Simpson, T.W., Kumara, S.R.T.: A methodology for knowledge discovery to support product family design. Ann. Oper. Res. **174**, 201–218 (2010)
7. Taylor, J.W.: Short-term electricity demand forecasting using double seasonal exponential smoothing. J. Oper. Res. Soc. **54**, 799–805 (2003)
8. Castillo, P.A., et al.: Applying computational intelligence methods for predicting the sales of newly published books in a real editorial business management environment. Knowl. Based Syst. **115**, 133–151 (2017)
9. Ma, Y., Wang, N., Che, A., Huang, Y., Jinpeng, X.: The bullwhip effect on product orders and inventory: a perspective of demand forecasting techniques. Int. J. Product. Res. **51**(1), 281–302 (2013). https://doi.org/10.1080/00207543.2012.676682
10. Sudheer, G., Suseelatha, A.: Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models. Elect. Power Energy Syst. **64**, 340–346 (2015)
11. Jiang, S., Yang, C., Guo, J., Ding, Z.: ARIMA forecasting of China's coal consumption, price and investment by 2030. Energy Sources Part B Econ. Planning Policy **13**(3), 190–195 (2018). https://doi.org/10.1080/15567249.2017.1423413
12. Aydin, G.: The development and validation of regression models to predict energy-related CO2 emissions in Turkey. Energy Sources Part B Econ. Planning Policy **10**(2), 176–182 (2015). https://doi.org/10.1080/15567249.2013.830662
13. Wang, Q., Li, S., Li, R.: Forecasting energy demand in China and India: using single-linear, hybrid-linear, and non-linear time series forecast techniques. Energy **161**, 821–831 (2018)
14. Vasquez, J.L., Perez, S.T., Travieso, C.M., Alonso, J.B.: Meteorological prediction implemented on field-programmable gate array. Cognit. Comput. **5**, 551–557 (2013)
15. Deb M., Kaur P., Sarma K.K.: Inventory control using fuzzy-aided decision support system. In: Bhatia, S., Mishra, K., Tiwari, S., Singh, V. (eds.), Advances in Intelligent Systems and Computing. Singapore, pp. 467–476. Springer (2017)

16. Davraz, M., Kilincaraslan, S., Ceylan, H.: Predicting the poisson ratio of lightweight concretes using artificial neural network. Acta Physica Polonica A **128**, 184–186 (2015)
17. Güven, A., Günal, A.Y., Günal, M.: Multi-output neural networks for estimation of synthetic unit hydrograph parameters: a case study of a catchment in Turkey. Acta Physica Polonica A 2017; 132:591–594
18. Gao, X., Huang, T., Wang, Z., Xiao, M.: Exploiting a modified gray model in back propagation neural networks for enhanced forecasting. Cognit. Comput. **6**, 331–337 (2014)
19. Nazemi, A., Nazemi, M.: A Gradient-based neural network method for solving strictly convex quadratic programming problems. Cognit. Comput. **6**, 484–495 (2014)
20. Veer, K., Sharma, T.: A novel feature extraction for robust EMG pattern recognition. J. Med. Eng. Technol. **40**, 149–154 (2016)
21. Jaipuria, S., Mahapatra, S.S.: An improved demand forecasting method to reduce bullwhip effect in supply chains. Expert Syst. Appl. **41**, 2395–2408 (2014)
22. Al-Saba, T., El-Ami, I.: Artificial neural networks as applied to long-term demand forecasting. Artif. Intell. Eng. **13**, 189–197 (1999)
23. Geem, Z.W., Roper, W.E.: Energy demand estimation of South Korea using artificial neural network. Energy Policy **37**, 4049–4054 (2009)
24. Paul, S., Azeem, A.: An artificial neural network model for optimization of finished goods inventory. Int. J. Industr. Eng. Comput. **2**, 431–438 (2011)
25. Kochak, A., Sharma, S.: Demand forecasting using for supply Chain management. Int. J. Mechn. Eng. Robot. Res. **4**, 96–104 (2015)
26. Han, J., Kamber, M.: Data Mining: Concepts and Techniques, 2nd edn. Morgan Kaufmann Publication, Waltham, USA (2006)
27. Berkhin, P.: A survey of clustering data mining techniques. In: Kogan, J., Nicholas, C., Teboulle, M. (eds.) Grouping Multidimensional Data. Springer, Berlin, Heidelberg (2006)
28. Thanh, N.D., Ali, M., Son, L.H.: A novel clustering algorithm in a neutrosophic recommender system for medical diagnosis. Cognit. Comput. **9**, 526–544 (2017)
29. Edla, D.R., Gondlekar, V., Gauns, V.: HK-means: a heuristic approach to initialize and estimate the number of clusters in biological data. Acta Physica Polonica A **130**, 78–82 (2016)
30. Zhang, G.B.E., Patuwo, M.Y., Hu, Y.: Forecasting with artificial neural networks: the state of the art. Int. J. Forecast. **14**, 35–62 (1998)
31. Rojas, R.: Neural Networks a Systematic Introduction. 1st ed. Berlin Heidelberg New York Hong Kong London Milan Paris Tokyo, pp. 151– 183. Springer (1996)
32. Turkish Construction Sector Report (2016). http://www.yapi.com.tr/TurkYapiSektoruRaporu2016/index. Accessed 1 Feb 2017
33. Ha, Ch., Seok, H., Ok, Ch.: Evaluation of forecasting methods in aggregate production planning: a cumulative absolute forecast error (CAFE). Comput. Industr. Eng. 118, 329–339. ISSN 0360–8352. DOI (2018). https://doi.org/10.1016/j.cie.2018.03.003

On the Grey Obligation Rules

O. Palancı, S. Z. Alparslan Gök, and Gerhard-Wilhelm Weber

Abstract In this paper, we extend obligation rules by using grey calculus. We introduce grey obligation rules for minimum grey cost spanning tree (mgcst) situations. It turns out that the grey obligation rule and the grey Bird rule are equal under suitable conditions. Further, we show that such rules are grey cost monotonic and induce population monotonic grey allocation schemes (pmgas). Moreover, if the game is concave, its (extended) grey Shapley value is a PMGAS. Some examples of pmgas, grey obligation rules and grey Shapley value for mgcst situations are also given.

Keywords Graph theory · Cooperative game theory · Minimum cost spanning tree situations · Grey data · Obligation rules · Population monotonic allocation schemes · Shapley value · Bird rule

1 Introduction

A connection problem arises in the presence of a group of agents, each of which needs to be connected directly or via other agents to a source. If connections among agents are costly, then each agent will evaluate the opportunity of cooperating with other agents in order to reduce costs. In fact, if a group of agents decides to cooperate,

O. Palancı (✉)
Faculty of Economics and Administrative Sciences, Department of Business and Administration, Süleyman Demirel University, 32260 Isparta, Turkey
e-mail: osmanpalanci@sdu.edu.tr

S. Z. Alparslan Gök
Faculty of Arts and Sciences, Department of Mathematics, Süleyman Demirel University, 32260 Isparta, Turkey

G.-W. Weber
Faculty of Engineering Management, Chair of Marketing and Economic Engineering, Poznan University of Technology, ul. Strzelecka 11, 60-965 Poznan, Poland
e-mail: gerhard.weber@put.poznan.pl

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_16

a configuration of links which minimizes the total cost of connection is provided by a minimum cost spanning tree.

The aim of minimum cost spanning tree problems is the building of a network of minimal cost which provides for every node in the network a connection with the source. Examples of minimum cost spanning tree problems are the problem of building a network of computers that connects every computer with some server or the problem of building a drainage system that connects every house in a city with the water purifier. The other example of a minimum cost spanning tree problem is the problem of carpooling. Moreover, the Bird rule [4] plays a special role.

Game theory is a mathematical theory dealing with models of conflict and cooperation. Game theory has many interactions with economics and with other areas such as Operational Research and social sciences. Game theory is divided into two parts: non-cooperative and cooperative. In this paper, we deal with cooperative game theory. Cooperative game theory deals with coalitions who coordinate their actions and pool their winnings. The cooperative game theory is widely used on interesting sharing cost/profit problems in many areas of Operational Research such as connection, routing, scheduling, production, inventory, transportation, etc. (see [5] for a survey on Operational Research Games).

In this paper by [4] provided the first game theoretical treatment of this problem by associating a coalitional game with transferable utility to minimum cost spanning tree (mcst) problems.

A generalized allocation scheme of a game specifies not only how to allocate the worth of full cooperation but also how to allocate the worth of each coalition of players. A population monotonic allocation scheme (PMAS) of a game is a generalized allocation scheme in which the payoff to each player cannot decrease as the coalition to which he/she belongs grows larger. Reference [15] names the set of all PMAS of a game the monotonic core.

Grey system theory is one of the new mathematical theories born out of the concept of the grey numbers. As a matter of fact, many researchers have handled this vagueness with the help of the grey numbers, one of the keystones in grey system theory. When decision information is given by a grey number, games are called cooperative grey games [7, 8, 11, 12]. There are many applications in cooperative game theory with grey uncertainty [1, 10].

As in the classical case, where edge costs are real numbers, also in the situation where edge costs are grey numbers, a cost allocation problem arises. With the goal to study this kind of cost allocation problems, in this paper we extend the notion of an obligation rule by using grey calculus, and we study some cost monotonicity properties. It turns out that cost monotonicity, under grey uncertainty, provides a population monotonic grey allocation scheme.

We note that [17] studied *minimum cost spanning tree* problems in which the connection costs are represented by random variables. In our paper, costs are not random variables, but instead, they are grey numbers.

This paper is organized as follows. Section 2 starts with some terminology on graph theory and the theory of cooperative grey game. In Sect. 3, we introduce minimum grey cost spanning tree situations and related grey games. In Sect. 4, grey

obligation rules are introduced; in the same section, the relation between the grey obligation rules and the grey Bird rule are given. In Sect. 5, we introduce pmgas. It is shown that interval obligation rules are grey cost monotonic and induce population monotonic grey allocation schemes. The conclusion of this work is given in the last section.

2 Preliminaries

In this section we give some preliminaries on graph theory, grey calculus and the theory of cooperative grey games [9, 13, 16].

An (undirected) *graph* is a pair $< V, E >$, where V is a set of vertices or nodes and E is a set of edges e of the form $\{i, j\}$ with $i, j \in V, i \neq j$. The *complete graph* on a set V of vertices is the graph $< V, E_V >$, where $E_V = \{\{i, j\} | i, j \in V \text{ and } i \neq j\}$. A *path* between i and j in a graph $< V, E >$ is a sequence of nodes $i = i_0, i_1, \ldots, i_k = j$, $k \geq 1$, such that all the edges $\{i_s, i_{s+1}\} \in E$, for $s \in \{0, \ldots, k-1\}$, are distinct. A *cycle* in $< V, E >$ is a path from i to i for some $i \in V$. Two nodes $i, j \in V$ are *connected* in $< V, E >$ if $i = j$ or if there exists a path between i and j in $< V, E >$. A *connected component* of V in a graph $< V, E >$ is a maximal subset of V with the property that any two nodes in this subset are connected in $< V, E >$.

A number denoted by $\otimes \in \left[\underline{\otimes}, \overline{\otimes}\right]$, where $\underline{\otimes}$ is called the *lower limit* and $\overline{\otimes}$ is called the *upper limit* for $\otimes$, is called a *grey number*. Now, we give some operations on grey numbers.

Let

$$\otimes_1 \in \left[\underline{\otimes_1}, \overline{\otimes_1}\right], \ \underline{\otimes_1} < \overline{\otimes_1} \text{ and } \otimes_2 \in \left[\underline{\otimes_2}, \overline{\otimes_2}\right], \ \underline{\otimes_2} < \overline{\otimes_2}.$$

The sum of $\otimes_1$ and $\otimes_2$, written as $\otimes_1 + \otimes_2$, is defined as follows:

$$\otimes_1 + \otimes_2 \in \left[\underline{\otimes_1} + \underline{\otimes_2}, \overline{\otimes_1} + \overline{\otimes_2}\right].$$

Let us assume that $\otimes \in \left[\underline{\otimes}, \overline{\otimes}\right]$, $\underline{\otimes} < \overline{\otimes}$, and k is a positive real number. The scalar multiplication of k and $\otimes$ is defined as follows:

$$k\otimes \in \left[k\underline{\otimes}, k\overline{\otimes}\right].$$

We denote by $\otimes(\mathbb{R})$ the set of grey numbers in $\mathbb{R}$. Let $\otimes_1, \otimes_2 \in \otimes(\mathbb{R})$ with $\otimes_1 \in \left[\underline{\otimes_1}, \overline{\otimes_1}\right]$, $\underline{\otimes_1} < \overline{\otimes_1}$; $\otimes_2 \in \left[\underline{\otimes_2}, \overline{\otimes_2}\right]$, $\underline{\otimes_2} < \overline{\otimes_2}$, $|\otimes_1| = \overline{\otimes_1} - \underline{\otimes_1}$ and $\alpha \in \mathbb{R}_+$. Then, $\otimes_1 + \otimes_2 \in \left[\underline{\otimes_1} + \underline{\otimes_2}, \overline{\otimes_1} + \overline{\otimes_2}\right]$ and $\alpha\otimes \in \left[\alpha\underline{\otimes}, \alpha\overline{\otimes}\right]$.

In general, the difference of $\otimes_1$ and $\otimes_2$ is defined as follows:

$$\mathcal{G}_1 \ominus \mathcal{G}_2 = \mathcal{G}_1 + (-\mathcal{G}_2) \in [a-d, b-c]\,,$$

(see [14]).

For example, let $\mathcal{G}_1 \in [6, 8]$, and $\mathcal{G}_2 \in [2, 5]$, then we have

$$\mathcal{G}_1 \ominus \mathcal{G}_2 \in [6-5, 8-2] = [1, 6]\,,$$
$$\mathcal{G}_2 \ominus \mathcal{G}_1 \in [2-8, 5-6] = [-6, -1]\,.$$

Different from the above subtraction we use a partial subtraction operator. We define $\mathcal{G}_1 - \mathcal{G}_2$, only if $|b-a| \geq |d-c|$, by $\mathcal{G}_1 - \mathcal{G}_2 \in [a-c, b-d]$. We recall that $[a, b]$ is weakly better than $[c, d]$, which we denote by $[a, b] \succcurlyeq [c, d]$, if and only if $a \geq c$ and $b \geq d$. We also use the reverse notation $[a, b] \preccurlyeq [c, d]$, if and only if $a \leq c$ and $b \leq d$ (for details see [2, 6]).

Notice that if we make a comparison with the above example, then in our case $[6, 8] - [2, 5]$ is not defined. But, $[2, 5] - [6, 8]$ is defined.

Let $\mathcal{G}_1 \in [2, 5]$, and $\mathcal{G}_2 \in [6, 8]$, $\mathcal{G}_1 - \mathcal{G}_2$ is defined since $|5-2| \geq |8-6|$, but $\mathcal{G}_2 - \mathcal{G}_1$ is not defined since $|8-6| = 2 \ngeq 3 = |5-2|$, then we have

$$\mathcal{G}_1 - \mathcal{G}_2 \in [2-6, 5-8] = [-4, -3]\,.$$

We recall that a cooperative grey cost game is an ordered pair $< N, c' >$, where $N = \{1, \ldots, n\}$ is the set of players, and $c' : 2^N \to \mathcal{G}(\mathbb{R})$ is the characteristic function such that $c'(\emptyset) \in [0, 0]$, the grey payoff function value $c'(S) \in [\underline{c'(S)}, \overline{c'(S)}]$ refers to the value of the grey expectation benefit belonging to a coalition $S \in 2^N$, where $\underline{c'(S)}$ and $\overline{c'(S)}$ represent the maximum and minimum possible profits of the coalition S. Grey solutions are useful to solve reward/cost sharing problems with grey data using cooperative grey games as a tool. The building blocks of grey solutions are grey payoff vectors, i.e., vectors whose components belong to $\mathcal{G}(\mathbb{R})$. We denote by $\mathcal{G}(\mathbb{R})^N$ the set of all such grey payoff vectors, and we designate by $\mathcal{G}G^N$ the family of all cooperative grey games.

We call a game $< N, c' >$ *grey size monotonic* if $< N, \left|c'\right| >$ is monotonic, i.e., $\left|c'\right|(S) \geq \left|c'\right|(T)$ for all $S, T \in 2^N$ with $S \subset T$. For further use we denote by $SM\mathcal{G}G^N$ the class of all grey size monotonic games with player set N.

The grey marginal operators and the grey Shapley value are defined on $SM\mathcal{G}G^N$.

Denote by $\Pi(N)$ the set of permutations $\sigma : N \to N$ of N. The *grey marginal operator* $m^\sigma : SM\mathcal{G}G^N \to \mathcal{G}(\mathbb{R})^N$ corresponding to σ, associates with each $c' \in SM\mathcal{G}G^N$ the *grey marginal vector* $m^\sigma(c')$ of c' with respect to σ defined by

$$m_i^\sigma(c') := c'(P^\sigma(i)) - c'(P^\sigma(i) \setminus \{i\}) \in [\underline{A_{P^\sigma(i)}} - \underline{A_{P^\sigma(i)\setminus\{i\}}}, \overline{A_{P^\sigma(i)}} - \overline{A_{P^\sigma(i)\setminus\{i\}}}],$$

for each $i \in N$, where

$$P^\sigma(i) := \left\{r \in N | \sigma^{-1}(r) < \sigma^{-1}(i)\right\},$$

and $\sigma^{-1}(i)$ denotes the entrance number of player i. For grey size monotonic games $< N, c' >$, $c'(T) - c'(S)$ is defined for all $S, T \in 2^N$ with $S \subset T$ since $\left|c'(T)\right| = \left|c'\right|(T) \geq \left|c'\right|(S) = \left|c'(S)\right|$. Now, we notice that for each $c' \in SM\mathcal{G}G^N$ the grey marginal vectors $m^\sigma(c')$ are defined for each $\sigma \in \Pi(N)$, because the monotonicity of $|c|$ implies $\overline{A_{S\cup\{i\}}} - \underline{A_{S\cup\{i\}}} \geq \overline{A_S} - \underline{A_S}$, which can be rewritten as $\overline{A_{S\cup\{i\}}} - \overline{A_S} \geq \underline{A_{S\cup\{i\}}} - \underline{A_S}$. So, $c'(S \cup \{i\}) - c'(S)$ is defined for each $S \subset N$ and $i \notin S$. We notice that all the grey marginal vectors of a grey size monotonic game are efficient grey payoff vectors.

The *grey Shapley value* $\Phi' : SM\mathcal{G}G^N \to \mathcal{G}(\mathbb{R})^N$ is defined by

$$\Phi'(c') = \frac{1}{n!}\sum_{\sigma\in\Pi(N)} m^\sigma(c') \in [\frac{1}{n!}\sum_{\sigma\in\Pi(N)} \underline{m^\sigma(c')}, \frac{1}{n!}\sum_{\sigma\in\Pi(N)} \overline{m^\sigma(c')}]$$

for each $c' \in SM\mathcal{G}G^N$ [16].

3 Minimum Grey Cost Spanning Tree Games

In this section, we introduce minimum grey cost spanning tree situation and related grey games.

A *minimum grey cost spanning tree* (*mgcst*) *situation* is a situation where $N = \{1, 2, \ldots, n\}$ is a set of agents who are willing to be connected as cheaply as possible to a source (i.e., a supplier of a service) denoted by 0, based on an grey-valued weight (or cost) function.

For each $S \subseteq N$, we also use the notation $S_0 = S \cup \{0\}$, and the notation W for the *grey weight function*, i.e., a map which assigns to each edge $e \in E_{N_0}$ a grey number $W(e) \in \mathcal{G}(\mathbb{R}_+)$. The grey cost $W(e)$ of each edge $e \in E_{N_0}$ ($N_0 = N \cup \{0\}$) will be denoted by $W(e) \in [\underline{W}(e), \overline{W}(e)]$. No probability distribution is assumed for edge costs. We denote an *mgcst situation* with set of users N, source 0, and grey weight function W by $< N_0, W >$ (or simply W). Further, we denote by $\mathcal{GW}^{N_0}$ the set of all *mgcst situations* $< N_0, W >$ (or W) with node set N_0.

The *cost of a network* $\Gamma \subseteq E_{N_0}$ in an mgcst situation $W \in \mathcal{GW}^{N_0}$ is $W(\Gamma) = \sum_{e\in\Gamma} W(e)$. A network Γ is a *spanning network* on $S_0 = S \cup \{0\}$, with $S \subseteq N$, if for every $e \in \Gamma$ we have $e \in E_{S_0}$ and for every $i \in S$ there is a path in $< S_0, \Gamma >$ from i to the source. For any *mgcst* situation $W \in \mathcal{GW}^{N_0}$ it is possible to determine at least one *spanning tree* on N_0, i.e., a spanning network without cycles on N_0, *of minimum grey cost* (such a network is also called an *mgcst on* N_0 *in* W or, shorter, an *mgcst for* W). Note that the number of edges which form a spanning tree on N_0 is n. In the following, we will denote by $\mathcal{T}_{N_0}$ the set of all spanning trees for N_0 and by $\mathcal{M}^W_{N_0} \subseteq \mathcal{T}_{N_0}$ the set of all *micst* for N_0 in W, for each $W \in \mathcal{GW}^{N_0}$.

An mgcst *game* $< N, c_W >$ (or simply c_W) corresponding to an mgcst situation $c_W \in \mathcal{GW}^{N_0}$ is defined by

$$c_W(T) := \min\{c_W(\Gamma) | \Gamma \text{ is a spanning network on } T_0\}$$

for every $T \in 2^N \backslash \{\emptyset\}$, with the convention that $c_W(\emptyset) \in [0, 0]$. Also, a *grey solution* is a map $\mathscr{F} : \mathscr{GW}^{N_0} \to I(\mathbb{R})^N$ assigning to every *mgcst* situation $W \in \mathscr{GW}^{N_0}$ a unique allocation in $\mathscr{G}(\mathbb{R})^N$.

In a mgcst game the number $c_W(S)$ is the grey cost of a cheapest network, which connects every member of S with the source and which uses only edges in $S \cup \{0\}$. Always a grey cheapest network without cycles, i.e. a tree, can be chosen.

4 Grey Obligation Rules

In this section, we introduce a class of allocation rules for minimum grey cost spanning tree situations, namely the class of Obligation grey rules.

Given an element $\mathbf{a} = (a_1, \ldots, a_n) \in (E_{N_0})^n$, we denote by $\mathbf{a}_{|j}$ the restriction of $\mathbf{a}$ to the first j components, that is $\mathbf{a}_{|j} = (a_1, \ldots, a_j)$ for each $j \in N$. Further, for each $j \in N$, we denote by $\Pi(\mathbf{a}_{|j})$ the partition of N_0 defined as

$$\Pi(\mathbf{a}_{|j}) = \{T \subseteq N_0 | T \text{ is a connected component in } < N_0, \{a_1, \ldots, a_j\} >\}.$$

In the following, we will use the notation $\Pi(\mathbf{a}_{|0})$ to denote the singleton partition of N_0.

For each $\Gamma \in \mathscr{T}_{N_0}$ and each $c_W \in \mathscr{GW}^{N_0}$, we denote by $\mathbf{A}^{\Gamma,W} \subseteq (E_{N_0})^n$ the set of vectors $\mathbf{a} = (a_1, \ldots, a_n)$ of n distinct edges in Γ such that $c_W(a_1) \preccurlyeq \ldots \preccurlyeq c_W(a_n)$,

Note that $c_W(a_i)$ is monotonically increasing with respect to " $\preccurlyeq$ " :

$$\mathbf{A}^{\Gamma,W} = \{\mathbf{a} \in (\Gamma)^n | c_W(a_1) \preccurlyeq \ldots \preccurlyeq c_W(a_n), a_j \neq a_k \text{ for all } j, k \in N\}.$$

Consider $\Delta(N)$ to be an usual simplex on N, defined by $\Delta(N) = \{x \in \mathbb{R}^N_+ | \sum_{i \in N} x_i = 1\}$. The sub-simplex $\Delta(S)$ of $\Delta(N)$ given by $\Delta(S) = \{x \in \Delta(N) | \sum_{i \in S} x_i = 1\}$ is called the set of *obligation vectors* of S. An *obligation function* is a map $O : 2^N \setminus \{\emptyset\} \to \Delta(N)$ assigning to each $S \in 2^N \setminus \{\emptyset\}$ an obligation vector

$$O(S) \in \Delta(S)$$

in such a way that for each $S, T \in 2^N \setminus \{\emptyset\}$ with $S \subset T$ and for each $i \in S$ it holds

$$O_i(S) \geq O_i(T).$$

Such an obligation function O on $2^N \setminus \{\emptyset\}$ induces an *obligation map* $\hat{O} : \Theta(N_0) \to \mathbb{R}^N$ such that

$$\hat{O}_i(\theta) := \sum_{S \in \theta, 0 \notin S} O_i(S),$$

for each $i \in N$ and each $\theta \in \Theta(N_0)$; here, $\Theta(N_0)$ is the family of partitions of N_0 (for details see [18]).

Obligation maps are basic ingredients for grey obligation rules. Now, we introduce the notion of grey obligation rules.

Definition 1 Let $\hat{O}$ be an obligation map on $\Theta(N_0)$. The *grey obligation rule* $\phi^{\hat{O}} : \mathscr{GW}^{N_0} \to \mathscr{G}(\mathbb{R})^N$ is defined by

$$\phi^{\hat{O}}(c_W) \in \left[\sum_{j=1}^{n} \underline{c_W(a_j)}(\hat{O}(\Pi(\mathbf{a}_{|j-1})) - \hat{O}(\Pi(\mathbf{a}_{|j}))), \sum_{j=1}^{n} \overline{c_W(a_j)}(\hat{O}(\Pi(\mathbf{a}_{|j-1})) - \hat{O}(\Pi(\mathbf{a}_{|j}))) \right]$$

for each *mgcst* situation $c_W \in \mathscr{GW}^{N_0}$, each $\Gamma \in \mathscr{M}_{N_0}^W$ and $\mathbf{a} \in \mathbf{A}^{\Gamma,W}$, and where $\Pi(\mathbf{a}_{|j-1}))$ and $\Pi(\mathbf{a}_{|j})$, for each $j = 1, \ldots, n$, are partitions of the set N_0.

Example 1 We consider a mgcst situation $< N, c_W >$ with three agents denoted by 1, 2, and 3 and the source 0. As depicted in Fig. 1, to each edge $e \in E_{\{0,1,2,3\}}$ is assigned a grey number $c_W(e) \in I(\mathbb{R}_+)$ representing the grey cost of edge e. For instance, $c_W(0,1) \in [20, 24]$, $c_W(2,3) \in [10, 13]$, etc.. Now we compute the grey obligation rule $\phi^{\hat{O}}(W)$. In this mgcst situation, c_W, $\Gamma = \{(0,1), (1,2), (2,3)\} \in \mathscr{M}_{N_0}^{c_W}$ and

$$\mathbf{a} = (a_1, a_2, a_3) = ((2,3), (1,2), (0,1)) \in \mathbf{A}^{\Gamma, c_W}.$$

Then,

$$\phi_1^{\hat{O}}(c_W) = \frac{2}{3}[15, 20] + \frac{1}{3}[20, 24] = \left[\frac{50}{3}, \frac{64}{3}\right],$$
$$\phi_2^{\hat{O}}(c_W) = \frac{1}{2}[10, 13] + \frac{1}{6}[15, 20] + \frac{1}{3}[20, 24] = \left[\frac{85}{6}, \frac{107}{6}\right],$$
$$\phi_3^{\hat{O}}(c_W) = \frac{1}{2}[10, 13] + \frac{1}{6}[15, 20] + \frac{1}{3}[20, 24] = \left[\frac{85}{6}, \frac{107}{6}\right].$$

More clearly,

$$\phi^{\hat{O}}(c_W) \in [10, 13] \cdot (0, 1/2, 1/2) + [15, 20] \cdot (2/3, 1/6, 1/6) + [20, 24] \cdot (1/3, 1/3, 1/3) = \left(\left[\frac{50}{3}, \frac{64}{3}\right], \left[\frac{85}{6}, \frac{107}{6}\right], \left[\frac{85}{6}, \frac{107}{6}\right]\right)$$
$$= (\phi_1^{\hat{O}}(c_W), \phi_2^{\hat{O}}(c_W), \phi_3^{\hat{O}}(c_W))$$

Remark 1 It is obvious that if the cost of the edge connecting to source is the cheapest cost, then the grey obligation rule equals the grey Bird rule which is defined by [3].

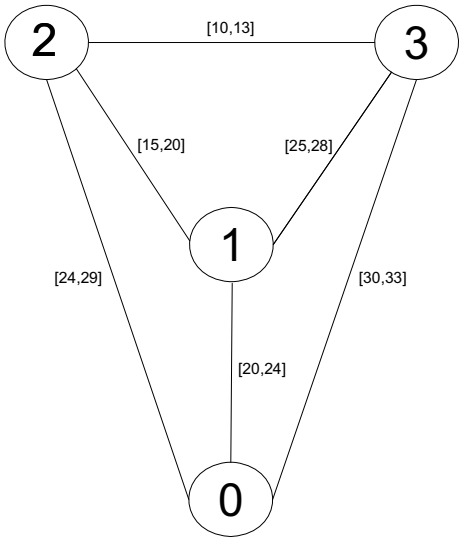


Fig. 1 A mgcst situation $< N_0, c_W >$

The classical Bird allocation is introduced in [4]. Next, we introduce the grey Bird allocation.

Definition 2 The grey Bird allocation is

$$\mathcal{GB}(N, \{0\}, A, c') \in (\left[\underline{\mathcal{GB}_1}, \overline{\mathcal{GB}_1}\right], \left[\underline{\mathcal{GB}_2}, \overline{\mathcal{GB}_2}\right], ..., \left[\underline{\mathcal{GB}_n}, \overline{\mathcal{GB}_n}\right]) \in \mathcal{G}(\mathbb{R})^N$$

with

$$\mathcal{GB}_k(N, \{0\}, A, c') = (c'(k, \widehat{b}(k)) \quad k = 1, 2, ..., n.$$

where,

$$\widehat{b}(k) = \underset{l \in T \cup \{0\}:(k,l) \in A}{\operatorname{argmin}} \underline{c'}(k, l),$$

the cheapest connection point of k in $T \cup \{0\}$ (for details see [3]).

Example 2 Figure 2 corresponding to mgcst situation $< N_0, c_W >$, the grey Bird allocation is

$$\mathcal{GB}(N, \{0\}, A, c_W) \in ([10, 13], [15, 20], [20, 24]) \text{ (see [3])}.$$

In this mgcst situation c_W, $\Gamma = \{(0, 1), (1, 2), (2, 3)\} \in \mathcal{M}_{N_0}^{c_W}$ and

$$\mathbf{a} = (a_1, a_2, a_3) = ((0, 1), (1, 2), (2, 3)) \in \mathbf{A}^{\Gamma, c_W}.$$

Then,

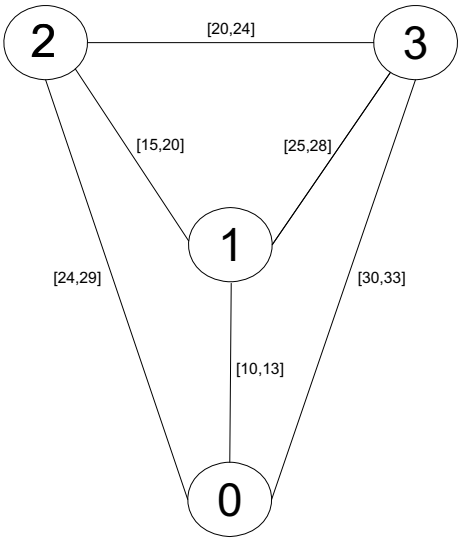


Fig. 2 A mgcst situation $< N_0, c_W >$

$$\phi_1^{\hat{O}}(c_W) = (1-0)[10,13] + (0-0)[15,20] + (0-0)[20,24] = [10,13],$$
$$\phi_2^{\hat{O}}(c_W) = (0-0)[10,13] + (1-0)[15,20] + (0-0)[20,24] = [15,20],$$
$$\phi_3^{\hat{O}}(c_W) = (0-0)[10,13] + (0-0)[15,20] + (1-0)[20,24] = [20,24].$$

It is clear that

$$\phi^{\hat{O}}(c_W) = \mathscr{G}B(N, \{0\}, A, c_W) \in ([10,13], [15,20], [20,24]).$$

5 Grey Cost Monotonicity and PMGAS

In this section, we introduce the notion of population monotonic grey allocation scheme (pmgas) for the game $< N, c' >$. We say that for a cost game c', a *scheme* $A = (A_{iS})_{i\in S, S\in 2^N\setminus\{\emptyset\}}$ with $A_{iS} \in \mathscr{G}(\mathbb{R})^N$ is a pmgas of c' if

(i) $\sum_{i\in S} A_{iS} = c'(S)$ for all $S \in 2^N \setminus \{\emptyset\}$, and
(ii) $A_{iS} \succcurlyeq A_{iT}$ for all $S, T \in 2^N$ and $i \in N$ with $i \in S \subset T$.

Moreover, we discuss some interesting grey monotonicity properties for the grey obligation rules and we show that if the game is concave, its (extended) grey Shapley value is a PMGAS.

First, we give the definition of grey cost monotonic solutions for mgcst situations.

Definition 3 A grey solution $\mathscr{F}$ is *a grey cost monotonic solution* if for all *mgcst* situations $c_W, c_{W'} \in \mathscr{GW}^{N_0}$ such that $c_W(e) \preccurlyeq c_{W'}(e)$ for each $e \in E_{N_0}$ it holds that $\mathscr{F}_i(c_W) \preccurlyeq \mathscr{F}_i(c_{W'})$ for each $i \in N$.

We prove in Theorem 1 that grey obligation rules are grey cost monotonic; the main step is the following lemma whose proof is straightforward.

Lemma 1 *Let $\hat{O}$ be an obligation map on $\Theta(N_0)$ and let $c_W \in \mathscr{GW}^{N_0}$. Let $\bar{e} \in E_{N_0}$ and let $h \succ c_W(\bar{e})$ be such that there is no $e \in E_{N_0}$ with $c_W(\bar{e}) \prec c_W(e) \prec h$. Define $\widetilde{c_W} \in \mathscr{GW}^{N_0}$ by $\widetilde{c_W}(e) := \widetilde{c_W}(e)$ if $e \in E_{N_0} \setminus \{\bar{e}\}$, and $\widetilde{c_W}(\bar{e}) = h$. Then, $\phi^{\hat{O}}(\widetilde{c_W}) \succcurlyeq \phi^{\hat{O}}(c_W)$.*

The proofs of the following theorems are straightforward (see [18]).

Theorem 1 *Grey obligation rules are grey cost monotonic.*

Theorem 2 *Let $\hat{O}$ be an obligation map on $\Theta(N_0)$ and let $\phi^{\hat{O}}$ the grey obligation rule with respect to $\hat{O}$, and $c_W \in \mathscr{GW}^{N_0}$. Then the table $[\phi^{\hat{O}_S}(c_{W|S_0})]_{S \in 2^N \setminus \{\emptyset\}}$ is a pmgas for the mgcst game $< N, c_W >$.*

Now, we give an example of grey cost monotonicity and pmgas.

Example 3 Consider again the mgcst situation $< N_0, c_W >$ as depicted in Fig. 1. Then, as the grey obligation rule $\phi^{\hat{O}}(c_W)$ previously introduced and the grey Shapley value, applied to each mgcst situation $< S_0, c_{W|S_0} >$, provides the following population monotonic grey allocation scheme:

$$[\phi^{\hat{O}_S}(c_{W|S_0})]_{S \in 2^N \setminus \{\emptyset\}} = \left\{ \begin{array}{|c|c|c|c|} \hline S & 1 & 2 & 3 \\ \hline \{1,2,3\} & [16\frac{4}{6}, 21\frac{2}{6}] & [14\frac{1}{6}, 17\frac{5}{6}] & [14\frac{1}{6}, 17\frac{5}{6}] \\ \{1,2\} & [17\frac{1}{2}, 22] & [17\frac{1}{2}, 22] & * \\ \{1,3\} & [20, 24] & * & [25, 28] \\ \{2,3\} & * & [17, 21] & [17, 21] \\ \{1\} & [20, 24] & * & * \\ \{2\} & * & [24, 29] & * \\ \{3\} & * & * & [30, 33] \\ \hline \end{array} \right. ,$$

Now, we show that if the game is concave, its (extended) grey Shapley value is a PMGAS.

A grey cost game $c' \in SM\mathscr{G}G^N$ is called concave iff

$$c'(S \cup T) + c'(S \cap T) \preccurlyeq c'(S) + c'(T)$$

for all $S, T \in 2^N$.

Proposition 1 *If c' is a concave game, then every extended vector of marginal grey contributions is a PMAS of c'.*

Proof Let c' be a concave game and $\sigma \in \Pi(N)$. Consider the extended vector of marginal contributions m^σ. Pick an arbitrary $S \in 2^N$ and rank all players $i \in S$ in increasing order of $\sigma(i)$. Let $i, j \in S$ be two players such that j immediately follows i (i.e., $\sigma(i) < \sigma(j)$ and there is no $k \in S$ such that $\sigma(i) < \sigma(k) < \sigma(j)$). Observe that

$$m_i^\sigma(c') = c'(P^\sigma(i)) - c'(P^\sigma(i) \setminus \{i\})$$

and

$$\begin{aligned} m_j^\sigma(c') &= c'(P^\sigma(j)) - c'(P^\sigma(j) \setminus \{j\}) \\ &= c'(P^\sigma(j)) - c'(P^\sigma(i)) \end{aligned}$$

Therefore

$$m_i^\sigma(c') + m_j^\sigma(c') = c'(P^\sigma(j)) - -c'(P^\sigma(i) \setminus \{i\}).$$

Repating this argument leads to $\sum_{i \in S} m_i^\sigma(c') = c'(S)$, which establishes the feasibility of m^σ. As for the monotonicity property, note that $i \in S \subseteq T \subseteq N$, then $S \cap P^\sigma(i) \subseteq T \cap P^\sigma(i)$ for all $i \in N$. Hence by concavity, $m_i^\sigma \succcurlyeq m_j^\sigma$. This completes the proof.

Corollary 1 *If c' is a concave game, then its extended grey Shapley value is a PMAS of c'.*

Proof The extended grey Shapley value is the extended vector of marginal grey contributions. By using Proposition 11, the proof is completed.

Example 4 Consider again the mgcst situation $< N_0, c_W >$ as depicted in Fig. 1. Then, the grey Shapley value, applied to each mgcst situation $< S_0, c_{W|S_0} >$, provides the following population monotonic grey allocation scheme:

$$[\Phi'(c_{W|S_0})]_{S \in 2^N \setminus \{\emptyset\}} = \left\{ \begin{array}{c|ccc} S & 1 & 2 & 3 \\ \hline \{1,2,3\} & [14\frac{2}{6}, 18\frac{4}{6}] & [11\frac{1}{6}, 16\frac{1}{6}] & [19\frac{1}{6}, 22\frac{1}{6}] \\ \{1,2\} & [15\frac{1}{2}, 19\frac{1}{2}] & [19\frac{1}{2}, 24\frac{1}{2}] & * \\ \{1,3\} & [17\frac{1}{2}, 21\frac{1}{2}] & * & [27\frac{1}{2}, 30\frac{1}{2}] \\ \{2,3\} & * & [14, 19] & [20, 23] \\ \{1\} & [20, 24] & * & * \\ \{2\} & * & [24, 29] & * \\ \{3\} & * & * & [30, 33] \end{array} \right.$$

6 Conclusion

This paper considers the class of grey obligation rules and studies their grey cost monotonicity properties. The grey obligation rules are grey cost monotonic and induce a pmgas. There are two important results of this study. One of them is, as

already stated in Remark 1, if the cost of the edges connecting to source is the cheapest cost, then the grey obligation rule equals the grey Bird rule which is defined by [3]. The other is that obligation rules have nice monotonicity properties: cost monotonicity and population monotonicity.

References

1. Alparslan Gök, S.Z., Palanci, O., Yücesan, Z.: Peer group situations and games with grey uncertainty. Handbook of Research on Emergent Applications of Optimization Algorithms, Chapter 11, 265–278, IGI Global, USA (2018)
2. Alparslan Gök, S.Z., Branzei, O., Branzei, R., Tijs, S.: Set-valued solution concepts using interval-type payoffs for interval games. J. Math. Econ. **47**, 621–626 (2011)
3. Alparslan Gök, S.Z., Palancı, O., Olgun, M.O.: Cooperative interval games: mountain situations with interval data. J. Comput. Appl. Math. **259**, 622–632 (2014)
4. Bird, C.G.: On cost allocation for a spanning tree: a game theoretic approach. Networks **6**, 335–350 (1976)
5. Borm, P., Hamers, H., Hendrickx, R.: Operations research games: a survey. TOP **9**, 139–216 (2001)
6. Branzei, R., Branzei, O., Alparslan Gök, S.Z., Tijs, S.: Cooperative interval games: a survey. Central Europ. J. Oper. Res. (CEJOR) **18**(3), 397–411 (2010)
7. Deng, J.: Control problems of grey systems. Syst. Control Lett. **5**, 288–294 (1982)
8. Deng, J.: Grey System Fundamental Method. Huazhong University of Science and Technology, Wuhan, China (1985)
9. Diestel, R.: Graph Theory. Springer (2000)
10. Ekici, M., Palanci, O., Alparslan Gök, S.Z.: The grey Shapley value: an axiomatization. IOP Conference Series: Materials Science and Engineering **300**, 1–8 (2018)
11. Fang, Z., Liu, S.F.: Grey matrix game model based on pure strategy. J. Nanjing Univ. Aeronaut. Astronaut. **35**(4), 441–445 (2003)
12. Kose, E., Forest, J.Y.L.: N-person grey game. Kybernetes **44**(2), 271–282 (2015)
13. Liu, S., Lin, Y.: Grey Information: Theory and Practical Applications. Springer, Germany (2006)
14. Moore, R.: Methods and applications of interval analysis. SIAM Stud. Appl. Math. (1995)
15. Moulin, H.: Cores and large cores when population varies. Int. J. Game Theory **19**(2), 219–232 (1990)
16. Palanci, O., Alparslan Gök, S.Z., Ergun, S., Weber, G.-W.: Cooperative grey games and the grey Shapley value. Optimization **64**(8), 1657–1688 (2015)
17. Suijs, J.: Cost allocation in spanning network enterprises with stochastic connection costs. Games Econ. Behav. **42**, 156–171 (2003)
18. Tijs, S., Branzei, R., Moretti, S., Norde, H.: Obligation rules for minimum cost spanning tree situations and their monotonicity properties. Europ. J. Oper. Res. **175**, 121–34 (2006)

Robustness Checks in Composite Indices: A Responsible Approach

Juan Diego Paredes-Gázquez, Eva Pardo, and José Miguel Rodríguez-Fernández

Abstract Scholars and practitioners are increasingly using composite indices. This is especially true in Corporate Social Responsibility (CSR) research, where composite indices are becoming a common measure. However, the robustness of these indices is not always tested or, if tested, it is mainly limited to the weighting scheme of the indicators of the index. The objective of this chapter is to highlight the importance of robustness checks in CSR composite indexes in order they can properly summarize CSR performance. We have constructed a CSR composite index, testing its robustness by using uncertainty and sensitivity analyses. The results evidence that missing data and normalization of indicators can highly affect the internal structure of CSR composite indices, making them unreliable sometimes. These results suggest that missing data treatment and the normalization of indicators deserve more attention by the CSR research community. Consequently, the robustness of CSR composite indexes should always be tested, especially if they include missing data.

Keywords Composite indices · Corporate social responsibility · Robustness · Missing data · Normalization · Weighting scheme

J. D. Paredes-Gázquez (✉) · E. Pardo
UNED, Facultad CC. Económicas y Empresariales, Paseo Senda del Rey 11, 28040 Madrid, Spain
e-mail: juandiegoparedes@cee.uned.es

E. Pardo
e-mail: epardo@cee.uned.es

J. M. Rodríguez-Fernández
Universidad de Valladolid, Facultad CC. Económicas y Empresariales, Avda del Valle Esgueva 6, 47011 Valladolid, Spain
e-mail: jmrodrig@eco.uva.es

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_17

1 Introduction

Composite indices (hereafter, indices):

> are formed when individual indicators are compiled into a single measure on the basis of an underlying model

(OECD [40], p. 13). They summarize the information of those individual indicators and provide support to decision-makers. Indices often measure multidimensional concepts, such as Corporate Social Responsibility (CSR) and sustainability performance. This chapter focuses on those indices designed specifically for measuring CSR or sustainability (hereafter, CSR indices).

There is not a unique and proper way to construct indices. The complexity of their construction process entails subjective choices that researchers should make carefully. Among these choices, Saisana et al. [47] identifies: selection of subindicators, data selection, data editing, data normalization, weighting scheme, weights' values and index indicators formula. Most of the choices deal with methodological issues. An inadequate combination of choices can lead to the misrepresentation or the manipulation of the index (OECD [40], p. 15). In order to avoid these problems, the construction process should be transparent and the robustness of the index should always be checked. Nevertheless, evidence suggests that the construction processes of indices might be not transparent. What is more, in many cases there is no evidence of robustness checks (Singh et al. [57]; Saisana et al. [47]).

The nature of this chapter is methodological. The objective is to show the importance of transparency and robustness checks in the construction of CSR indices. This research field was chosen because it is common to measure CSR and related concepts with these indices (Capelle-Blancard and Petit [6]; Escrig-Olmedo et al. [14]). We construct a CSR index for illustrative purpose, evidencing how missing data and normalization could threat the robustness of the index.

The chapter is structured as follows. Section 2 provides a brief literature review about CSR, while Sect. 3 describes the construction of the CSR index. The robustness of this index is tested in Sect. 4. We focus our analyses on three of the main choices researchers have to adopt in the construction process: missing data treatment, data normalization and weighting scheme. Finally, Sect. 5 provides the conclusions.

2 Literature Review

Sometimes CSR and sustainability are regarded as interchangeable terms. These concepts are different in some aspects, but they share a common basis: both aim to balance economic prosperity, social integrity, and environmental responsibility (Montiel [33]).[1]

[1] For a discussion of similarities and differences between sustainability and CSR see Ghobadian et al. [18], Hahn et al. [22], Baumgartner [1] and Montiel [33].

The European Commission defines CSR as:

> the responsibility of enterprises for their impacts on society, (European Commission [13], p. 6).

and it involves social, environmental, economic and even governance implications (Hediger [23]; Jamali et al. [26]; Rodríguez-Fernández [46]). The definition of sustainability has evolved from a concept limited to the environmental dimension of business, to a broader one which also includes economic and social dimensions (Montiel [33]).

As the chapter does not intend to enter in a theoretical discussion on CSR and sustainability, the two terms are considered as synonyms. This interpretation relies on previous researches that state that: (1) CSR and sustainability share a common basis (Montiel [33]), (2) CSR may be understood as the implementation of business Sustainability (Escrig-Olmedo et al. [14]) and (3) both concepts share metrics and indicators, for example, the Global Reporting Initiative provides Sustainability Reporting guidelines which are a reference is CSR research (Daub [10]). Other related concepts are Corporate Social Performance (CSP), corporate citizenship, corporate sustainability and shared value, among others (Ghobadian et al. [18]; Hahn et al. [22]).

The measurement of CSR and sustainability requires multiple indicators in order to reflect the performance in so diverse areas as environment, governance and social impacts. This is the reason why CSP is regarded as the measure of the social responsibilities defined by Carroll [7]. The identification of these sets of indicators has been done by scholars focusing on CSR measurement, i.e. De la Cuesta et al. [11], Zhao et al. [67] and O'Connor and Spangenberg [39], as well as by others focused on measuring sustainability such as Sardain et al. [53], Fernández-Sánchez and Rodríguez-Lopez [15].

National and international organizations have also proposed indicators of CSR and sustainability issues. Perhaps the most popular organization is the Global Reporting Initiative (GRI), an international independent standards organization. GRI launched the first Sustainability Reporting guidelines in 2000. The last version of the guidelines, published in 2013, proposed more than 90 indicators (GRI [20]). In 2008, the United Nations Conference on Trade and Development (UNCTAD) also proposed a set of 16 indicators divided into six groups for measuring CSR (UNCTAD [61]). These indicators, listed in Table 1, originate from a consultation process endorsed by international organizations. Companies may voluntarily adopt some or all of these indicators in order to report their CSR policies, practices and outcomes. Although some of these initiatives are focused on transparency, the set of indicators they provide is a basis for measuring CSR and sustainability performance (Gallego [17]; Vigneau et al. [63]).

The huge number and the diversity of possible indicators make it difficult to get a whole picture of the status of a company's CSR performance. This is the reason why CSR indicators are often summarized into a single index. These indices receive different names depending on their theoretical underpinning. One name widely accepted is Corporate Social Performance (CSP) (Pierick 2004). Table 2 shows studies that use indices for measuring CSR and related terms.

Table 1 UNCTAD CSR indicators

Group	UNCTAD Code[a]	Indicator
Trade, investment and linkages	TRA1	1. Total revenues
	TRA2	2. Value of imports versus exports
	TRA3	3. Total new investments
	TRA4	4. Local purchasing
Employment creation and labour practices	EM1	5. Total workforce with breakdown by employment type, employment contract and gender
	EM2	6. Employee wages and benefits with breakdown by employment type and gender
	EM3	7. Total number and rate of employee turnover broken down by gender
	EM4	8. Percentage of employees covered by collective agreements
Technology and human resource development	TEC1	9. Expenditure on research and development
	TEC2	10. Average hours of training per year per employee broken down by employee category
	TEC3	11. Expenditure on employee training per year per employee broken down by employee category
Health and safety	HS1	12. Cost of employee health and safety
	HS2	13. Work days lost due to occupational accidents, injuries and illness
Government and community contributions	GOV1	14. Payments to Government
	GOV2	15. Voluntary contributions to civil society
Corruption	COR1	16. Number of convictions for violations of corruption related laws or regulations and amount of fines paid/payable

[a] Codes proposed by the authors
Source UNCTACD (2008)

Some of the analysis on CSR performance take the index from an external source (i.e. sustainability rating agencies), while others construct their own index. Whatever the case, the construction process is not always transparent, although transparency is a key issue on indices construction. Indeed, some studies have noted the theoretical and methodological weaknesses of aggregate measures of CSR or CSP (Wood [66]; Scalet and Kelly [54]).

Table 2 Some studies including composites of CSR and related terms[a]

Term	Studies
CSR	Nollet et al. [38], Lima-Crisóstomo et al. [29], Choi et al. [9], Gjølberg [19]
Sustainability	Morse [34], Bondarchik et al. [3], Singh et al. [57], Pulido-Fernández and Sánchez-Rivero [45]
CSP[b]	Wang et al. [64], Ioannou and Serafeim [25], Chen and Delmas [8]
Others	Spangenberg [60], Searcy [56]

[a]This literature survey should be considered illustrative rather than exhaustive
[b]CSP has a solid theoretical background (Wood [66]; Pierick et al. [43]). The authors cited in the table conceptualize a narrow CSP or outcomes-based measure of CSR.
Source own elaboration

Researchers have paid much attention to the effects of the weighting scheme on indices (Mikulić et al. [30]; Salvati and Zitti [52]; Munda and Nardo [36]). However, the debates about normalization and missing data have received less attention. While normalization is often a pure methodological issue, the availability of data is a concern in any CSR research (Sardain et al. [53]; Gray [21]; Levett [28]). These two aspects, normalization and missing data, deserve a deeper analysis.

3 Index Construction

This section describes the construction of the CSR index of our study. The construction follows the recommendations suggested by the OECD in its Handbook on Constructing Index indicators (OECD [40], p. 20). The handbook proposes a process of eight steps: 1. Theoretical framework selection; 2. Data selection; 3. Imputation of missing data; 4. Multivariate analysis; 5. Normalization; 6. Weighting and aggregation; 7. Uncertainty and sensitivity analysis; 8. Back to the data. Since our study focuses on robustness checks based on uncertainty and sensitivity analysis, we will finish our analysis on step 7.

GRI [20] is a popular source of CSR indicators. The initiative includes more than 90 indicators covering all the dimensions of CSR. However, the more indicators in an index, the more variability, and hence the less robust could it be. Moreover, some studies evidence that GRI indicators may provide a biased and positive view of the CSR achievements of a company (Boiral [2]; Moneva et al. [32]). In order to construct the index, we chose the CSR indicators proposed by UNCTAD [61] because it is a balanced initiative: it has a reduced number of indicators while considering the most relevant aspects of CSR. Our index includes twelve of the sixteen CSR indicators proposed by UNCTAD [61], since we excluded those indicators indirectly related to

Table 3 Link between UNCTAD and Asset4 indicators

UNCTAD Code	Asset 4 indicator	Sign
EM1	Number of both full and part time employees of the company	Positive
EM2	Average salaries and benefit in US dollars (Salaries and Benefits (US dollars) /Total Number of Employees)	Positive
EM3	Percentage of employee turnover	Negative
EM4	Percentage of employees represented by independent trade union organizations or covered by collective bargaining agreements	Positive
TEC1	Research and development costs divided by net sales or revenue	Positive
TEC2	Average hours of training per year per employee	Positive
TEC3	Training costs per employee in US dollars	Positive
HEA1	Total number of injuries and fatalities including no-lost-time injuries relative to one million hours worked	Negative
HEA2	Total lost days at work divided by total working days. (Refers to an employee absent from work because of incapacity of any kin)	Negative
GOV1	The Effective Tax Rate is defined as Income Taxes (Credit) divided by Income Before Taxes and expressed as a percentage	Positive
GOV2	Total amount of all donations divided by net sales or revenue	Positive
COR1	All real or estimated penalties, fines from lost court cases, settlements or cases not yet settled regarding controversies linked to business ethics in general, political contributions or bribery and corruption, price-fixing or anti-competitive behaviour, tax fraud, parallel imports or money laundering in US dollars	Negative

Source own elaboration

CSR or those seldom reported by companies (indicators TRA1, TRA2, TRA3 and TRA4).

The source of data is Asset4, a Thomson Reuters non-financial information database that provides both raw data and CSR indices. Empirical studies such as those by Miras-Rodríguez et al. [31] or Ioannou and Serafeim [25] use Asset4 database as a source of CSR or CSP indicators. Table 3 shows the description of the Asset4 indicators included in the index, as well as the link between UNCTAD and Asset4 indicators. The sign indicates how the indicator contributes to the index, i.e. if a higher value of the indicator increases (positive sign) or decreases (negative sign) the compromise of the company with CSR practices. The data requested are for year 2014, the last year available at that moment.

The sample includes 48 companies, which are all the electric utilities available in the Asset4 database for the year 2014 (Sector classification according to Standard Industrial Classification, code 4911). We choose electric utilities because this sector faces many CSR risks on different CSR dimensions (Wilde-Ramsing [65]). Focusing on a specific sector eliminates any possible bias among sectors due to different CSR impacts.

The index construction process we chose, which combines standardization as normalization procedure, principal component analysis as weighting method and

linear aggregation, is a quite common procedure for constructing indices (Cano-Orellana and Delgado-Cabeza [5]; Salvati and Carlucci [51]; Hosseini and Kaneko [24]). Prior to data normalization and the selection of the weighting scheme, the outliers were checked. After a visual inspection of quantile-quantile plots (QQ-plots), six outliers were removed from the data set. Indicators with negative sign were reversed. Missing data were handled through multiple imputation, using the Markov Chain Monte Carlo (MCMC) method proposed by Van Buuren [62].

Prior to set the indicators' weighs factors, sample adequacy should also be tested. The Barlett test is significant (Bartlett χ^2_{66}= 100.89), while the Kaiser-Meyer-Olkin (KMO) statistic is 0.56. The value of the KMO statistic is low but acceptable (Kaiser [27]), indicating that data can be summarized through principal component analysis.

A varimax rotated principal component analysis provided the initial weights of the indicators. The estimation of the final weights followed the guidelines proposed by Nicoletti et al. [37]. The aggregation scheme was linear and compensatory. The companies of the sample are ranked according to their final score in the index (Table 4). This score was the reference for the robustness checks of the index.

4 Robustness Checks

The robustness of new indices should be always checked (OECD [40], p. 117). In this section, we test if the index is robust or not. In order to do that, firstly we should define the input factors. These factors are necessary to perform uncertainty and sensitivity analyses, which are the core of the robustness checks. Secondly, we show the results of these robustness checks.

4.1 Input Factor Definition

We test the robustness of the ranks of the companies, assessing how the ranks would have changed if the construction of the index had been different. Uncertainty and sensitivity analyses allow us to analyze these changes (Saltelli et al. [47, 49]; Saltelli et al. [50]). Uncertainty analysis evidences how the score or the rank of the companies changes across the alternative indices. The more the scores or the ranks change, the less robust is the index. Sensitivity analysis allows us to identify the input factor causing the rank changes. It is based on the computation of sensitivity indices.

Uncertainty and sensitivity analyses require different combinations of input factors. These factors determine how the index is calculated, triggering the subjective choices adopted in the construction of new alternative indices. In our analysis, we define three input factors: missing data, normalization and weighting scheme. Each of the input factors has three different triggers. As this study does not aim to explore all the methodological options possible, the triggers considered represent some recurrent choices adopted by scholars when constructing indices.

Table 4 Rank of the companies

Company	Rank	Company	Rank
E On	1	Nisource	25
Alstom	2	Elia System Operator	26
Cameco	3	Transcanada	27
Datang International P.G.	4	Hera	28
Norsk Hydro	5	Energy Development	29
Federal Grid Company U.E.S.	6	P.G.C. India	30
Xcel Energy	7	Iberdrola	31
Enel	8	Jindal Steel and Power	32
A2A	9	Hess	33
CLP Holdings	10	Pinnacle West Capital	34
Verbund	11	Tenaga Nasional	35
Aboitiz Equity Ventures	12	Drax Group	36
Southern	13	Manila Electric	37
Centrica	14	China Resources P.H.	38
China Longyuan P.G.	15	NRG Energy	39
Duke Energy	16	SSE	40
Ameren	17	Sempra Energy	41
Acea	18	Transalta	42
Fortum	19	Origin Energy (ex Boral)	43
JSW Steel	20	Reliance Infrastructure	44
National Grid	21	Abengoa	45
RWE	22	Exelon	46
Dominion Resources	23	Tata Power	47
Huaneng P.I.	24	Sembcorp Industries	48

Source own elaboration from UNCTAD [61] indicators and Asset4 data

For the first input factor, missing data, we consider three common approaches. The first one is to replace missing data by zero. The second one is to replace them by the mean. The last one is the multiple imputation (MCMC method) proposed by Van Buuren [62]. The second input factor concerns data normalization. We work with three of the main normalization techniques among those summarized by Munda ([35], p. 88): distance from the mean, distance from the worst performer and standard deviation from the mean. The third input factor refers to the weighting scheme. Different methods are also available for setting the contribution of the indicators to the index: equal-weighting, opinion-based methods, distance-based methods and multivariate analysis, among others (Domínguez-Serrano et al. [12]; OECD [40], p.

Table 5 Input factors

Input factor	Input factor distribution	Triggers	
Missing data	$(X_1) \sim U$ (0-1)	$X_1 \leq 0.33$	Replace by cero
		$0.33 < X_1 \leq 0.66$	Replace by mean
		$0.66 < X_1 < 1$	Multiple imputation (MCMC)
Normalization	$(X_2) \sim U$ (0-1)	$X_2 \leq 0.33$	Distance from the mean
		$0.33 < X_2 \leq 0.66$	Distance from the worst performer
		$0.66 < X_2 < 1$	Standard deviation from the mean
Weighting	$(X_3) \sim U$ (0-1)	$X_3 \leq 0.33$	Equal weighting
		$0.33 < X_3 \leq 0.66$	Rotated multiple factor analysis
		$0.66 < X_3 < 1$	Rotated principal component analysis

Source own elaboration

89). In our analysis we consider equal weighting, varimax rotated multiple factor analysis and varimax rotated principal component analysis.

Table 5 summarizes the input factors defined for the analysis. The distribution of each input factor determines its value, while the trigger sets how the index is constructed. For example, if the value of the input factor $X_1 = 0.430$, missing data are replaced by the mean. A combination of the three input factors determines the construction of the alternative indices. For example, a combination of input factors where $X_1 = 0.254$, $X_2 = 0.986$ and $X_3 = 0.234$ gives an alternative index with missing data replace by zero ($X_1 \leq 0.33$), standard deviation from the mean normalization ($0.66 < X_2 < 1$) and equal weitghing scheme ($X_3 \leq 0.33$). So forth for each combination of input factors.

The combinations of input factors were created using a quasi-random sampling procedure (Sobol [58, 59]; Saltelli [48]). This sampling was generated using SimLab (Saltelli et al. [50]; version 4.0.). The sample size was n = 256 for each input factor. We get a total of 2n (k +1) = 2048 simulations, where k = 3 is the number of input factors. Different combinations of input factors determine the construction of alternative indices. For each company it is calculated the difference between the rank in the original index (the reference) and the rank in the alternative index. Uncertainty and sensitivity analyses were performed using the combinations of input factors as input and the difference between ranks as output.

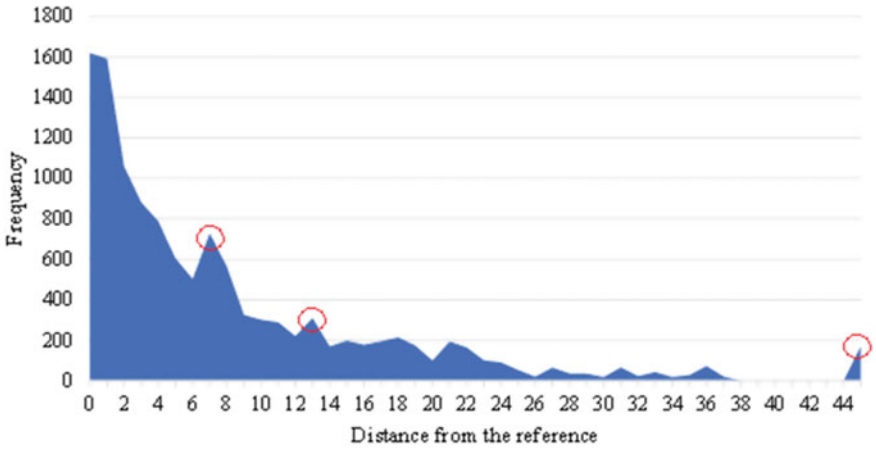


Fig. 1 Frequency of the distances from the reference. *Source* own elaboration

4.2 Results

Figures 1 and 2 summarize the results of the uncertainty analysis. Figure 1 shows the distance from the reference of the alternative indices across simulations for all the companies. The figure provides a general picture of the rank's changes. A distance of zero from the reference means that the rank order of the alternative index and the rank order of the reference are equal. The more the distance, the more the rank order changes across simulations. Common values for distance from the reference are zero, one or two positions. However, there are abnormal peaks in the distances of seven, thirteen and forty-five positions from the reference (circles in Fig. 1). These peaks indicate that the rank change is affecting some companies specially.

Figure 2 summarizes the distance from the reference across simulations for each company. The length of the lines represents the range of the distances. The figure provides company-focused picture of the rank changes. The greater the range, the more the rank of a company changes across simulations. The dark boxes represent the median. For ten companies, the median is above ten. This indicates that, in at least the 50% of simulations, the rank change of these companies is above ten positions or more. There are eight companies with a range greater that thirty. The median of two of them is below five, while the range is 45. This means that, while a 50% of the distance is between zero and five, the other 50% is between five and 45. These results evidence the index is not robust for many companies.

We estimate both first-order effect and total effect indices. The first-order effect index indicates the relative importance of an individual input factor in driving the total uncertainty or variance of the output (García Aguña and Kovacevic [16]). The total effect index includes the first-order effect and the uncertainty of the interaction of the input factor with others input factors. Saltelli et al. [50] describe the calculation methods of these indices, based on decomposition of variance techniques.

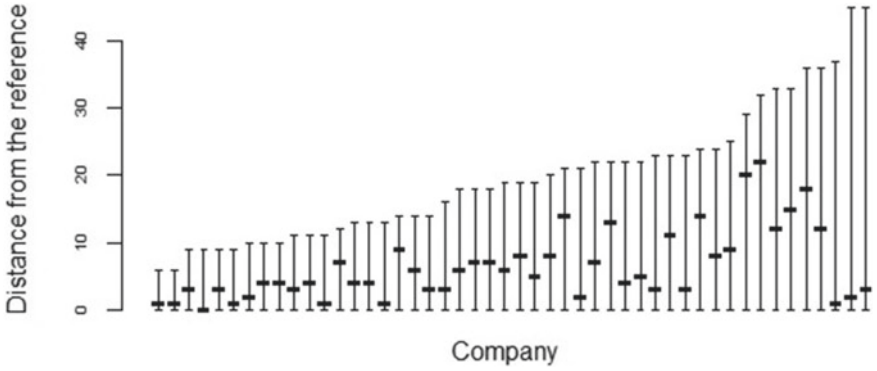


Fig. 2 Distances from the reference for each company. *Source* own elaboration

Table 6 Sobol' sensitivity measures

Input factor	First-order effect (S_i)	Total effect (S_{Ti})	S_{Ti}–S_i
Missing data	0.4289	0.4305	0.0016
Normalization	0.5866	0.6018	0.0152
Weighting	0.0142	0.0315	0.0173

Source own elaboration

Table 6 contains the first-order and total effect Sobol' sensitivity estimated measures. Saisana et al. [47] consider that an input factor is important if it explains more than 1/k of the output variance. This threshold is 1/3 = 0.333. First order effects of our analysis show that normalization is the most influential input factor, followed by missing data. The values of these indices are above the threshold 1/k, so they are influential input factors. These results reveal that normalization and missing data highly affects the internal structure of the indices, undermining their robustness. The effects of the weighting input factor are residual compared to the effects of normalization and missing data.

In order to test the results of Table 5, we computed the sensitivity indices using different sample sizes. Additionally, we estimated a different sensitivity analysis using the effective algorithm for computing global sensitivity indices (EASI) (Plischke [44]). EASI estimates first order sensitivity indices from given data using Fast Fourier Transformations. Figure 3 shows the new indices. The results are in line with the previous Sobol' sensitivity measures. A striking result is that, when sample size is 512, the input factor of missing data is just below the threshold (0.333); nevertheless its value is still important.

These results lead us to agree with Cabello (2014), who states that nowadays it is impossible to obtain a completely objective measurement of sustainability. This unreliability is due to aspects such as missing data, normalization and weighting scheme. In order to reduce subjectivity, researchers should construct transparent indices, providing as much information as they can on the construction process of

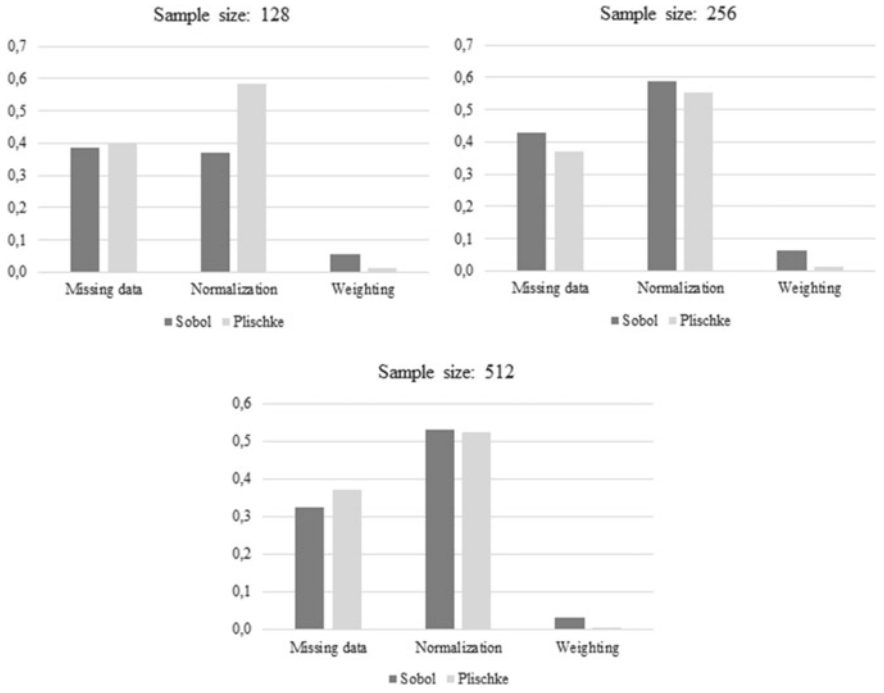


Fig. 3 First order sensitivity indices for three sample sizes *Source* own elaboration

the indexes. Most of the scholars are focusing their efforts on those issues related to the weighting scheme. However, the results of this research point that the uncertainty in the rank change is mainly due to normalization and missing data factors.

How can we reduce uncertainty? Concerning normalization, Paruolo et al. [42] suggest that the normalization method should be chosen attending to the distribution of the data. Concerning the effect of missing data, it is a more complex issue. The source of the data is key in this sense, since a complete source would minimize or remove the missing data problem. If the source of data has missing data, it is necessary to analyze their pattern, mechanism and distribution in order to adopt an adequate solution (Paredes-Gázquez et al. [41]). In any case, uncertainty will always exist, but the researchers should minimize it by adopting the methodological choices the data require.

5 Conclusion

This research highlights why it is important to check the robustness of CSR indices. In order to do so, first we have constructed a CSR index based on indicators promoted by

the UNCTAD. Second, we have identified three design choices affecting the internal structure of the index: missing data, normalization and weighting scheme. Finally, we have tested the robustness of the index.

By now, the literature about indices has mainly focused on the effect of the weighting scheme on the construction and interpretation of indices. Specifically, CSR literature reveals that aggregated CSR and CSP measures may have theoretical and methodological weakness (Capelle-Blancard and Petit [6]; Escrig-Olmedo et al. [14]; Wood [66]; Scalet and Kelly [54]). The robustness checks of our analysis allow us to identify the reasons behind these methodological weaknesses, evidencing that missing data and normalization treatment affects significantly the internal structure of the index and, therefore, its results. The missing data problem is especially important, since lack of data is common in CSR research (Paredes-Gázquez et al. [41]). The results of our analysis suggest that more attention should be paid to those issues related to the imputation of missing data and the normalization of indicators. This conclusion raises awareness of the importance of adopting a responsible approach when constructing CSR indices.

This research also shows that CSR researchers should take a keen effort on transparency when constructing indices. Transparency allows researchers to detect deficiencies in the construction of the index and then, make it possible to correct them. Although the present study is limited to just one economic sector and by the fact that the index only has an illustrative purpose, it evidences the robustness problems that indices may have.

Future work will involve the application of robustness checks to CSR indices, especially in those indices that include indicators with missing data. Future studies could also include additional input factors such as aggregation method, or new methods in the input factors we defined, thus covering additional potential sources of uncertainty.

References

1. Baumgartner, R.J.: Managing corporate sustainability and CSR: a conceptual framework combining values, strategies and instruments contributing to sustainable development. Corpor. Soc. Responsib. Environ. Manag. **21**(5), 258–271 (2014)
2. Boiral, O.: Sustainability reports as simulacra? a counter-account of A and A GRI reports. Account. Audit. Account. J. **6**(7), 1036–1071 (2013)
3. Bondarchik, J., Jabłońska-Sabuka, M., Linnanen, L., et al.: Improving the objectivity of sustainability indices by a novel approach for combining contrasting effects: happy planet index revisited. Ecol. Indicators **69**, 400–406 (2016). https://doi.org/10.1016/j.ecolind.2016.04.044
4. Cabello, J., Navarro, E., Prieto, F., et al.: Multicriteria development of synthetic indicators of the environmental profile of the Spanish regions. Ecol. Ind. **39**, 10–23 (2014)
5. Cano-Orellana, A., Delgado-Cabeza, M.: Local ecological footprint using principal component analysis: a case study of localities in Andalusia (Spain). Ecolog. Ind. **57**, 573–579 (2015)
6. Capelle-Blancard, G., Petit, A.: The weighting of CSR dimensions: one size does not fit all. Bus. Soc. **56**(6), 919–943 (2015)
7. Carroll, A.B.: A three-dimensional conceptual model of corporate performance. Acad. Manag. Rev. **4**(4), 497–505 (1979)

8. Chen, C., Delmas, M.: Measuring corporate social performance: an efficiency perspective. Product. Oper. Manag. **20**(6), 789–804 (2011)
9. Choi, J., Kwak, Y., Choe, C.: Corporate social responsibility and corporate financial performance: Evidence from Korea. Austr. J. Manag. **35**(3), 291–311 (2010)
10. Daub, C.-H.: Assessing the quality of sustainability reporting: an alternative methodological approach. J. Cleaner Product. **15**, 75–85 (2007)
11. de la Cuesta González, M., Pardo Herrasti, E., Paredes Gázquez, J.D.: Identificación de indicadores relevantes del desempeño RSE mediante la utilización de técnicas multicriterio. Innovar **25**(55), 75–88 (2015)
12. Domínguez-Serrano, Blancas-Peral, F.J., Guerrero-Casas, F.M., González-Lozano, M.: Una revisión crítica para la construcción de indicadores sintéticos. Revista de métodos cuantitativos para la economía y la empresa **11**, 41–70 (2011)
13. European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A renewed EU strategy 2011–2014 for Corporate Social Responsibility. Brussels: European Commission (2011)
14. Escrig-Olmedo, E., Munoz-Torres, M.J., Fernández-Izquierdo, M.A., Rivera-Lirio, J.M.: Lights and shadows on sustainability rating scoring. Rev. Manager. Sci. **8**, 559 (2014)
15. Fernández-Sánchez, G., Rodríguez-López, F.: A methodology to identify sustainability indicators in construction project management?Application to infrastructure projects in Spain. Ecol. Ind. **10**(6), 1193–1201 (2010). https://doi.org/10.1016/j.ecolind.2010.04.009
16. García Aguña, C., Kovacevic, M.: Uncertainty and sensitivity analysis of the human development index. Human Development Research Paper 2010/47, United Nations Development Programme (2011)
17. Gallego, I.: The use of economic, social and environmental indicators as a measure of sustainable development in Spain. Corpor. Soc. Respon. Environ. Manag. **13**, 78–97 (2005)
18. Ghobadian, A., Money, K., Hillenbrand, C.: Corporate responsibility research past-present-future. Group Organ. Manag. **40**(3), 271–294 (2015)
19. Gjølberg, M.: Measuring the immeasurable?: constructing an index of CSR practices and CSR performance in 20 countries. Scandinavian J. Manag. **25**(1), 10–22 (2009). https://doi.org/10.1016/j.scaman.2008.10.003
20. Global Reporting Initiative: Sustainability Reporting Guidelines Version 4.0 (G4) (2013)
21. Gray, R.: Taking a long view on what we now know about social and environmental accountability and reporting. Issues Soc. Environ. Account. **1**(2), 169–198 (2007)
22. Hahn, T., Figge, F., Aragón-Correa, J.A. et al.: Advancing research on corporate sustainability off to pastures new or back to the roots?. Bus Soc:0007650315576152 (2015). https://doi.org/10.1177/0007650315576152
23. Hediger, W.: Welfare and capital-theoretic foundations of corporate social responsibility and corporate sustainability. J. Socio-Econ. **39**(4), 518–526 (2010). https://doi.org/10.1016/j.socec.2010.02.001
24. Hosseini, H.M., Kaneko, S.: Dynamic sustainability assessment of countries at the macro level: a principal component analysis. Ecol. Indic. **11**, 811–823 (2011)
25. Ioannou, I., Serafeim, G.: What drives corporate social performance? the role of nation-level institutions. J. Int. Bus. Stud. **43**(9), 834–864 (2012)
26. Jamali, D., Safieddine, A.M., Rabbath, M.: Corporate governance and corporate social responsibility synergies and interrelationships. Corpor. Govern. Int. Review. **16**(5), 443–459 (2008). https://doi.org/10.1111/j.1467-8683.2008.00702.x
27. Kaiser, H.F.: An index of factorial simplicity. Psychometrika **39**(1), 31–36 (1974)
28. Levett, R.: Sustainability indicators-integrating quality of life and environmental protection. J. Royal Stat. Soc. Series A (Statistics in Society) **161**(3), 291–302 (1998). https://doi.org/10.1111/1467-985X.00109
29. Lima Crisóstomo, V., de Souza, Freire F., Cortes de Vasconcellos, F.: Corporate social responsibility, firm value and financial performance in Brazil. Social Respon. J. **7**(2), 295–309 (2011)

30. Mikulić, J., Kožić, I., Krešić, D.: Weighting indicators of tourism sustainability: a critical note. Ecol. Ind. **48**, 312–314 (2015)
31. Miras-Rodríguez, M., Carrasco-Gallego, A., Escobar-Pérez, B.: Has the CSR engagement of electrical companies had an effect on their performance? A closer look at the environment. Bus. Strategy Environ. **24**(8), 819–835 (2015)
32. Moneva, J.M., Archel, P., y Correa, C. : GRI and the camouflaging of corporate unsustainability. Account. Forum **30**(2), 121–137 (2006)
33. Montiel, I.: Corporate social responsibility and corporate sustainability separate pasts, common futures. Organ. Environ. **21**(3), 245–269 (2008)
34. Morse, S.: Measuring the success of sustainable development indices in terms of reporting by the global press. Soc. Indic. Res. **125**(2), 359–375 (2016)
35. Munda, G.: Social Multi-criteria Evaluation for a Sustainable Economy. Springer, Berlin (2008)
36. Munda, G., Nardo, M.: Constructing consistent composite indicators: the issue of weights. EUR 21834 EN (2005)
37. Nicoletti, G., Scarpetta, S., Boylaud, O.: Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation. OECD Economics Department Working Papers n 226 (2000). https://doi.org/10.1787/215182844604
38. Nollet, J., Filis, G., Mitrokostas, E.: Corporate social responsibility and financial performance: a non-linear and disaggregated approach. Econ. Model. **52**(part B), 400–407 (2016)
39. O'Connor, M., Spangenberg, J.H.: A methodology for CSR reporting: assuring a representative diversity of indicators across stakeholders, scales, sites and performance issues. J. Clean Prod. **16**(13), 1399–1415 (2008)
40. OECD: Handbook on constructing composite indicators. Methodology and user guide, OECD, Paris (2008)
41. Paredes-Gázquez, J.D., Rodríguez-Fernández, J.M., de la Cuesta-González, M.: Measuring corporate social responsibility using composite indices: Mission impossible? The case of the electricity utility industry. Revista de Contabilidad **19**(1), 142–153 (2016). https://doi.org/10.1016/j.rcsar.2015.10.001
42. Paruolo, P., Saisana, M., Saltelli, A.: Ratings and rankings: voodoo or science? J. Royal Stat. Soc. Series A (Statistics in Society) **176**(3), 609–634 (2013)
43. Pierick, E.T., Beekman, T., an Der Week, C.N., Meeusen, M.J.G. y De Graaff, R.P.M.: A framework for analyzing corporate social performance, beyond the wood model. The Hague: Agricultural Economic Research Institute (LEI) (2004)
44. Plischke, E.: An effective algorithm for computing global sensitivity indices (EASI). Reliab. Eng. Syst. Saf. **95**(4), 354–360 (2010)
45. Pulido Fernández, J.I., Sánchez Rivero, M.: Measuring tourism sustainability: proposal for a composite index. Tourism Econ. **15**, 277–296 (2009). https://doi.org/10.5367/000000009788254377
46. Rodríguez-Fernández, J.M.: Stakeholder model and social responsibility: a global corporate governance. M@n@gement **11**(2), 81–111 (2008)
47. Saisana, M., Saltelli, A., Tarantola, S.: Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. J. Royal Stat. Soc. Series A (Statistics in Society) **168**(2), 307–323 (2005)
48. Saltelli, A.: Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun. **145**(2), 280–297 (2002)
49. Saltelli, A., Ratto, M., Andres, T., et al.: Global Sensitivity Analysis: the Primer. Wiley, New York (2008)
50. Saltelli, A., Tarantola, S., Campolongo, F., et al.: Sensitivity Analysis in Practice: a Guide to Assessing Scientific Models. Wiley, New York (2004)
51. Salvati, L., Carlucci, M.: A composite index of sustainable development at the local scale: Italy as a case study. Ecol. Ind. **43**, 162–171, ISSN 1470-160X (2014). https://doi.org/10.1016/j.ecolind.2014.02.021
52. Salvati, L., Zitti, M.: Substitutability and weighting of ecological and economic indicators: Exploring the importance of various components of a synthetic index. Ecol. Econ. **68**(4), 1093–1099 (2009)

53. Sardain, A., Tang, C., Potvin, C.: Towards a dashboard of sustainability indicators for Panama: a participatory approach. Ecol. Ind. **70**, 545–556 (2016). https://doi.org/10.1016/j.ecolind.2016.06.038
54. Scalet, S., Kelly, T.F.: CSR rating agencies: What is their global impact? J. Bus. Ethics **94**, 69–88 (2010)
55. Schaltegger, S., Etxeberria, I., Álvarez and Ortas, E. : Innovating corporate accounting and reporting for sustainability ? attributes and challenges. Sustain. Dev. (2017). https://doi.org/10.1002/sd.1666
56. Searcy, C.: Corporate sustainability performance measurement systems: a review and research agenda. J. Bus. Ethics **107**(3), 239–253 (2012)
57. Singh, R.K., Murty, H.R., Gupta, S.K., Dikshit, A.K.: An overview of sustainability assessment methodologies. Ecol. Ind. **15**(1), 281–299 (2012)
58. Sobol, I.M.: On the distribution of points in a cube and the approximate evaluation of integrals. USSR Comput. Math. Phys. **7**, 86–112 (1967)
59. Sensitivity analysis for non-linear mathematical models. Math. Modll. Comput. Exp. **1**, 407–414
60. Spangenberg, J.H.: The corporate human development index CHDI: a tool for corporate social sustainability management and reporting. J. Clean Prod. **134**, Part A:414–424 (2016). https://doi.org/10.1016/j.jclepro.2015.12.043
61. UNCTAD: Guidance on Corporate Responsibility Indicators in annual reports. United Nations Conference On Trade And Development (UNCTAD), United Nations (2008)
62. van Buuren, S.: Multiple imputation of discrete and continuous data by fully conditional specification. Stat. Methods Med. Res. **16**(3), 219–242 (2007)
63. Vigneau, L., Humphreys, M., Moon, J.: How do firms comply with international sustainability standards? processes and consequences of adopting the global reporting initiative. J. Bus. Ethics **131**(2), 469–486 (2014)
64. Wang, Q., Dou, J., Ase, J.: A meta-analytic review of corporate social responsibility and corporate financial performance: the moderating effect of contextual factors. Bus. Soc. **8**, 1083–1121 (2016)
65. Wilde-Ramsing, J.: Quality Kilowatts: a normative-empirical approach to the challenge of defining and providing sustainable electricity in developing countries. Doctoral Dissertation, University of Twente (2013)
66. Wood, D.J.: Measuring corporate social performance: a review. Int. J. Manag. Rev. **12**(1), 50–84 (2010)
67. Zhao, Z., Zhao, X., Davidson, K., et al.: A corporate social responsibility indicator system for construction enterprises. J. Clean Prod. **29–30**, 277–289 (2012). https://doi.org/10.1016/j.jclepro.2011.12.036

A Logic-Based Approach to Incremental Reasoning on Multi-agent Systems

Elena V. Ravve, Zeev Volkovich, and Gerhard-Wilhelm Weber

Abstract We introduce the notion of strongly distributed multi-agent systems and present a uniform approach to incremental automated reasoning on them. The approach is based on systematic use of two logical reduction techniques: Feferman-Vaught reductions and syntactically defined translation schemes. The distributed systems are presented as logical structures $\mathcal{A}$'s. We propose a uniform template for methods, which allow for certain cost evaluation of formulae of logic $\mathcal{L}$ over $\mathcal{A}$ from values of formulae over its components and values of formulae over the index structure $\mathcal{I}$. Given logic $\mathcal{L}$, structure $\mathcal{A}$ as a composition of structures $\mathcal{A}_i$, $i \in I$, index structure $\mathcal{I}$ and formula ϕ of the logic to be evaluated on $\mathcal{A}$, the question is: what is the reduction sequence for ϕ if any. We show that if we may prove preservation theorems for $\mathcal{L}$ as well as if $\mathcal{A}$ is a strongly distributed composition of its components then the corresponding reduction sequence for $\mathcal{A}$ may be effectively computed. We show that the approach works for lots of extensions of *FOL* but not all. The considered extensions of *FOL* are suitable candidates for modeling languages for components and services, used in incremental automated reasoning, data mining, decision making, planning and scheduling. A short complexity analysis of the method is also provided.

Keywords Incremental reasoning · Multi-agent system · Strongly distributed system · Logical reduction sequence

E. V. Ravve (✉) · Z. Volkovich
Software Engineering Department, Ort Braude College, Rehov Snunit 51, POB 78, Karmiel 2161002, Israel
e-mail: cselena@braude.ac.il

Z. Volkovich
e-mail: vlvolkov@braude.ac.il

G.-W. Weber
Faculty of Engineering Management, Poznan University of Technology, ul. Jacka Rychlewskiego 2, 60-965 Poznan, Poland
e-mail: gerhard.weber@put.poznan.pl

IAM, METU, 06800 Ankara, Turkey

A. Pinto and D. Zilberman (eds.), *Modeling, Dynamics, Optimization and Bioeconomics IV*, Springer Proceedings in Mathematics & Statistics 365,
https://doi.org/10.1007/978-3-030-78163-7_18

1 Introduction

Multi-agent systems (*MAS*) have numerous civilian, security and military applications. Centralized quantitative and qualitative modeling, analysis, constraint satisfaction, maintenance and management seem to be too rigid for these systems. On the other hand, the distributed and incremental reasoning on the systems seems to be more scalable, robust, and flexible. Incremental automated reasoning, data mining, decision making, planning and scheduling for multi-agent systems is under massive attack in the last decades; cf. [39, 102, 111, 118].

In this contribution, we propose a method to solve the following problem. *Given a method of construction of a MAS, from several components. Given either boolean or quantitative property, constraint or description of the desired behaviour of the MAS. We want to define* **algorithmically** *a derived set of properties (constraints, behaviours, etc.) on the components, and a post-processing procedure such that the procedure is capable to evaluate the property of MAS from the locally evaluated properties of components.*

1.1 A Motivating Example: Alternation of Data Streams

Let us consider a toy example, inspired by [109], in order to illustrate the main idea. Assume that we are given two streams: the *First* stream and the *Second* stream. These streams are ordered sequences of their elements, which may be labeled. The first element of a stream is labeled by a star ($\star$).

These streams are combined in a resulting stream by alternation. Finite constant number of elements of each stream are labeled by $\updownarrow$ that is aimed to define the switch points of the streams, for treatment of stochastic switches cf. [108, 117]. The combination of these streams is defined as follows.

Combination rules: The resulting stream is composed from the fragment of the first stream from its first element, labeled by $\star$ till its first element, labeled by $\updownarrow$. This element, labeled by the $\updownarrow$, is the last element of the first stream that is included into the resulting stream. Then the resulting stream includes all the elements from the second stream from its first element, labeled by $\star$ until its first element, labeled by $\updownarrow$. Then we switch back to the first stream an so on, see Fig. 1.

Some elements of the streams may be also labeled by $\oplus$ and some elements of the streams are labeled by $\otimes$. Assume that our *MAS* is aimed to compute and analyze a global quantitative property on the resulting stream. Assume that the desired property counts the number of pairs, for which the first element is labeled by $\oplus$ and the second element is labeled by $\otimes$. Each agent monitors one component and transfers locally observed results to the centralized final proceeding.

In order to propagate the computation to the components (two original streams), we observe, that a pair of such elements of the resulting stream is labeled by $\oplus\otimes$ in two cases.

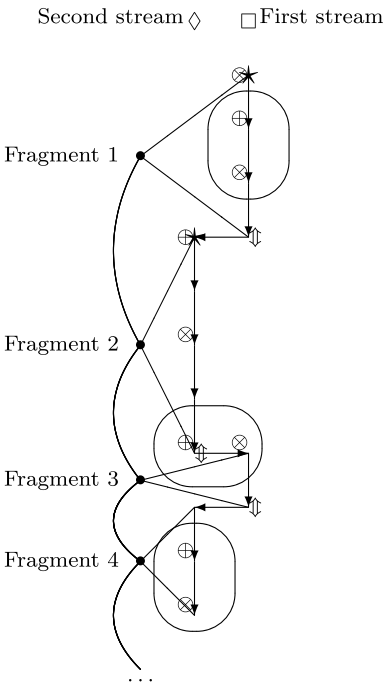


Fig. 1 Alternation of two data streams

Case 1: Both elements belong to the same original stream. The first condition is translated to computation of the number of $\oplus\otimes$ sequences in the original streams. There are two such pairs on Fig. 1: one in Fragment 1 and one more in Fragment 4. They are marked by two vertical ovals.
Case 2: These elements are coming from different original streams. The second condition is translated to computation of the number of pairs of elements x and y from different streams, which satisfy the following properties:
- Last element of a fragment is labeled by $\oplus$:
x is labeled by both $\oplus$ and $\updownarrow$;
- First element of a fragment is labeled by $\otimes$:
y is labeled by $\otimes$ and either it is also labeled by $\star$ or its immediate predecessor is labeled by $\updownarrow$.
- The fragments are succeeding:
x belongs to a fragment that is an immediate predecessor of the fragment of y in the resulting stream.

There exists one[1] pair of elements on Fig. 1, which satisfies all the conditions of the second case: on the border between Fragments 2 and 3. This pair is marked by the horizontal oval.

[1] The first element of the *First* stream satisfy the condition: **First element of a fragment is labeled by** $\otimes$. However, its fragment does not satisfy the condition: **The fragments are succeeding**.

The final calculation sums all the local results: three ovals on Fig. 1.
From the consideration, we observe:

1. The propagation is exact: the final calculation on the results of the evaluation of the derived local properties **is equivalent** to the evaluation of the global one.
2. The derived properties **do not depend** upon particular streams but rather upon the combination rules. For all particular streams, which are combined, according to the predefined rules, the same derived properties must be checked.
3. The global and the derived properties may coincide (case 1).
4. The global and the derived properties might not coincide (case 2).
5. The derived properties may be checked on the original data streams: **Last element of a fragment is labeled by** $\oplus$ or **First element of a fragment is labeled by** $\otimes$.
6. The derived properties may be checked on the index structure of the fragments: **The fragments are succeeding**.

Fortunately, the above counting properties are expressible in *Weighted Monadic Second Order Logic* (*WMSOL*). More surprisingly, such reductions of computation of a global property on a resulting data stream to the derived properties on components may be sometimes converted to an algorithm. In this contribution, we prove that such an algorithm exists for any *strongly distributed multi-agent system.*

1.2 Novelty of the Contribution: Logical Tools

We propose a logic based approach to incremental reasoning on *MAS* that is new in the context. The approach allows unification of the investigation of *MAS* as well as significant reduction of the communication load between the components. We do not use any approximation, our results are exact. Our approach is applicable to reasoning of any property, expressible in lots of extensions of *First Order Logic* (*FOL*). We systematically exploit logical reduction techniques to big *MAS*'s. We assume that the reader has general logical background as may be found in [33, 34].

Logical reduction techniques come in two favors:

1. **Feferman-Vaught reductions**, which are applied in situations of distributed data. The reduction describes how the queries over a distributed data can be computed from queries over the components and queries over the index set. Feferman-Vaught reductions were first introduced in model theory in [38]. Their use in computer science was seemingly first suggested in [100] in the context of formal verification and model checking, and in [16] in the context of graph algorithms for graphs of bounded clique width. For the algorithmic uses of the Feferman-Vaught Theorem see also [75]. The reductions are some analogue of *local violations* of [57].
2. **The syntactically defined translation schemes**, known also in model theory as interpretations, studied in [50]. They describe transformations of data and queries.

They give rise to two induced maps: *translations* and *transductions*. Transductions describe the induced transformation of data instances and the translations describe the induced transformations of queries. The fundamental property of translation schemes describes how to compute transformed queries in the same way Leibniz' Theorem describes how to compute transformed integrals. The fundamental property of the syntactically defined translation schemes has a long history, but was first properly stated by Rabin; cf. [91]. One of the most recent uses of interpretations in the field of database theory may be found in [43, 99].

In this paper we compile, summarize, adopt and extend our previous publications in the field, in particular [93, 95, 98]. We provide details and explanations, which were omitted there due to the space limitations. Moreover, we include in our framework *fuzzy* systems as well.

1.3 Strongly Distributed MAS's

Combination and adaptation of these techniques allow us to introduce the notion of *strongly distributed MAS's*. For such *MAS*'s, we extend and generalize the known techniques of incremental reasoning. For the strongly distributed *MAS*'s, we provide the precise definition of *locality* as well as improved *complexity analysis*, which generalizes such models of parallel and distributed computations and communications, as *BSP* of [114], and *LogP* of [20].

For the strongly distributed *MAS*'s, we derive the following main steps of the reasoning:

1. Computation of the derived properties on the components;
2. Computation of the derived properties on the index;
3. Final evaluation, based on the obtained results.

Our use of the logical reduction techniques in the field of the incremental reasoning on *MAS*'s is new. The notion of *strongly distributed systems* was introduced in [22] in the context of information systems.

1.4 Incremental Reasoning

Our method shows how a global reasoning property, expressed as formula ϕ on strongly distributed systems, may be syntactically reduced to incremental computations of effectively derived properties on components, the index $\mathcal{I}$ and some post-proceeding. The exact formulation of this is rather involved and is explained in details in Sect. 7. It is an extension of the Feferman–Vaught Theorem; cf. [38], for *FOL*. For *FOL*, the Feferman–Vaught Theorem covers a very wide class of gener-

alized products and sums of structures and is extremely powerful. From our main Theorem 9, we derive a method for computing ϕ on strongly distributed systems, which proceeds as follows:

Preprocessing: Given ϕ and translation scheme Φ that describes combination of the local agents to the system, but not the strongly distributed systems themselves; we construct a sequence of formulae $\psi_{\iota,J}$ and an evaluation function $F_{\Phi,\phi}$.

Incremental Computation: We compute the local values $\psi_{\iota,J}$ for each local component.

Final Solution: Theorem 9 now states that ϕ on the strongly distributed system may be effectively computed from $\psi_{\iota,J}$, using $F_{\Phi,\phi}$.

We emphasize the following:

Communication load: The only values, transferred between different components, are the (boolean or quantitative) values of $\psi_{\iota,J}$, that significantly reduces the communication load.

Confidentiality: All meaningful information is still corresponds to the components in the secure way and is not transferred. The transferred values $\psi_{\iota,J}$, as a rule, are meaningless without the knowledge about the final proceeding.

We systematically apply two logical reduction techniques to the field of reasoning on distributed *MAS*. The *MAS*'s are presented as logical structures $\mathcal{A}$'s, built from components. Properties to be analyzed, constraints to be satisfied, maintenance and management of *MAS* are presented as logical formulae. The boolean properties of $\mathcal{A}$ are expressed as different extensions of First Order Logic (*FOL*). The quantitative properties of $\mathcal{A}$ are expressed in Weighted Monadic Second Order Logic (*WMSOL*); cf. [31].

We propose a uniform template for methods, which allow for a certain cost evaluation of formulae of logic $\mathcal{L}$ over $\mathcal{A}$ from values of formulae over its components and values of formulae over the index $\mathcal{I}$. We show that our approach works for a great variety of extensions of *FOL* and *WMSOL*, but not all.

1.5 Structure of the Contribution

The paper is organized as follows. In Sect. 2, we quote results, concerning incremental reasoning on (fuzzy) multi-agent systems. Modeling of *IA*'s as logical structures is recalled in Sect. 3. An already published motivating example is presented in Sect. 4. Section 5 describes logical presentation of Disjoint Union and Shuffling of Structures. In Sect. 6, we give the general framework of the Syntactically Defined Translation Schemes. Strongly Distributed *MAS*'s are introduced in Sect. 7. Section 8 is dedicated to the complexity analysis of our method. Section 9 is the main section of the paper. Section 10 summarizes the paper. Sections. 11 and 12 provide the main logical background, used in the contribution. In Sect. 13, we give our original proof of a part of our main Theorem 4.

2 Related Work

In this section, we sketch results, concerning incremental reasoning on multi-agent systems, including fuzzy systems.

Incremental *heuristic* methods of reasoning are well developed and practically implemented. These methods find and reuse information from previous runs to reach solutions faster than by solving each problem from the very beginning. For combination of incremental and heuristic search, see [59]. A multi-criteria variation of the approach may be found in [88]. In this paper, we show rather an *algorithmic* (non-stochastic) approach to incremental treatment of *MAS*. The algorithm allows the *exact* (not approximation of) analysis, maintenance and management of *MAS*, based on the derived requirements for the individual intelligent agents (*IA*'s). Our method also allows massive reuse of results from previous runs as it is shown, particularly, in Sect. 8.4.

Multi-agent Partially Observable Markov Decision Processes (*MPOMDP*) provide a powerful framework for optimal decision making under the assumption of instantaneous communication. The model with delayed communication setting (*MPOMDP*$-$*DC*), in which broadcasted information is delayed by at most one time step, is investigated in [112]. The authors demonstrate that computation of the *MPOMDP*$-$*DC* backup can be structured as *a tree*. Moreover, they introduce two novel *tree-based* pruning techniques that exploit this structure in an effective way. In this paper, we show how that *tree-based* approach may be generalized using *weighted labeled trees*, as introduced in [31], or even more generally as *behavior of finite state machines*, as described in [19].

Distributed reinforcement learning in cooperative multi-agent decision processes was considered, especially, in [64]. In this scenario, an ensemble of simultaneously and independently acting agents tries to maximize a discounted sum of rewards. As a rule, one assumes that each agent has *no information* about its team mates' behaviour. In this paper, we introduce the notion of *strongly distributed MAS*, which allows some communication between the agents, well defined by a translation scheme on their disjoint union. An example of modeling and analysis of such systems is presented in Sect. 4. In this case, three *IA*'s minimize their total run time, when some communication between them is allowed.

In constraint satisfaction, local consistency conditions are properties of constraint satisfaction problems related to the consistency of subsets of variables or constraints. Several such conditions exist, the best known being node consistency, arc consistency, and path consistency. Local consistency can be enforced via transformations of the problem called constraint propagation; cf. [103]. New distributed algorithms for incremental maintaining of partial path consistency have been recently presented in [8]. In our paper, we show how the propagation technique may be effectively applied to **any** *strongly distributed MAS*.

The Point-Based Incremental Pruning Toolbox is a platform dedicated to approximate solution methods for decentralized stochastic control problems, represented as different variations of *POMDP*. This toolbox serves as a platform for developing

approximate algorithms with no-provable bounds for solving decentralized stochastic control problems and associated applications; cf. [25]. Our approach allows a precise definition and evaluation of the cost of our incremental reasoning. For a detailed complexity analysis of the method, see Sect. 8.

One of the successful attempts to specify the distributed multi-agent reasoning system (*dMARS*) is presented in [27]. In general, there are two schools of thought on reasoning about distributed systems: one is following interleaving based semantics, and the other one follows partial-order/graph based semantics; cf. [23]. The second one seems to be more promising.

Dimensionality reduction using different techniques and pre-processing for two classification algorithms were investigated in [89]. Recently, in [4], an approach, based on local methods of smooth optimization to solve the clustering problems, was proposed. The incremental approach is applied in order to generate starting points for cluster centers. However, to our best knowledge, no general, **logically based**, approach to incremental reasoning multi-agent systems has yet been proposed.

Decomposition and incremental reasoning on switch systems goes back to the early 60's, cf. [3, 21]. Since Zadeh introduced the fuzzy set theory in [124], by exploiting the concept of membership grade, numerous attempts to investigate fuzzy systems and their properties have been applied. In this context, the theory of disjunctive decompositions of the [3, 21] and others, was shown to be insufficient. From the pioneering works, investigating the problem, we refer only to [56], where an approach for obtaining simple disjunctive decompositions of fuzzy functions is described. However, the approach is not scalable to large functions and hardly implemented. The summary of the first results may be found in [55]. See [87] for the next contributions in the field.

Jumping to the 90's, we mention [116], which deals with the problem of general max-min decomposition of binary fuzzy relations defined in the Cartesian product of finite spaces. In [125], a new method to derive the membership functions and reference rules of a fuzzy system was developed. Using this method, a complicated *Multiple Input Single Output* (*MISO*) system can be obtained from combination of several *Single Input Single Output* (*SIMO*) systems with a special coupling method. Moreover, it was shown how the decomposition and coupling method reduces complexity of the network, used to represent the fuzzy system. Theoretical results on structural decomposition of general *MIMO* fuzzy systems are presented in [123]. Some recent results on a decomposition technique for complex systems into hierarchical and multi-layered fuzzy logic sub-systems may be found in [83].

For α-decomposition of [121] (originated from $max - min$ composition of [107]), in [122], it was shown that every fuzzy relation R is always generally effectively α-decomposable. Moreover, calculating of

$$\rho(R) = \min\{|Z| : R = Q\alpha T, Q \in F(X \times Z), T \in F(Z \times Y)\}$$

is an NP-complete problem. A new concept for the decomposition of fuzzy numbers into a finite number of α-cuts is provided in [110].

In [7], the *Multiagent Simple Temporal Problem* was formulated. The proposed foundational algorithms for scheduling agents allow reasoning over the distributed but interconnected scheduling problems of multiple individuals. The method combines bottom-up and top-down approaches. In the bottom-up phase, an agent externalizes constraints that compactly summarize how its local subproblem affects other agents' subproblems, whereas in the top-down phase an agent proactively constructs and internalizes new local constraints that decoup'le its subproblem from others'. In our paper, we propose an approach that allows to solve incrementally any problem, expressible in a logic. Expressive power of the logics is rather discussed in Sects. 3.2 and 12.2.

Recently, in [14], the problem of maintaining the temporal consistency of distributed plans during execution, when temporal constraints may be updated was investigated. New incremental algorithms for managing dynamic Multi-agent Simple Temporal Network (*MaSTN*). These algorithms help each agent know, as soon as possible, whether the distributed plan to be executed is still temporally consistent. In these latter days, in [82], a new distributed algorithm for decoupling the *MaSTN* problem was proposed. The agents cooperatively decouple the *MaSTN* while simultaneously optimizing a sum of concave objectives local to each agent. Again, the approaches deal with particular problems, while we propose a more general approach that allows to solve incrementally any problem, expessible in a logic.

Application of logical apparatus to multi-agent systems is not really new. We go directly to [105], distributed default logic (*DDL*), the formalism for multi-agent knowledge representation and reasoning, was introduced. Afterwards, in [5], quantified interpreted systems, a semantics to reason about knowledge in multi-agent systems in a first-order setting, was introduced. Moreover, first order modal axiomatisations for different settings was defined. It was shown that they are sound and complete with respect to the corresponding semantical classes. In our paper, we propose a *generalized* purely theoretical approach, which may be directly applied to incremental reasoning even on fuzzy and quantitative distributed systems.

In fact, we may consider fuzzy logic as a infinite-valued (infinitely-many valued) logic, in which the law of excluded middle does not hold. In fact, the truth function for an extension of *FOL* relation R with a fuzzy relation is a mapping in the interval [0, 1]. History of many-valued logics (a propositional calculus, in which there are more than two truth values) comes back to early 20's of the previous century, cf. [68, 90]. One of the first formalizations of such a view may be found in [28]. The approach leads to the following definition of a fuzzy truth-value lattice: *A fuzzy truth-value lattice* of [13] is a lattice of fuzzy sets on [0, 1] that includes two complete sublattices $\mathbf{T}$ and $\mathbf{F}$ such that:

1. $\forall v_1 \in \mathbf{T}\, \forall v_2 \in \mathbf{F} : v_1$ and v_2 incomparable;
2. $\forall S \in \mathbf{T} :\ lub(S) \in \mathbf{T}$ and $glb(S) \in \mathbf{T}$,
   $\forall S \in \mathbf{F} :\ lub(S) \in \mathbf{F}$ and $glb(S) \in \mathbf{F}$;
3. $\forall v \in \mathbf{T}\, \forall \epsilon \in [0, 1] :$ if $\exists v^* \in \mathbf{T} :\ v^* \leq_l v + \epsilon$ then $v + \epsilon \in \mathbf{T}$;
   $\forall v \in \mathbf{F}\, \forall \epsilon \in [0, 1] :$ if $\exists v^* \in \mathbf{F} :\ v^* \leq_l v + \epsilon$ then $v + \epsilon \in \mathbf{F}$.

Here, **T** and **F**, respectively, denote the set of all **TRUE**-characteristic truth-values and the set of all **FALSE**-characteristic false-values in the lattice; *lub* and *glb* are the labels of the *least upper bound* and the *greatest lower bound*.

In a particular definition of a truth-value lattice, *lub* and *glb* are interpreted by specific operations. There exists a variety of fuzzy set intersection and union definitions, cf. [58], and *lub* and *glb* can be defined to be any corresponding ones of them. Moreover, systems based on real numbers in [0, 1] having truth-characteristics distinguished, cf. [28], commonly use 0.5 as the splitting point between **FALSE**- and **TRUE**-characteristic regions, where 0.5 is considered an **UNKNOWN**-characteristic truth-value. In such a case, *lub* corresponds to *max*, *glb* corresponds to *min* and $\leq$ is the usual real number less-than-or-equal-to relation. For a possible connection of fuzzy logic and graph grammars, see [48].

In our paper, we generalize and extend the coupling method of [125] by systematic application of two logical reduction techniques to the field of reasoning on fuzzy distributed systems. The fuzzy systems are modeled by different variations of *WMSOL*-structures.

In [42], a generalized model of finite automata, which is able to work over arbitrary structures, was introduced and studied However, it turns out that many elementary properties known from classical finite automata are lost. The concept was improved in [79, 80]. The new model is restricted in that computed information is deleted after a fixed period in time. Thus the new model seems to reflect more adequately what might be considered as a finite automata over the reals and similar structures. In this paper, we use *WTA* to compute the quantitative coverage properties. These automata are related to *WMSOL* and allow for incremental evaluation of the formula in many cases.

3 Modeling Intelligent Agents as Logical Structures

In this section, we recall how we use *Finite State Machine* (*FSM*), as introduced in [19], in order to model (behaviour of) *MAS*; cf. [95].

An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives; cf. [119]. This type of units originates from hardware and some sort of software: control systems, software demons, etc. In our understanding of *Intelligent Agents* (*IA*'s), we mostly follow [120], where the following list of capabilities was suggested in order to characterize *AI*'s:

Reactivity: Intelligent agents are able to perceive their environment, and respond in a timely fashion to changes that occur in it in order to satisfy their design objectives.

Proactiveness: Intelligent agents are able to exhibit goal-directed behaviour by taking the initiative in order to satisfy their design objectives.

Social ability: Intelligent agents are capable of interacting with other agents (and possibly humans) in order to satisfy their design objectives.

In computer science, the most popular ways to model such units are: *FSM*, *Kripke*−model, *Petri*−nets, etc. For hardware units, *FSM* is even effectively extractable from the Hardware Descriptive Language code of the design, cf. [54].

We use the modeling in order to present *MAS* as logical structures $\mathcal{A}$'s (agents), built from components. Each component is also presented as a logical structure $\mathcal{A}_i$ (agent i). Properties, constraints and the desired behavior of *MAS* are presented as logical formulae ϕ's. The formal presentation of the problem that we want to solve is:

> Is it possible to evaluate ϕ on $\mathcal{A}$, using some effectively derived formulae on $\mathcal{A}_i$, and (may be) some additional manipulations?

A simple agent program can be defined mathematically as an *agent function*, cf. [104]. The function maps every possible percepts sequence either to a possible action, which the agent can perform, or to a coefficient, feedback element, function or constant, which affects eventual actions: $f : Percept^* \rightarrow Act$. Agent function is an abstract concept as it could incorporate various principles of decision making like calculation of utility, deduction over logic rules, etc.; cf [106]. Note that the *agent function* is a special case of a logical binary relation.

In order to show how computation may be expressed as logical formulae, we take an example from [53], where addition was shown to be expressible in *FOL* in the following way. Let $\tau = \langle A, B, k \rangle$ be a vocabulary consisting of two unary relation and a constant k. In structure $\mathfrak{M}$ of the vocabulary, A and B are binary strings of length $k = |\mathfrak{M}|$. $\mathfrak{M}$ satisfies the additional property if the kth bit of the sum of A and B is one:

$$CARRY(x) = (\exists y < x)[A(y) \wedge B(y)(\forall z\, y < z < x)A(z) \vee B(z)],$$

$$PLUS(x) = A(x) \oplus B(x) \oplus CARRY(x).$$

The last sentence expresses the addition property.

3.1 Finite State Machine and Its Behaviour as Logical Structures

In this work, we describe in great details how *FSM*'s may be used as models of *IA*'s in *MAS*. However, our approach works for *any* suitable modeling of *IA*'s as logical structures. We use logical formulae over the logical structures to describe computations, properties, constraints, maintenance and management of *MAS*. The *FSM* may be defined as a set of states and transitions between them in the following way.

Definition 1 (*Finite State Machine,* [19])
$\mathcal{M}_{FSM} = \langle S \cup T, P_S, P_T, \mathbf{init}, \mathbf{src}, \mathbf{tgt}, P_1, \ldots, P_k \rangle$, where

– S and T are two sets, called the set of *states* and the set of *transitions*, respectively;
– P_S is a unary predicate for the states;
– P_T is a unary predicate for the transitions;
– **init** is a state called the *initial state*;
– **src** and **tgt** are two mappings from T to S that define respectively the *source* and the *target* of a transition;
– $P_1, \ldots, P_k$ are subsets either of S or of T, that specify properties of the states or of the transitions, respectively.

State predicate P_ℓ may denote *Idle* state, when there is nothing for the *FSM* to do, *Accepting* states, *Rejecting* states, *Error* states, etc. Transition predicate P_{Ack} may stand for perception of an input *Acknowledgment* signal, and so on.

FSM is not really interesting as an objects under consideration. Given *FSM*, we are rather interested in different properties of its runs, which are affected by different percepts sequences $Percept^*$.

Definition 2 (*Paths of runs of FSM*)

1. A *path of a run* of *FSM* $\mathcal{M}$ is a finite sequence of transitions $(t_1, t_2, \ldots, t_n)$, such that the source of t_1 is the initial state and for every $i = 1, \ldots, n-1$: $t_i = (s_{i_1}, s_{i_2}) \in \mathbf{T}$.
2. $\mathbf{Paths}(\mathcal{M})$ is the set of paths of runs of *FSM* $\mathcal{M}$ and $\prec$ is the *prefix order* on $\mathbf{Paths}(\mathcal{M})$.[2]

Now, we define the notion of the *behaviour* of a *FSM*.

Definition 3 (*Behavior of FSM*)
The *behavior* of $\mathcal{M}$ is a logical structure

$$\mathbf{bhv}(\mathcal{M}) = \langle \mathbf{Paths}(\mathcal{M}), \prec_1, P_1^*, \ldots, P_k^* \rangle,$$

where each P_i^* is the property of a path saying that the last state of this path satisfies P_i.[3]

We note that the *agent function*, as defined above, falls properly in the general framework the behavior of *FSM*.

3.2 Expressive Power of MSOL on Finite State Machines

One of the ways to express properties, constraints and the desired behavior of *IA*'s is the use of *logical* formulae. Monadic Second Order Logic (*MSOL*) is an extension of

[2] The prefix order means here the prefix order on strings, which is a special kind of substring relation. $p \prec_1 p'$ iff $p \prec p'$ and p' has exactly one more transition than p.

[3] In the definition, the prefix order $\prec_1$ generalizes the intuitive concept of a tree by introducing the possibility of continuous progress and continuous branching.

FOL, when quantification on unary relations is allowed. This logic has considerable expressive power on *FSM*'s and their behaviour. Most of the logics used in different applications of computer science are sublogics of it. Among these we have:

- **Monadic Fixed Point Logic** (with inflationary fixed points): This Logic is *MSOL* expressible on the *FSM*. Fixed Point Logic, in general, is definable in *SOL*.
- **Propositional Dynamic Logic** [47]: Actually, it is definable in Monadic Fixed Point Logic on the *FSM*.
- **The Fixed Point Definable Operations** on the power set algebra of trees, [2]. They are *MSOL* expressible on the *FSM*.
- **Linear Temporal Logic** (*LTL*): Actually, *LTL* expressible already in First Order Logic on the behavior of the *FSM*, as was shown by Kamp; [12].
- **Computation Tree Logic** (*CTL*) [65]: Actually, it is in Monadic Fixed Point Logic on behavior of the *FSM*.
- **Fragments of μ and ν Calculus**: The operations, which can be defined without alternation on μ and ν, cf. [2]. Actually, they are definable in Weak Monadic Second Order Logic on behavior of the *FSM*.

Moreover, using powerful tools based on ideas related to Rabin's Theorem on the decidability of *MSOL* on infinite trees, [92], Courcelle [17], and Courcelle and Walukiewicz [18] proved:

Theorem 1 (Courcelle and Niwiński, [18])
Every MSOL expressible property of behavior of the FSM is equivalent to some MSOL expressible property of the FSM.

3.3 Parallel Runs of Intelligent Agents with Message Passing

Let us consider two *FSM*s (FSM^1 and FSM^2), which can communicate via two one-way channels (message passing parallel computation); see Fig. 2. Both asynchronous and synchronous communication may be modeled in this way. For synchronous communication, after a state, labeled by P_{Send}, the *FSM* goes to a state, labeled by P_{Idle}, until perception of an input signal *Ack* is observed. For more detailed investigation of the example see [96].

We assume that the vocabularies of the FSM^i ($i \in \{1, 2\}$) in the simplest case contain at least two unary predicates, which indicate P_{Send} and $P_{Receive}$ for communication states. The communication with the external environment is omitted for simplicity and

$$\mathcal{M}^i_{FSM} = \langle S^i \cup T^i, P^i_S, P^i_T, \mathbf{init}^i, \mathbf{src}^i, \mathbf{tgt}^i, P^i_{Receive}, P^i_{Send} P^i_1, \ldots, P^i_{k^i} \rangle.$$

The resulting *FSM* may be described as

$$\mathcal{M}_{FSM} = \langle S^1 \dot{\cup} S^2 \dot{\cup} T^1 \dot{\cup} T^2 \cup \{(P^1_{Send}, P^2_{Receive}), (P^2_{Send}, P^1_{Receive})\},$$

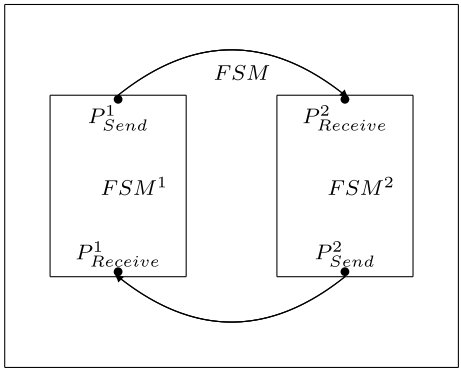


Fig. 2 Composition of two *FSM*'s

$$\cup\{(\mathbf{init}, \mathbf{init}^1), (\mathbf{init}, \mathbf{init}^2)\}, P^1_S, P^1_T, P^2_S, P^2_T, P_T,$$

$$\mathbf{init}, \mathbf{src}, \mathbf{tgt}, \mathbf{init}^1, \mathbf{src}^1, \mathbf{tgt}^1, \mathbf{init}^2, \mathbf{src}^2, \mathbf{tgt}^2,$$

$$P^1_{Receive}, P^1_{Send}, \ldots, P^2_{Receive}, P^2_{Send}\rangle.$$

Here, the new added transitions and predicates are:

- The new added transitions are: $(P^1_{Send}, P^2_{Receive})$, $(P^2_{Send}, P^1_{Receive})$ and $(\mathbf{init}, \mathbf{init}^1)$, $(\mathbf{init}, \mathbf{init}^2)$;
- The initial state of $\mathcal{M}_{FSM}$ is labeled by **init**;
- The unary predicate for the new added transitions is P_T;
- The new functions **src** and **tgt** map the new transitions to their sources and targets.

4 Motivating Example: Cooperation of Three Intelligent Agents

The example is taken verbatim form [94]. Assume that we are given a system with three intelligent agents, which may communicate according to predefined rules. Assume that *the agent number one* may call both agents: *the agent number two* and *the agent number three*. Moreover, *the agent number two* can not call any agent, while *the agent number three* may call back *the agent number one*; see Fig. 3. We use the following formalization: the runs of the agents are accumulated in weighted labeled trees: the weights are put on the edges of the trees. The vertices also may be labeled. A run is a path in the tree, as introduced in Definition 2.

Assume that run tree T of the multi-agent system is presented on Fig. 4. In this figure, we omit the weights, put on the edges. The meaning of the labels on the vertices is as follows:

Fig. 3 A system with 3 agents

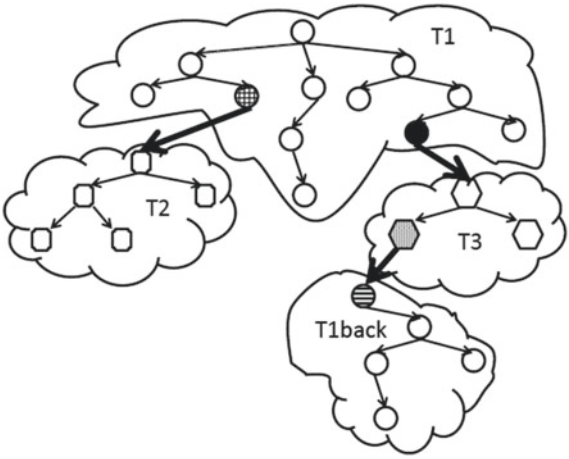


Fig. 4 Run tree T of a system with 3 agents

- the gridded vertex in $T1$ corresponds to the call to *the agent number two* by *the agent number one* (the bold edge goes to the root of $T2$);
- the filled vertex in $T1$ corresponds to the call to *the agent number three* by *the agent number one* (the bold edge goes to the root of $T3$);
- the doted vertex in $T3$ corresponds to the call to *the agent number one* by *the agent number three* (the bold edge does **NOT** go to the root of $T1$ but rather goes back to a vertex of $T1$ labeled by strips).

Assume that we want to optimize (minimize) the runs in the tree. On the one hand, we may use one of the optimization algorithms on the complete tree T that will give quantitative result $\mathcal{R}$. On the other hand, we observe that we may receive the optimal result in the following way:

1. Find the optimal run $\mathcal{R}_1$ in $T1$.
2. Find **ALL** labeled runs Λ_{i_1} in $T1$: there are two such runs Λ_{1_1} and Λ_{2_1}.
3. Find the optimal run $\mathcal{R}_2$ in $T2$. For the optimization, we will use $\Lambda_{1_1} + \mathcal{R}_2$.
4. Find the optimal run $\mathcal{R}_3$ in $T3$. For the optimization, we will use $\Lambda_{2_1} + \mathcal{R}_3$.
5. Find **ALL** labeled runs Λ_{i_3} in $T3$: there is one such a run Λ_{1_3}.
6. Find the optimal run $\mathcal{R}_{1_{back}}$ in $T1back$. We will use $\Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}$.
7. Finally, we find $\min\{\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}\}$.

8. We observe that $\mathcal{R} = \min\{\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}\}$.

In order to generalize the obtained observations of the example, we need:

1. The precise definition of languages, which describe different problems: cf. Sect. 11.
2. The precise definition of the weighted labeled trees and computations on them: cf. Sect. 12.
3. Tree T is **NOT** a disjoint union of its sub-trees. We need a formal framework to deal with such objects, which is rather presented in Sect. 7.
4. Let $\mathfrak{T}_{old}(N)$ denote the time to solve the problem directly (N stands for the size of the coding of T).
   - $\mathfrak{E}_I$ denotes time to extract index structure I from T. We have four ordered numbers to distinguish sub-trees in our example.
   - $\mathfrak{E}_\iota$ denotes time to extract each T_ι from T. We have four numbered sub-trees in our example.
   - $\mathfrak{C}_\iota(n_\iota)$ denotes time to compute values $\mathcal{R}_1, \Lambda_{1_1}, \Lambda_{2_1}$ on $T1$; $\mathcal{R}_2$ on $T2$; $\mathcal{R}_3, \Lambda_{1_3}$ on $T3$ and $\mathcal{R}_{1_{back}}$ on $T1back$.[4]
   - $\mathfrak{T}_F$ denotes time to build the sentence like $\min\{\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}\}$.
   - $\mathfrak{T}_{comp}$ denotes time to compute $\min\{\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}\}$.

The new computation time is:

$$\mathfrak{T}_{new} = \mathfrak{E}_I + \sum_{\iota \in I} \mathfrak{E}_\iota + \sum_{\iota \in I} \mathfrak{C}_\iota + \mathfrak{T}_F + \mathfrak{T}_{comp}.$$

The question is: When is $\mathfrak{T}_{new} < \mathfrak{T}_{old}$? The analysis is provided in Sect. 8. Moreover, we want the construction of the sentence like

$$\min\{\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}\}$$

depends only upon the property to be optimized and the predefined "communication" rules but **NOT** upon the given T. We will call them $F_{\Phi,\phi}$. Values $\mathcal{R}_1, \Lambda_{1_1} + \mathcal{R}_2, \Lambda_{2_1} + \mathcal{R}_3, \Lambda_{2_1} + \Lambda_{1_3} + \mathcal{R}_{1_{back}}$ will be referenced as evaluation of $\psi_{1,1}, \ldots, \psi_{1,j_1}, \ldots, \psi_{\beta,1}, \ldots, \psi_{\beta,j_\beta}$.

5 Disjoint Union and Shuffling of Structures

The first reduction technique that we use is *Feferman-Vaught reductions*. Feferman-Vaught reduction sequence (or simply, reduction) is a set of formulae such that

[4] n_ι is the size of the coding of T_ι.

each such a formula can be evaluated locally in some component or index structure. Next, from the local values, received from the components, and possibly some additional information about the components, we compute the value for the given global formula. In the logical context, the reductions are applied to relational structure $\mathcal{A}$ distributed over different components with structures $\mathcal{A}_i$, $i \in I$. The reductions allow the formulae over $\mathcal{A}$ be computed from formulae over the $\mathcal{A}_i$'s and formulae over the index structure $\mathcal{I}$.

In this section, we start to discuss different ways of obtaining structures from components. The *Disjoint Union* of a family of structures is the simplest example of juxtaposing structures over index structure $\mathcal{I}$ with universe I, where none of the components are linked to each other. In such a case, the index structure $\mathcal{I}$ may be replaced by index set I.

Definition 4 (*Disjoint Union*)
Let $\tau_i = \langle R_1^i, \ldots, R_{j^i}^i \rangle$ be a vocabulary of structure $\mathcal{A}_i$. In the general case, the resulting structure is $\mathcal{A} = \dot{\bigsqcup}_{i \in I} \mathcal{A}_i$, where

$$\mathcal{A} = \langle I \cup \dot{\bigcup}_{i \in I} A_i, P(\iota, v), Index(x), R_j^I (1 \leq j \leq j^I), R_{j^i}^i (i \in I, 1 \leq j^i \leq j^{i'}) \rangle$$

for all $i \in I$, where $P(i, v)$ is true iff element a came from A_i; $Index(x)$ is true iff x came from I.

Definition 5 (*Partitioned Index Structure*)
Let $\mathcal{I}$ be an index structure over τ_{ind}. $\mathcal{I}$ is called *finitely partitioned* into ℓ parts if there are unary predicates I_α, $\alpha < \ell$, in the vocabulary τ_{ind} of $\mathcal{I}$ such that their interpretation forms a partition of the universe of $\mathcal{I}$.

Using *Ehrenfeucht-Fraïssé* games for *MSOL*, cf. [35], it is easy to see that

Theorem 2 *Let $\mathcal{I}, \mathcal{J}$ be two (not necessary finitely) partitioned index structures over the same vocabulary such that for $i, j \in I_\ell$ and $i', j' \in J_\ell$, $\mathfrak{A}_i$ and $\mathfrak{A}_j$ ($\mathfrak{B}_{i'}$ and $\mathfrak{B}_{j'}$) are isomorphic.*

1. *If $\mathcal{I} \equiv^n_{MSOL} \mathcal{J}$, and $\mathfrak{A}_i \equiv^n_{MSOL} \mathfrak{B}_i$ then $\dot{\cup}_{i \in I} \mathfrak{A}_i \equiv^n_{MSOL} \dot{\cup}_{j \in J} \mathfrak{B}_j$.*
2. *If $\mathcal{I} \equiv^n_{MSOL} \mathcal{J}$, and $\mathfrak{A}_i \equiv^n_{FOL} \mathfrak{B}_i$ then $\dot{\cup}_{i \in I} \mathfrak{A}_i \equiv^n_{FOL} \dot{\cup}_{j \in J} \mathfrak{B}_j$.*

If, as in most our applications, there are only finitely many different components, we can prove a stronger statement, dealing with formulae rather than theories.

Theorem 3 *Let $\mathcal{I}$ be a finitely partitioned index structure. Let $\mathcal{A} = \dot{\bigsqcup}_{i \in I} \mathcal{A}_i$ be a τ–structure, where each $\mathcal{A}_i$ is isomorphic to some $\mathcal{B}_1, \ldots, \mathcal{B}_\ell$, over the vocabularies $\tau_1, \ldots, \tau_\ell$, in accordance to the partition (ℓ is the number of the classes). For every $\phi \in MSOL(\tau)$ there are:*

- *a boolean function $F_\phi(b_{1,1}, \ldots, b_{1,j_1}, \ldots, b_{\ell,1}, \ldots, b_{\ell,j_\ell}, b_{I,1}, \ldots, b_{I,j_I})$,*
- *MSOL–formulae $\psi_{1,1}, \ldots, \psi_{1,j_1}, \ldots, \psi_{\ell,1}, \ldots, \psi_{\ell,j_\ell}$,*
- *MSOL–formulae $\psi_{I,1}, \ldots, \psi_{I,j_I}$,*

such that for every $\mathcal{A}$, $\mathcal{I}$ and $\mathcal{B}_i$ as above with $\mathcal{B}_i \models \psi_{i,j}$ iff $b_{i,j} = 1$ and $\mathcal{B}_I \models \psi_{I,j}$ iff $b_{I,j} = 1$ we have

$$\mathcal{A} \models \phi \text{ iff } F_\phi(b_{1,1}, \ldots, b_{1,j_1}, \ldots, b_{\ell,1}, \ldots, b_{\ell,j_\ell}, b_{I,1}, \ldots, b_{I,j_I}) = 1.$$

Moreover, F_ϕ and the $\psi_{i,j}$ are computable from ϕ, ℓ and vocabularies alone, but the number of $\psi_{i,j}$ is tower exponential in the quantifier rank of ϕ.[5]

Proof By analyzing the proof of Theorem 2 and careful but rather routine tedious book keeping. See, in particular, [15].

Now, we introduce an abstract preservation property of *XX*-combination of logics $\mathcal{L}_1, \mathcal{L}_2$, denoted by $XX - PP(\mathcal{L}_1, \mathcal{L}_2)$. *XX* may mean, for example, Disjoint Union. The property says roughly that if two *XX*-combinations of structures $\mathfrak{A}_1, \mathfrak{A}_2$ and $\mathfrak{B}_1, \mathfrak{B}_2$ satisfy the same sentences of $\mathcal{L}_1$; then the disjoint unions $\mathfrak{A}_1 \sqcup \mathfrak{A}_2$ and $\mathfrak{B}_1 \sqcup \mathfrak{B}_2$ satisfy the same sentences of $\mathcal{L}_2$.

The reason, we look at this abstract property, is that it can be proven for various logics using their associated pebble games. The proofs usually depend on the detail on the particular pebble games. However, the property $XX - PP(\mathcal{L}_1, \mathcal{L}_2)$ and its variants play an important role in our development of the Feferman-Vaught style theorems. This abstract approach was initiated by [38] and further developed in [73, 74].

Now, we spell out various ways in which the theory of a disjoint union depends on the theory of the components. We first look at the case where the index structure is fixed.

Definition 6 (*Preservation Properties with Fixed Index Set*) For two logics $\mathcal{L}_1$ and $\mathcal{L}_2$ we define:

Input of operation: Indexed set of structures;
Preservation Property: if for each $i \in I$ (index set) $\mathfrak{A}_i$ and $\mathfrak{B}_i$ satisfy the same sentences of $\mathcal{L}_1$ then the disjoint unions $\bigsqcup_{i\in I} \mathfrak{A}_i$ and $\bigsqcup_{i\in I} \mathfrak{B}_i$ satisfy the same sentences of $\mathcal{L}_2$.
Notation: $DJ - PP(\mathcal{L}_1, \mathcal{L}_2)$.

Disjoint Pair

Input of operation: Two structures;
Preservation Property: if two pairs of structures $\mathfrak{A}_1, \mathfrak{A}_2$ and $\mathfrak{B}_1, \mathfrak{B}_2$ satisfy the same sentences of $\mathcal{L}_1$ then the disjoint unions $\mathfrak{A}_1 \sqcup \mathfrak{A}_2$ and $\mathfrak{B}_1 \sqcup \mathfrak{B}_2$ satisfy the same sentences of $\mathcal{L}_2$.
Notation: $P - PP(\mathcal{L}_1, \mathcal{L}_2)$.

Disjoint Union

Input of operation: Indexed set of structures;

[5] However, in practical applications, the complexity, as a rule, is simply exponential.

Preservation Property: if for each $i \in I$ (index set) $\mathfrak{A}_i$ and $\mathfrak{B}_i$ satisfy the same sentences of $\mathcal{L}_1$ then the disjoint unions $\bigsqcup_{i\in I} \mathfrak{A}_i$ and $\bigsqcup_{i\in I} \mathfrak{B}_i$ satisfy the same sentences of $\mathcal{L}_2$.
Notation: $DJ - PP(\mathcal{L}_1, \mathcal{L}_2)$.

The *Disjoint Union* of a family of structures is the simplest example of juxtaposing structures where none of the components are linked to each other. Another way of producing a new structure from several given structures is by mixing (shuffling) structures according to a (definable) prescribed way along the index structure.

Definition 7 (*Shuffle over Partitioned Index Structure*) Let $\mathcal{I}$, be a partitioned index structure into β parts, using unary predicates I_α, $\alpha < \beta$. Let $\mathcal{A}_i$, $i \in I$ be a family of structures such that, for each $i \in I_\alpha$, it holds: $\mathcal{A}_i \cong \mathcal{B}_\alpha$, according to the partition. In this case, we say that $\bigsqcup_{i\in I} \mathcal{A}_i$ is the *shuffle of* $\mathcal{B}_\alpha$ *along the partitioned index structure* $\mathcal{I}$, and denote it by $\biguplus^{\mathcal{I}}_{\alpha<\beta} \mathcal{B}_\alpha$.

Note that the shuffle operation, as defined here, is a special case of the disjoint union, and that the disjoint pair is a special case of the finite shuffle.

In the case of variable index structures and of *FOL*, Feferman and Vaught observed that it is not enough to look at the *FOL*-theory of the index structures, but one has to look at the *FOL*-theories of expansions of the Boolean algebras $PS(\mathcal{I})$ and $PS(\mathcal{J})$, respectively, where $PS(X)$ denotes the power set of X. Gurevich suggested another approach, by looking at the *MSOL* theories of structures $\mathcal{I}$ and $\mathcal{J}$. This is really the same, but more in the spirit of the problem, as the passage from I to an expansion of $PS(\mathcal{I})$ remains on the semantic level, whereas the comparison of theories is syntactic. There is not much freedom in choosing the logic in which to compare the index structures, so we assume it always to be *MSOL*.

Definition 8 (*Preservation Properties with Variable Index Structures*)
For two logics $\mathcal{L}_1$ and $\mathcal{L}_2$ we define:

Disjoint Multiples

Input of operation: Structure and Index structure;
Preservation Property: Given two pairs of structures $\mathfrak{A}$, $\mathfrak{B}$ and $\mathcal{I}$, $\mathcal{J}$ such that $\mathfrak{A}, \mathfrak{B}$ satisfy the same sentences of $\mathcal{L}_1$ and $\mathcal{I}, \mathcal{J}$ satisfy the same *MSOL*-sentences. Then the disjoint unions $\bigsqcup_{i\in I} \mathfrak{A}$ and $\bigsqcup_{j\in J} \mathfrak{B}$ satisfy the same sentences of $\mathcal{L}_2$.
Notation: $Mult - PP(\mathcal{L}_1, \mathcal{L}_2)$.

Shuffles

Input of operation: A family of structures $\mathfrak{B}_\alpha : \alpha < \beta$ and a (finitely) partitioned index structure $\mathcal{I}$ with I_α a partition.
Preservation Property: Assume that for each $\alpha < \beta$ the pair of structures $\mathfrak{A}_\alpha, \mathfrak{B}_\alpha$ satisfy the same sentences of $\mathcal{L}_1$, and $\mathcal{I}, \mathcal{J}$ satisfy the same *MSOL*-sentences. Then the schuffles $\biguplus^{\mathcal{I}}_{\alpha<\beta} \mathfrak{A}_\alpha$ and $\biguplus^{\mathcal{J}}_{\alpha<\beta} \mathfrak{B}_\alpha$ satisfy the same sentences of $\mathcal{L}_2$.

Notation: $Shu - PP(\mathcal{L}_1, \mathcal{L}_2)$ $(FShu - PP(\mathcal{L}_1, \mathcal{L}_2))$.

Here, we provide some observations:

Observation 1 *Assume that for two logics $\mathcal{L}_1$, $\mathcal{L}_2$ we have the preservation property $XX - PP(\mathcal{L}_1, \mathcal{L}_2)$ and $\mathcal{L}_1'$ is an extension of $\mathcal{L}_1$, $\mathcal{L}_2'$ is a sub-logic of $\mathcal{L}_2$, then $XX - PP(\mathcal{L}_1', \mathcal{L}_2')$ holds as well.*

Observation 2 *For two logics $\mathcal{L}_1$, $\mathcal{L}_2$ the following implications between preservation properties hold:*

1. *$DJ - PP(\mathcal{L}_1, \mathcal{L}_2)$ implies $P - PP(\mathcal{L}_1, \mathcal{L}_2)$ and, for fixed index structures, $Mult - PP(\mathcal{L}_1, \mathcal{L}_2)$, $Shu - PP(\mathcal{L}_1, \mathcal{L}_2)$ and $FShu - PP(\mathcal{L}_1, \mathcal{L}_2)$.*
2. *For variable index structures, we have: $Shu - PP(\mathcal{L}_1, \mathcal{L}_2)$ implies $FShu - PP(\mathcal{L}_1, \mathcal{L}_2)$ and $Mult - PP(\mathcal{L}_1, \mathcal{L}_2)$.*

Definition 9 (*Reduction Sequence for Shuffling*)
Let $\mathcal{I}$ be a finitely partitioned τ_{ind}-index structure and $\mathcal{L}$ be logic.
Let $\mathcal{A} = \biguplus_{\alpha<\beta}^{\mathcal{I}} \mathcal{B}_\alpha$ be the τ–structure which is the finite shuffle of the τ_α-structures $\mathcal{B}_\alpha$ over $\mathcal{I}$. *$\mathcal{L}_1$-reduction sequence for shuffling* for $\phi \in \mathcal{L}_2(\tau_{shuffle})$ is given by

1. a boolean function $F_\phi(b_{1,1}, \ldots, b_{1,j_1}, \ldots, b_{\beta,1}, \ldots, b_{\beta,j_\beta}, b_{I,1}, \ldots, b_{I,j_I})$,
2. set Υ of $\mathcal{L}_1$–formulae $\Upsilon = \{\psi_{1,1}, \ldots, \psi_{1,j_1}, \ldots, \psi_{\beta,1}, \ldots, \psi_{\beta,j_\beta}\}$,
3. *MSOL*–formulae $\psi_{I,1}, \ldots, \psi_{I,j_I}$,

and has the property that for every $\mathcal{A}, \mathcal{I}$ and $\mathcal{B}_\alpha$ as above with $\mathcal{B}_\alpha \models \psi_{\alpha,j}$ iff $b_{\alpha,j} = 1$ and $\mathcal{B}_I \models \psi_{I,j}$ iff $b_{I,j} = 1$ we have

$$\mathcal{A} \models \phi \text{ iff } F_\phi(b_{1,1}, \ldots, b_{1,j_1}, \ldots, b_{\beta,1}, \ldots, b_{\beta,j_\beta}, b_{I,1}, \ldots, b_{I,j_I}) = 1.$$

Note that we require that F_ϕ and formulae $\psi_{\alpha,j}$ depend only on ϕ, β and $\tau_1, \ldots, \tau_\beta$, but not on the structures involved.

Remark 1 If $\mathcal{I}$ is finite and fixed, then the *MSOL*–formulae $\psi_{I,1}, \ldots, \psi_{I,j_I}$ in the reduction sequences can be hidden in the function F.

In many applications, we want to keep the quantifier rank $m = rank(\phi)$ or the number of variables $k = var(\phi)$ of the formulae ϕ under control. Let $\mathcal{L}^{m,k}(\tau) \subseteq \mathcal{L}(\tau)$ be the set of formulae of quantifier rank $\leq m$ and total number of variables $\leq k$. Frequently, a property $XX - PP(\mathcal{L}_1, \mathcal{L}_2)$ holds uniformly also for subsets of formulae of bounded quantifier rank and total number of variables, i.e. $XX - PP(\mathcal{L}_1^{m_1,k_1}, \mathcal{L}_2^{m_2,k_2})$ holds, where m_1 is a simple function of m_2 and k_1 is a simple function of k_2. A strong form of such a uniform version of $XX - PP(\mathcal{L}_1^{m_1,k_1}, \mathcal{L}_2^{m_2,k_2})$ is given by Theorem 4.

For our various notions of combinations, preservation theorems can be proven by suitable Pebble games, which are generalizations of *Ehrenfeucht-Fraïssé* games. This gives usually a version of $XX - PP(\mathcal{L}_1^{m_1,k_1}, \mathcal{L}_2^{m_2,k_2})$ for suitable chosen definitions of quantifier rank and counting of variables.

Usually, a close examination of the winning strategies for Pebble games gives more: we can get an algorithm, which, for each ϕ, produces the corresponding reduction sequence. However, for several logics, the preservation theorems can be proven directly, and the proofs give the transparent way to build the corresponding reduction sequence, cf. [15]. Such a proof for Theorem 2 is given in [101]; see also [84] and [15]. Now, we list which Preservation Properties hold for which logics.

Theorem 4 *Let $\mathcal{I}$ be an index structure and $\mathcal{L}$ be any of FOL, $FOL^{m,k}$, $L^{\omega}_{\omega_1,\omega}$, $L^{k}_{\omega_1,\omega}$, $MSOL^m$, MTC^m, $MLFP^m$, or $FOL[\mathbf{Q}]^{m,k}$ ($L_{\omega_1,\omega}[\mathbf{Q}]^k$) with unary generalized quantifiers. Then $DJ - PP(\mathcal{L},\mathcal{L})$ and $FShu - PP(\mathcal{L},\mathcal{L})$ hold . This includes $DJ - PP(FOL^{m,k}, FOL^{m,k})$ and $FShu - PP(FOL^{m,k}, FOL^{m,k})$ with the same bounds for both arguments, and similarly for the other logics.*

Proof We first list the cases known from the literature.

FOL and $\mathbf{FOL^{m,k}}$: The proofs for *FOL* and *MSOL* are classical, see in particular [15]. Extension for $FOL^{m,k}$ can be done directly from the proof for *FOL*.
MLFP and $\mathbf{MLFP^m}$: The proof for *MLFP* was given in [11].
$\mathbf{L_{!_1,!}(Q)^k}$: The proof was given in [24].

Our **original** proof for MTC^m is explicitly provided in Sect. 13.

Theorem 5 *Let $\mathcal{L}$ be any of FOL, $FOL^{m,k}$, $L^{\omega}_{\omega_1,\omega}$, $L^{k}_{\omega_1,\omega}$ $MSOL^m$, MTC^m, $MLFP^m$, or $FOL[\mathbf{Q}]^{m,k}$ with unary generalized quantifiers. There is an algorithm, which for given $\mathcal{L}$, τ_{ind}, τ_α, $\alpha < \beta$, $\tau_{shuffle}$ and $\phi \in \mathcal{L}(\tau_{shuffle})$ produces a reduction sequence for ϕ for $(\tau_{ind}, \tau_{shuffle})$-shuffling. However, F_ϕ and the $\psi_{\alpha,j}$ are tower exponential in the quantifier rank of ϕ and F depends on the MSOL–theory of the index structure restricted to the same quantifier rank as ϕ.*[6]

Proof By analyzing the proof of Theorem 4. A special case was analyzed in Gurevich's work [46].

We finally show that our restriction to unary generalized quantifiers (*MTC* and *MLFP*) is necessary.

Proposition 1 *Theorem 4 does not hold for* 2*-TC or* 2*-LFP.*

Proof Let $I = \{0, 1\}$ and let the components be finite linear orders. Using a counting argument, it is easy to produce arbitrary large pairs of linear orders $\mathcal{A}_0$, $\mathcal{A}_1$, which are $2\text{-}TC - m$–equivalent, but of different cardinalities. Now consider the structure $\mathcal{B}_0 = \mathcal{A}_0 \sqcup \mathcal{A}_0$ and $\mathcal{B}_1 = \mathcal{A}_0 \sqcup \mathcal{A}_1$. The $2\text{-}TC$–formula, which distinguishes $\mathcal{B}_0$ from $\mathcal{B}_1$ is the formula θ, which asserts that the two components have the same cardinality. θ can be written as

$$2\text{-}TC x_0, x_1, y_0, y_1; first_0, first_1, last_0, last_1 (succ_0(x_0, y_0) \wedge succ_1(x_1, y_1)),$$

[6] Note that this includes reduction sequences for disjoint pairs and disjoint multiples.

where $succ_i$ is the *FOL* formula, expressing the successor in the ith component, and $first_i$, $last_i$ are the constant symbols, which are interpreted by the first, respectively, the last element in the ith component.

Now, we discuss various ways of obtaining weighted labeled trees from components, as in [94].

Definition 10 (*Finite Disjoint Union of Weighted Labeled Trees*)
Let $\tau_i = \langle label^{\tau}_{a\,i}, edge^{\tau}_{i\,i}\rangle$, be a vocabulary of a weighted labeled tree $\mathcal{T}_i$ over Σ. In the general case, the tree over $\Sigma \cup I$ is

$$\mathcal{T} = \dot{\bigsqcup}_{i\in I}\mathcal{T}_i = \langle \dot{\bigcup}_{i\in I}\mathcal{B}_i, D; label_i(u)(i \in I), label^{\tau}_{a\,i}(i \in I), edge^{\tau}_{i\,\iota}(\iota \in I)\rangle$$

for all $i \in I$, where $label_i(u)$ is true iff u came from $\mathcal{B}_i$, I is finite and each element in I is of rank 1.

The following theorem can be stated, cf. [72, 94]:

Theorem 6 *Let I be a finite index set with ℓ elements. Let $\mathcal{T} = \dot{\bigsqcup}_{i\in I}\mathcal{T}_i$ be a weighted labeled tree. Then for every $\varphi \in WMSOL(\tau)$ over boolean semi-rings, there are:*

– *a computation over weighted WMSOL formulae*

$$F_\varphi(\varpi_{1,1}, \ldots, \varpi_{1,j_1}, \ldots, \varpi_{\ell,1}, \ldots, \varpi_{\ell,j_\ell}),\ \textit{and}$$

– *WMSOL–formulae $\psi_{1,1}, \ldots, \psi_{1,j_1}, \ldots, \psi_{\ell,1}, \ldots, \psi_{\ell,j_\ell}$,*

such that for every $\mathcal{T}_i$ and I as above with $\varpi_{i,j} = \varrho_{i,j}$ iff $[\psi_{i,j}] = \varrho_{i,j}$, we have

$$[\varphi] = \varrho \textit{ iff } F_\varphi(\varpi_{1,1}, \ldots, \varpi_{1,j_1}, \ldots, \varpi_{\ell,1}, \ldots, \varpi_{\ell,j_\ell}) = \varrho.$$

Moreover, F_φ and the $\psi_{i,j}$ are computable from φ, ℓ and vocabularies alone, but are tower exponential in the quantifier rank of φ.

We list also some other options of commutative semi-rings to choose as follows:

Theorem 7 *In addition, the following semi-rings satisfy Theorem 6:*

- *Subset Semi-ring:* $(PS(A), \cap, \cup, \emptyset, A)$. *The proof by analyzing and extension of the proof in [38].*
- *Fuzzy Semi-ring:* $([0, 1], \vee, \wedge, 0, 1)$. *The proof by analyzing and extension of the proof in [72].*
- *Extended natural number:* $(\mathbf{N} \cup \{\infty\}, +, \cdot, 0, 1)$. *The proof by analyzing and extension of the proof in [72].*
- *Tropical Semi-ring:* $(\mathbf{R}_+ \cup \{+\infty\}, \min, +, +\infty, 0)$, *cf. [94].*
- *Arctic Semi-ring:* $(\mathbf{R}_+ \cup \{-\infty\}, \max, +, -\infty, 0)$. *The proof by analyzing and extension of the proof in [94].*

Therewith, Theorem 7 opens the door to incremental reasoning on fuzzy distributed systems.

6 Syntactically Defined Translation Schemes

The second logical reduction technique that we use is *the syntactically defined translation schemes*, which describe transformations of logical structures. The notion of abstract translation schemes comes back to Rabin, cf. [91]. They give rise to two induced maps, translations and transductions. Transductions describe the induced transformation of logical structures and the translations describe the induced transformations of logical formulae.

Definition 11 (*Translation Schemes Φ*) Let τ_1 and τ_2 be two vocabularies and $\mathcal{L}$ be a logic. Let $\tau_2 = \{R_1, \ldots, R_m\}$ and let $\rho(R_i)$ be the arity of R_i. Let $\Phi = \langle \varphi, \psi_1, \ldots, \psi_m \rangle$ be formulae of $\mathcal{L}(\tau_1)$. Then Φ is named *κ–feasible for τ_2 over τ_1* if φ has exactly κ distinct free variables and each ψ_i has $\kappa\rho(R_i)$ distinct free variables. Such a $\Phi = \langle \varphi, \psi_1, \ldots, \psi_m \rangle$ is also called a *κ–τ_1–τ_2–translation scheme* or, shortly, a *translation scheme*, if the parameters are clear in the context. If $\kappa = 1$ we speak of *scalar* or *non–vectorized* translation schemes.

The above definition as a rule assumes *one sorted* logical structures. However, if we deal with weighted logics like *WMSOL* this is not a case. In general, if $\mathcal{L}$ is defined over particular kinds of $\mathcal{L}$-objects (like graphs, words, etc.), then the exact logical presentation of the objects must be explicitly provided. *WMSOL* is defined over the weighted labeled trees and we present them as logical structures in the following way:

Definition 12 (*Weighted Labeled Tree over a ptv-monoid D*) Given a ptv-monoid D, the weighted labeled tree over D is the following logical many-sorted structure $\mathcal{T} = \langle \mathcal{B}, D; label_a, edge_i \rangle$, where

Universe of $\mathcal{T}$ is *many-sorted*: $\mathcal{B}$ is the tree domain and D comes from the monoid.

Relations of $\mathcal{T}$:

$-label_a$ is a unary relation, which for each $a \in \Sigma$ means that $u \in \mathcal{B}$ is labeled by a, and

$-edge_i$ is a binary relation, which for each $1 \leq i \leq \max_\Sigma$ and two $u_1, u_2 \in \mathcal{B}$ means that u_2 is an immediate prefix of u_1.

In the context of *WMSOL*, Definition 11 may be paraphrased as follows:

Definition 13 (*Translation Schemes Φ_D on Weighted Labeled Trees*)
Let τ_1 and τ_2 be two vocabularies of weighted labeled trees.
Let $\tau_1 = \langle label_a^{\tau_1}, edge_i^{\tau_1} \rangle$, over ptv-monoid D.
Let $\Phi = \langle \phi_\mathcal{B}, \phi_D; \psi_{label_a}, \psi_{edge_i} \rangle$ be almost boolean WMSOL formulae (for each $a \in \Sigma$ and $1 \leq i \leq \max_\Sigma$).
We say that Φ_D is *feasible for τ_2 over τ_1*, if

– $\phi_\mathcal{B}$ has exactly 1 distinct free first order variable over $\mathcal{B}$,

– ϕ_D is a tautology with exactly one free variable over D,
– each ψ_{label_a} has exactly 1 distinct free first order variable over $\mathcal{B}$,
– each ψ_{edge_i} has exactly 2 distinct free first order variables over $\mathcal{B}$.

In general, Definition 11 must be adopted to the given logic $\mathcal{L}$, if it is not straightforward.

For $\mathcal{L}$ like *FOL*, *MSOL*, *MTC*, *MLFP* or *FOL* with unary generalized quantifiers, with a translation scheme Φ we can *naturally* associate a (partial) function Φ^* from τ_1–structures to τ_2–structures.

Definition 14 (*The induced map* Φ^*)
Let $\mathcal{A}$ be a τ_1–structure with universe A and Φ be κ–feasible for τ_2 over τ_1. The structure $\mathcal{A}_\Phi$ is defined as follows:

1. The universe of $\mathcal{A}_\Phi$ is the set $A_\Phi = \{\bar{a} \in A^\kappa : \mathcal{A} \models \varphi(\bar{a})\}$.
2. The interpretation of R_i in $\mathcal{A}_\Phi$ is the set

$$A_\Phi(R_i) = \{\bar{a} \in {A_\Phi}^{\rho(R_i)\cdot\kappa} : \mathcal{A} \models \psi_i(\bar{a})\}.$$

Note that $\mathcal{A}_\Phi$ is a τ_2–structure of cardinality at most $|A|^\kappa$.
3. The partial function $\Phi^* : Str(\tau_1) \to Str(\tau_2)$ is defined by $\Phi^*(\mathcal{A}) = \mathcal{A}_\Phi$. Note that $\Phi^*(\mathcal{A})$ is defined iff $\mathcal{A} \models \exists\bar{x}\varphi$.

The case of *WMSOL* is a bit more tricky. The (partial) function Φ_D^* from τ_1–trees to τ_2–trees is subsequently defined:

Definition 15 (*The induced map* Φ_D^*)
Let $\mathcal{T}^{\tau_1}$ be a τ_1–tree and Φ_D be feasible for τ_2 over τ_1. The structure $\mathcal{T}^{\tau_2}_{\Phi_D}$ is defined as follows:

- The many-sorted universe $\mathcal{B}, D$: are the sets:
  1. $\mathcal{B}_{\Phi_D} = \{u \in \mathcal{B}^{\tau_1} : \mathcal{T}^{\tau_1} \models \phi_{\mathcal{B}}(u)\}$;
  2. $D_{\Phi_D} = D$.
- Relations $label_a^{\tau_2}, edge_i^{\tau_2}$: For each $a \in \Sigma$ and $1 \le i \le \max_\Sigma$:
  1. The interpretation of each $label_a$ in ${\mathcal{T}^{\tau_2}}_{\Phi_D}$ is the set

$${\mathcal{T}^{\tau_2}}_{\Phi_D}(label_a) = \{u \in \mathcal{B}^{\tau_1} : \mathcal{T}^{\tau_1} \models \psi_{label_a}(u)\};$$

  2. The interpretation of each $edge_i$ in is the set of pairs

$${\mathcal{T}^{\tau_2}}_{\Phi_D}(edge_i) = \{(u_1, u_2) \in {\mathcal{B}^{\tau_1}}^2 : \mathcal{T}^{\tau_1} \models \psi_{edge_i}(u_1, u_2)\};$$

- ${\Phi_D}^*$: The partial function ${\Phi_D}^* : Trees(\tau_1) \to Trees(\tau_2)$ is defined by

$${\Phi_D}^*(\mathcal{T}^{\tau_1}) = {\mathcal{T}^{\tau_2}}_{\Phi_D}.$$

Note that $\Phi_D{}^*(\mathcal{T}^{\tau_1})$ is defined iff $\mathcal{T}^{\tau_1} \models \phi_{\mathcal{B}}(u)$.

Again, for $\mathcal{L}$ like *FOL*, *MSOL*, *MTC*, *MLFP* or *FOL* with unary generalized quantifiers, with a translation scheme Φ we can also *naturally* associate a function $\Phi^{\#}$ from $\mathcal{L}(\tau_2)$–formulae to $\mathcal{L}(\tau_1)$–formulae.

Definition 16 (*The induced map* $\Phi^{\#}$)
Let θ be a τ_2–formula and Φ be κ–feasible for τ_2 over τ_1. The formula θ_Φ is defined inductively as follows:

1. For $R_i \in \tau_2$ and $\theta = R(x_1, \ldots, x_m)$ let $x_{j,h}$ be new variables with $i \leq m$ and $h \leq \kappa$ and denote by $\bar{x}_i = \langle x_{i,1}, \ldots, x_{i,\kappa}\rangle$. We put $\theta_\Phi = \psi_i(\bar{x}_1, \ldots, \bar{x}_m)$.
2. For the boolean connectives the translation distributes; i.e., if $\theta = (\theta_1 \vee \theta_2)$, then $\theta_\Phi = (\theta_{1\Phi} \vee \theta_{1\Phi})$ and if $\theta = \neg\theta_1$ then $\theta_\Phi = \neg\theta_{1\Phi}$, and similarly for $\wedge$.
3. For the existential quantifier, we use relativization; i.e., if $\theta = \exists y\theta_1$, let $\bar{y} = \langle y_1, \ldots, y_\kappa\rangle$ be new variables. We put $\theta_\Phi = \exists\bar{y}(\varphi(\bar{y}) \wedge \theta_{1\Phi})$.
4. For (monadic) second order variables U of arity ℓ ($\ell = 1$ for *MSOL*) and $\bar{v}$ being a vector of length ℓ of first order variables or constants we translate $U(\bar{v})$ by treating U like a relation symbol above and put

$$\theta_\Phi = \exists V(\forall\bar{v}(V(\bar{v}) \rightarrow (\phi(\bar{v_1}) \wedge \ldots \phi(\bar{v_\ell}) \wedge (\theta_1)_\Phi))).$$

5. For generalized quantifiers, if $\theta = Q_i v^1, v^2, \ldots, v^m\theta_1(v^1, v^2, \ldots, v^m, \ldots)$, then let $\bar{v}^j = \langle v^j_1, \ldots, v^j_k\rangle$ be new variables for v^j. We make

$$\theta_\Phi = Q_i\bar{v}^1, \bar{v}^2, \ldots, \bar{v}^m(\theta_1(\bar{v}^1, \bar{v}^2, \ldots, \bar{v}^m, \ldots)_\Phi).$$

6. For infinitary logics, if $\theta = \bigwedge\Psi$, then $\theta_\Phi = \bigwedge\Psi_\Phi$.
7. For *LFP*, if $\theta = n\text{-}LFP\bar{x}, \bar{y}, \bar{u}, \bar{v}\theta_1$, then $\theta_\Phi = (n \cdot \kappa)\text{-}LFP\bar{x}, \bar{y}, \bar{u}, \bar{v}\theta_{1\Phi}$.
8. For *TC*, if $\theta = n\text{-}TC\bar{x}, \bar{y}, \bar{u}, \bar{v}\theta_1$, then $\theta_\Phi = (n \cdot \kappa)\text{-}TC\bar{x}, \bar{y}, \bar{u}, \bar{v}\theta_{1\Phi}$.
9. For weighted formulae over ptv-monoid D:
   (a) for d we do nothing;
   (b) for a boolean formula β we put $\zeta_{\Phi_D} = \beta$;
   (c) for boolean connectives and quantifiers the translation distributes.
10. The function $\Phi^{\#} : \mathcal{L}(\tau_2) \rightarrow \mathcal{L}(\tau_1)$ is defined by $\Phi^{\#}(\theta) = \theta_\Phi$.

Note that the case of weighted formulae over ptv-monoid D is the most complicated in the definition. In general, given $\mathcal{L}$, Definition 16 must be adopted to all well-formed formulae of the logic.

Observation 3 *If we use MSOL and* Φ^* *is over MSOL too, and it is vectorized, then we do not obtain MSOL for* $\mathcal{A}_\Phi$. *However, in most of feasible applications, we have that* Φ^* *is not vectorized, but not necessarily; cf. [6, 71].*

Observation 4 *1.* $\Phi^{\#}(\theta) \in FOL$ (SOL, TC, LFP) *provided* $\theta \in FOL$ (SOL, TC, LFP), *even for vectorized* Φ.

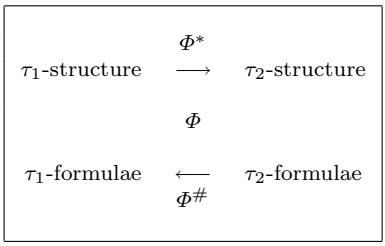


Fig. 5 Translation scheme and its components

2. $\Phi^{\#}(\theta) \in$ *MSOL provided* $\theta \in$ *MSOL, but only for scalar (non–vectorized)* Φ.
3. $\Phi^{\#}(\theta) \in$ *nk-TC(nk-LFP) provided* $\theta \in$ *n-TC(n-LFP) and* Φ *is a k–feasible.*
4. $\Phi^{\#}(\theta) \in TC^{kn}(LFP^{kn}, L^{kn}_{\infty\omega})$ *provided* $\theta \in TC^{n}(LFP^{n}, L^{n}_{\infty\omega})$ *and* Φ *is a k–feasible.*

The following fundamental theorem is easily verified for correctly defined $\mathcal{L}$ translation schemes, see Fig. 5. Its origins go back at least to the early years of modern logic; cf. [49, page 277 ff.] and see also [33].

Theorem 8 *Let* $\Phi = \langle \varphi, \psi_1, \ldots, \psi_m \rangle$ *be a* κ–τ_1–τ_2–*translation scheme,* $\mathcal{A}$ *a* τ_1-*structure and* θ *a* $\mathcal{L}(\tau_2)$–*formula. Then*

$$\mathcal{A} \models \Phi^{\#}(\theta) \text{ iff } \Phi^{*}(\mathcal{A}) \models \theta.$$

7 Strongly Distributed Systems

The disjoint union and shuffles as such are not very interesting. However, combining it with translation schemes gives as a much richer repertoire of composition techniques. Let τ_0, τ_1, τ be finite vocabularies. For a τ_0-model $\mathcal{I}$ (serving as index model), τ_1-structures are pairwise disjoint for simplicity $\mathcal{A}_i$ $(i \in I)$ and a τ-structure $\mathcal{A}$ is the disjoint union of $\langle \mathcal{A}_i : i \in I \rangle$ with $\mathcal{A} = \dot{\bigsqcup}_{i \in I} \mathcal{A}_i$. Now, we generalize the disjoint union or shuffling of structures to *strongly distributed systems* in the following way.

Definition 17 (*Strongly Distributed Systems*)
Let $\mathcal{I}$ be a finitely partitioned index structure and $\mathcal{L}$ be any like *FOL*, *MSOL*, *WMSOL*, *MTC*, *MLFP*, or *FOL* with unary generalized quantifiers. Let $\mathcal{A} = \bigsqcup_{i \in I} \mathcal{A}_i$ be a τ–structure, where each $\mathcal{A}_i$ is isomorphic to some $\mathcal{B}_1, \ldots, \mathcal{B}_\beta$ over the vocabularies $\tau_1, \ldots, \tau_\beta$, in accordance with the partition. For a Φ be a scalar (non–vectorized) τ_1–τ_2 $\mathcal{L}$–translation scheme, the Φ–*Strongly Distributed System*, composed from $\mathcal{B}_1, \ldots, \mathcal{B}_\beta$ over $\mathcal{I}$ is the structure $\mathcal{S} = \Phi^{*}(\mathcal{A})$, or rather any structure isomorphic to it.

The above definition is pretty general. If $\mathcal{A} = \bigsqcup_{i \in I} \mathcal{A}_i$ models a multi-agent system, then we are talking about a *strongly distributed multi-agent system.*

Now, our main theorem can be formulated as follows:

Theorem 9 *Let $\mathcal{I}$ be a finitely partitioned index structure, $\mathcal{L}$ be any of FOL, MSOL, MTC, MLFP, MSOL or FOL with unary generalized quantifiers. Let $\mathcal{S}$ be a Φ–Strongly Distributed System, composed from $\mathcal{B}_1,\ldots,\mathcal{B}_\beta$ over $\mathcal{I}$, as above. For every $\phi \in \mathcal{L}(\tau)$ there are:*

1. *a boolean function $F_{\Phi,\phi}(b_{1,1},\ldots,b_{1,j_1},\ldots,b_{\beta,1},\ldots,b_{\beta,j_\beta},b_{I,1},\ldots,b_{I,j_I})$,*
2. *$\mathcal{L}$–formulae $\psi_{1,1},\ldots,\psi_{1,j_1},\ldots,\psi_{\beta,1},\ldots,\psi_{\beta,j_\beta}$, and*
3. *MSOL–formulae $\psi_{I,1},\ldots,\psi_{I,j_I}$,*

such that for every $\mathcal{S}$, $\mathcal{I}$ and $\mathcal{B}_i$ as above with $\mathcal{B}_i \models \psi_{i,j}$ iff $b_{i,j} = 1$ and $\mathcal{I} \models \psi_{I,j}$ iff $b_{I,j} = 1$ we have

$$\mathcal{S} \models \phi \quad \text{iff} \quad F_{\Phi,\phi}(b_{1,1},\ldots,b_{1,j_1},\ldots,b_{\beta,1},\ldots,b_{\beta,j_\beta},b_{I,1},\ldots,b_{I,j_I}) = 1.$$

Moreover, $F_{\Phi,\phi}$ and $\psi_{i,j}$ are computable from $\Phi^{\#}$ and ϕ, but are tower exponential in the quantifier rank of ϕ.

Proof By analyzing the proof of Theorem 5 and using Theorem 8.

Moreover, in [94], the following was proven for *WMSOL*.

Theorem 10 *Let $\mathcal{I}$ be a finite index structure and let $\mathcal{T}$ be Φ-Strongly Distributed over $\mathcal{T}_1,\ldots,\mathcal{T}_\ell$, over I, as above. For every $\varphi \in WMSOL(\tau)$ that satisfies Theorem 6 there are:*

– *a computation over weighted formulae*

$$F_{\Phi,\varphi}(\varpi_{1,1},\ldots,\varpi_{1,j_1},\ldots,\varpi_{\ell,1},\ldots,\varpi_{\ell,j_\ell}), \text{ and}$$

– *WMSOL–formulae $\psi_{1,1},\ldots,\psi_{1,j_1},\ldots,\psi_{\ell,1},\ldots,\psi_{\ell,j_\ell}$,*

such that for every $\mathcal{T}_i$ and $\mathcal{I}$ as above with $\varpi_{i,j} = \varrho_{i,j}$ iff $[\psi_{i,j}] = \varrho_{i,j}$, we have

$$[\varphi] = \varrho \text{ iff } F_{\Phi,\varphi}(\varpi_{1,1},\ldots,\varpi_{1,j_1},\ldots,\varpi_{\ell,1},\ldots,\varpi_{\ell,j_\ell}) = \varrho.$$

Moreover, $F_{\Phi,\varphi}$ and $\psi_{i,j}$ are computable from $\Phi^{\#}$ and φ, but are tower exponential in the quantifier rank of φ.

8 Complexity Analysis

In this section, we discuss under what conditions our approach improves the complexity of computations, when measured in the size of the composed structures only. Our scenarios are as follows: A strongly distributed system is now submitted to a computation unit and we want to know, how long it takes to check whether ϕ is true on it. Now, we give the general complexity analysis of the computation on strongly distributed systems.

Assume that $\mathcal{A}$ is a strongly distributed system. Its components are $\mathcal{A}_\iota$ with index structure $\mathcal{I}$, and we want to check whether ϕ is true in $\mathcal{A}$. Assume that:

- $\mathfrak{T}(N)$ or $\mathfrak{T}_{old}(N)$ denotes the time to solve the problem by the traditional sequential way (N stands for the size of the coding of $\mathcal{A}$);
- $\mathfrak{E}_I$ denotes the time to extract index structure $\mathcal{I}$ from $\mathcal{A}$;
- $\mathfrak{E}_\iota$ denotes the time to extract each $\mathcal{A}_\iota$ from $\mathcal{A}$;
- $\mathfrak{C}_I(n_I)$ denotes the time to compute all values of $b_{I,J}$, where n_I is the size of I;
- $\mathfrak{C}_\iota(n_\iota)$ denotes the time to compute all values of $b_{\iota,J}$, where n_ι is the size of A_ι;
- $\mathfrak{T}_{F_{\Phi,\phi}}$ denotes the time to build $F_{\Phi,\phi}$;
- $\mathfrak{T}_S$ denotes the time to achieve one result of $F_{\Phi,\phi}$.

According to these symbols, the new computation time is:

$$\mathfrak{T}_{new} = \mathfrak{E}_I + \sum_{\iota \in I} \mathfrak{E}_\iota + \mathfrak{C}_I + \sum_{\iota \in I} \mathfrak{C}_\iota + \mathfrak{T}_{F_{\Phi,\phi}} + \mathfrak{T}_{\mathcal{S}}. \tag{1}$$

Now, the question to answer is: When does hold $\mathfrak{T}_{old} > \mathfrak{T}_{new}$.

8.1 Scenario A: Single Computation on Repetitive Structures

In many practical applications, all agents, their missions and environments are identical. In this section, we consider the complexity gain of our approach in this case.

8.1.1 Complexity Gain for *MSOL*

Assume that our system is presented as a *FSM*, such that:

- N is a size of $\mathcal{A}$, n is a size of $\tilde{\mathcal{A}}$ and l is a size of index structure I.
- The decomposition is given: $\mathcal{E}_\mathcal{I} = \mathcal{E}_\iota = 0$.
- The computation is exponential in the form: $\mathfrak{T} = e^{g(x)}$.

In this case,

$$\mathfrak{T}_{new} = P^p(\mathfrak{T}(n), \mathfrak{T}(l)),$$

where P^p denotes polynomial of degree p, and

$$\mathfrak{T}_{old} = \mathfrak{T}(l \cdot n).$$

The question to answer is: when $f(n \cdot l) > P^p(f(n), f(l))$. According to our assumptions, we obtain that the comparison of the computation times in (1) looks like:

$$e^{g(n \cdot l)} > a_p(e^{p \cdot g(n)} + e^{p \cdot g(l)}).$$

Assume that $n = l$. Then, $g(n^2) > p \cdot g(n) + \ln 2 + \ln(a_p)$. Assume that $g(x) = \ln^2(x)$, then $f(x) = x^{\ln(x)}$. In this case, we obtain that (1) is transformed to:

$$\ln^2(n^2) > p \cdot \ln^2(n) + \ln 2 + \ln(a_p)$$

or

$$\ln^2(n) > \frac{\ln(2 \cdot a_p)}{2 - p}.$$

8.1.2 Complexity Gain for Polynomial Checker

If we a logic, where the computational procedure is polynomial in the sizes of A and in I and each $\mathcal{A}_i$ too, then we do not obtain any time gain.

8.2 Scenario B: Incremental Re-computations

Assume that we change several times (let us denote the number of times by ς) some fixed component $\breve{\mathcal{A}}$ of $\mathcal{A}$. We check each time whether $\mathcal{A} \models \phi$.

8.2.1 Complexity Gain for Polynomial Checker

Let $\Im_{old}$ be time to solve the given problem by the traditionally applied way. It should be clear that $\Im_{old} = \varsigma \cdot \Im(N)$. Let $\Im_{new}$ be time to solve the same problem, when structure $\mathcal{A}$ is viewed as a strongly distributed system. It is easy to see that

$$\Im_{new}(N, n) = \Im(N - n) + \varsigma \cdot \Im(n) + \Im_{F_{\Phi,\phi}} + \varsigma \cdot \Im_S.$$

The question to answer is: Which value of n provides that $\Im_{old} > \Im_{new}$. Assume that $\Im(x) = x^2$, then (1) becomes to be:

$$\varsigma \cdot N^2 > (N - n)^2 + \varsigma \cdot n^2 + \Im_{F_{\Phi,\phi}} + \varsigma \cdot \Im_S,$$

$$N^2 - 2 \cdot n \cdot N + n^2(\varsigma + 1) + \Im_{F_{\Phi,\phi}} + \varsigma \cdot \Im_S - \varsigma \cdot N^2 < 0,$$

$$n_{1,2} = \frac{N \pm \sqrt{N^2 + (\varsigma+1)(N^2(\varsigma-1) - \Im_{F_{\Phi,\phi}} - \varsigma \cdot \Im_S)}}{\varsigma+1}.$$

If $n_1 \leq n \leq n_2$, then $\Im_{old} > \Im_{new}$ and

$$n_2 = \frac{N + \sqrt{\varsigma^2(N^2 - \Im_S) - \varsigma(\Im_S + \Im F_{\Phi,\phi}) - \Im_{F_{\Phi,\phi}}}}{\varsigma+1},$$

$$\lim_{\varsigma \to \infty} n_2 = \sqrt{N^2 - \Im_S}.$$

8.2.2 Complexity Gain for Other Logics

Let $\mathcal{L}$ be any proper sub-logic of *MSOL* stronger that *FOL*. Our theorems do not hold in the following: if we apply it, then $\psi_{\iota,J}$ are not necessary in $\mathcal{L}$.

8.3 Scenario C: Parallel Computations

In (1), the new computation time is calculated as:

$$\mathfrak{T}_{new} = \mathfrak{E}_I + \sum_{\iota \in I} \mathfrak{E}_\iota + \mathfrak{C}_I + \sum_{\iota \in I} \mathfrak{C}_\iota + \mathfrak{T}_{F_{\Phi,\phi}} + \mathfrak{T}_S$$

for the case, when all the computations are done sequentially on a single computational unit. In fact, now even personal computers and smartphones have several cores. In this case, the computation may be done in the following way (we assume that there exist enough computational units for total parallelism):
Extraction Super Step: The extraction of the index structure I from G and each G_ι from G may be done in parallel as well as the building of $F_{\Phi,\phi}$. We denote the extraction time by:

$$\mathfrak{E} = \max\{\mathfrak{E}_I, \max_{\iota \in I}\{\mathfrak{E}_\iota\}, \mathfrak{T}_{F_{\Phi,\phi}}\}.$$

Computational Super Step: The computation of all values of $b_{I,J}$ and $b_{\iota,J}$ may be done in parallel as well. We denote it by

$$\mathfrak{C} = \max\{\mathfrak{C}_I(n_I), \max_{\iota \in I}\{\mathfrak{C}_\iota(n_\iota)\}\}.$$

In fact, at this step, even more parallelism may be reached, if we compute all $b_{\iota,J}$ in parallel.
Final Proceeding: $\mathfrak{T}_S$ still denotes time to search one result of $F_{\Phi,\phi}$.

The new computation time for the case of full parallelism is:

$$\mathfrak{T}^{BSP}_{new} = \mathfrak{E} + \mathfrak{C} + \mathfrak{T}_S.$$

The computation model fails in the general framework of BSP; cf. [114].

8.3.1 Complexity Gain for *MSOL*

The computation is exponential in the form: $\mathfrak{T} = e^{g(x)}$. In this case,

$$\mathfrak{T}_{old} = \mathfrak{T} = f(N) = e^{g(N)} \text{ and } \mathfrak{T}^{BSP}_{new} = \mathfrak{E} + P^p(e^{g(\frac{N}{k})}) + \mathfrak{T}_S,$$

and the question to answer is: When $f(n \cdot k) > P^p(f(n))$. According to our assumptions, we obtain:

$$e^{g(n \cdot k)} > \mathfrak{E} + a_p \cdot e^{p \cdot g(n)} + \mathfrak{T}_S.$$

Assume that $k = n$, which means that there exist enough computation units for full parallelization. In this case, the condition of the effective computation looks like:

$$e^{g(n^2)} > \mathfrak{E} + a_p \cdot e^{p \cdot g(n)} + \mathfrak{T}_S.$$

8.3.2 Complexity Gain for Polynomial Checker

Assume that again $\mathfrak{T}(x) = x^2$, and all G_ι are of the same size $\frac{N}{k}$ then:

$$\mathfrak{T}_{old} = N^2 \text{ and } \mathfrak{T}^{BSP}_{new} = \mathfrak{E} + (\tfrac{N}{k})^2 + \mathfrak{T}_S,$$

$$N^2 > \mathfrak{E} + (\tfrac{N}{k})^2 + \mathfrak{T}_S \; ; N^2 - (\tfrac{N}{k})^2 > \mathfrak{E} + \mathfrak{T}_S.$$

Now, the condition of the effective computation looks like:

$$N^2 \cdot \tfrac{(k^2-1)}{k^2} > \mathfrak{E} + \mathfrak{T}_S.$$

Complexity consideration for other logics $\mathcal{L}$, which are proper sub-logics of *MSOL* stronger than First Order Logic, are similar to the one, given in Sect. 8.2.2.

8.4 Scenario D: Parallel Activities on Distributed Environments

In this section, we consider the underlying structure, the formula and the modularity as well as possible applications of our method for parallel activities on distributed environments.

8.4.1 Complexity Gain

The full computation process is composed now from the following steps (the above $\mathfrak{E} = 0$):

Computational Super Step The computation of all values of $b_{I,J}$ and $b_{\iota,J}$ is done in parallel in the corresponding sites. We again denote by

$$\mathfrak{C} = \max\{\mathfrak{C}_I(n_I), \max_{\iota \in I}\{\mathfrak{C}_\iota(n_\iota)\}\}.$$

Recall that in each site, the $b_{\iota,J}$ still may be computed in parallel if the corresponding computer has several cores.

Communication Super Step The results $b_{I,J}$ and $b_{\iota,J}$ must be sent for the final proceeding. We denote by $\mathfrak{T}_I$ the time to transfer all values of $b_{I,J}$, and by $\mathfrak{T}_\iota$ the time to transfer all values of $b_{\iota,J}$. The communication time now is

$$\mathfrak{T} = \max\{\mathfrak{T}_I, \max_{\iota \in I}\{\mathfrak{T}_\iota)\}\}.$$

Final Proceeding $\mathfrak{T}_S$ still denotes the time to search one result of $F_{\Phi,\phi}$. The new computation time of (1) for the case of the distributed storage and computation is:

$$\mathfrak{T}^{distr}_{new} = \mathfrak{C} + \mathfrak{T} + \mathfrak{T}_S.$$

If the computations and the data transfer in each site may be done in parallel, then we may combine two first super steps in the above model in one step that, in fact, leads to some variation of LogP model, introduced in [20]. We denote

$$\mathfrak{D} = \max\{(\mathfrak{C}_I(n_I) + \mathfrak{T}_I), \max_{\iota \in I}\{(\mathfrak{C}_\iota(n_\iota) + \mathfrak{T}_\iota)\}\}.$$

Now, the corresponding computation time is:

$$\mathfrak{T}^{LogP}_{new} = \mathfrak{D} + \mathfrak{T}_S.$$

Observation 5 ***Communication Load*** *Note that the only values transferred between different computational sites are $b_{I,J}$ and $b_{\iota,J}$, which are binary.*

Confidentiality *All meaningful information is still stored in the corresponding locations in the secure way and t is not transferred.*

For more details we refer to [97].

9 Description of the Method

Our general scenario is as follows: given logic $\mathcal{L}$, structure $\mathcal{A}$ as a composition of structures $\mathcal{A}_i$, $i \in I$, index structure $\mathcal{I}$ and formula ϕ of the logic to be evaluated on $\mathcal{A}$. The question is: *What is the reduction sequence of ϕ if any?* We propose a general approach to try to answer the question and to investigate the computation gain of the incremental evaluations. The general template is defined as follows:

1. **Prove preservation theorems**

   Given logic $\mathcal{L}$.

   (a) **Define disjoint union of $\mathcal{L}$-structures** The logic may be defined for arbitrary structures or rather for a class of structures like graphs, (directed) acyclic graphs, trees, words, (Mazurkiewicz) traces, cf. [26], or (lossy) message

sequence charts, etc. In the general case, we use Definition 4 that provides a logical definition of *disjoint union* of the components: $\mathcal{A} = \dot{\bigsqcup}_{i \in I} \mathcal{A}_i$. Adaptation of Definition 4 to the case of Weighted Monadic Second Order Logic (*WMSOL*), which is introduced over trees, is presented in Definition 10. If logic $\mathcal{L}$ is introduced over another class of structures then Definition 4 must be aligned accordingly.

(b) **Define preservation property** $XX - PP$ **for** $\mathcal{L}$ After we defined the appropriate disjoint union of structures, we define the notion of a *(XX) preservation property (PP)* for logics; see Definitions 6 and 7.

(c) **Prove the preservation property** $XX - PP$ **for** $\mathcal{L}$ Now, we try to prove the corresponding preservation property for $\mathcal{L}$. As a rule, such preservation theorem can be proven by suitable Pebble games, which are generalizations of Ehrenfeucht-Fraïssé games. Our Theorem 4 shows that the preservation theorems hold for lots of extensions of *FOL*. However, theorems like Theorem 4 are not always true. In Proposition 1, we show that our restriction to unary generalized quantifiers (*MTC* and *MLFP*) is necessary. Moreover, Theorem 7 shows for which semi-rings *WMSOL* possess the preservation property.

2. **Define Translation Schemes**
   Given logic $\mathcal{L}$.
   Definition 11 gives the classical syntactically defined translation schemes. Definition 12 is an adaption of Definition 11 to the case of a many-sorted structures. In general, Definition 11 must be adopted in the similar way to the given logic $\mathcal{L}$. Definition 11 gives rise to two induced maps, translations and transductions. Transductions describe the induced transformation of $\mathcal{L}$-structures and the translations describe the induced transformations of $\mathcal{L}$-formulae: see Definitions 15 and 16. Again, the presented adaptation of the definitions to the case of *WMSOL* is a bit tricky. If the $\mathcal{L}$-translation scheme is defined correctly, then the proof of the corresponding variation of Theorem 8 is easily verified.
3. **Strongly Distributed System**
   Given $\mathcal{L}$-structure $\mathcal{A}$.
   If the given $\mathcal{L}$-structure $\mathcal{A}$ is a Φ-strongly distributed composition of its components, then we may apply $\mathcal{L}$-variation of our main Theorem 9 to it. Theorem 9 shows how effectively compute the reduction sequences for different logics, under investigation, for the strongly distributed systems.

Finally, we derive a method for evaluating $\mathcal{L}$-formula ϕ on $\mathcal{A}$, which is a Φ-strongly distributed composition of its components. The method proceeds as follows:

Preprocessing: Given ϕ and Φ, but not $\mathcal{A}$, we construct a sequence of formulas $\psi_{i,j}$ and an evaluation function $F_{\Phi,\phi}$ as in Theorem 9.

Incremental Computation: We compute the local values $b_{i,j}$ for each component of the $\mathcal{A}$.

Solution: Theorem 9 now states that ϕ, expressible in the corresponding logic $\mathcal{L}$, on $\mathcal{A}$ may be effectively computed from $b_{i,j}$, using $F_{\Phi,\phi}$.

The detailed complexity analysis of the method may be found in Sect. 8.

10 Conclusion and Outlook

In this contribution, we introduced the notion of strongly distributed systems and presented a uniform logical approach to incremental automated reasoning on such systems. The approach is based on systematic use of two logical reduction techniques: Feferman-Vaught reductions and the syntactically defined translation schemes.

Our general scenario is as follows: given logic $\mathcal{L}$, structure $\mathcal{A}$ as a composition of structures $\mathcal{A}_i$, $i \in I$, index structure $\mathcal{I}$ and formula ϕ of the logic to be evaluated on $\mathcal{A}$. The question is: What is the reduction sequence for ϕ, if any?

We showed that if we may prove preservation theorems for $\mathcal{L}$ as well as if $\mathcal{A}$ is a strongly distributed composition of its components then the corresponding reduction sequence for $\mathcal{A}$ may be effectively computed. In such a case, we derive a method for evaluating $\mathcal{L}$-formula ϕ on $\mathcal{A}$, which is a Φ-strongly distributed composition of its components. First of all, given ϕ and Φ, but not a $\mathcal{A}$, we construct a sequence of formulas $\psi_{i,j}$ and an evaluation function $F_{\Phi,\phi}$. Next, we compute the local values $b_{i,j}$ for each component of the $\mathcal{A}$. Finally, our main theorem states that ϕ, expressible in the corresponding logic $\mathcal{L}$, on $\mathcal{A}$ may be effectively computed from $b_{i,j}$, using $F_{\Phi,\phi}$.

We showed that the approach works for lots of extensions of *FOL* but not all. The considered extensions of *FOL* are suitable candidates for modeling languages for components and services, used in incremental automated reasoning, data mining, decision making, planning and scheduling.

We plan to apply the proposed methodology to the incremental reasoning, based on the promising variations of *WMSOL* as introduced recently in [61, 63, 85] (see also [62]).

11 Extensions of First Order Logic

In this section, we consider several extensions of First Order Logic (*FOL*). *FOL* is not powerful enough to express many useful properties. This obstacle can be overcome by adding different operators as well as by richer quantification. Second Order Logic (*SOL*) is like First Order Logic (*FOL*) but allows quantification over relations. If the arity of the relation restricted to 1 then we deal with Monadic Second Order Logic (*MSOL*).

In our further considerations, we will need some additional logical tools and notations. For all logics we define:

Definition 18 (*Quantifier Rank of Formulae*)
Quantifier rank of formula φ ($rank(\varphi)$) can be defined as follows:

- for φ without quantifiers $rank(\varphi) = 0$;
- if $\varphi = \neg\varphi_1$ and $rank(\varphi_1) = n_1$, then $rank(\varphi) = n_1$;
- if $\varphi = \varphi_1 \cdot \varphi_2$, where $\cdot \in \{\vee, \wedge, \rightarrow\}$, and $rank(\varphi_1) = n_1$, $rank(\varphi_2) = n_2$, then $rank(\varphi) = max\{n_1, n_2\}$;

– if $\varphi = Q\varphi_1$, where Q is a quantifier, and $rank(\varphi_1) = n_1$, then $rank(\varphi) = n_1 + 1$.

We use the following notation for arbitrary logics $\mathcal{L}$: $\mathcal{A} \equiv^n_{\mathcal{L}} \mathcal{B}$ means that all formulae of logic $\mathcal{L}$ with quantifier rank n have in these structures $\mathcal{A}$ and $\mathcal{B}$ the same truth value.

It is well known that the expressive power of *FOL* is very limited. For example, the transitive closure is not defined in this logic. The source of this defect is the lack of counting or recursion mechanism in this logic. Several attempts to augment the expressive power of *FOL* were done in this direction. For example, Immerman, cf. [51], introduced the *counting quantifier* $\exists ix$, that can be read as: "there are at lest i elements x such that ...". On the other hand, these attempts were inspired by the work by Mostowski, cf. [86], when he introduced the notion of *cardinality quantifiers* (for example: "there are infinitely many elements"), and Tarski, cf. [113], who studied the *infinitary languages*. The next development of this subject was done in the works by Linström, cf. [66, 67], which introduced *generalized quantifiers*. In this contribution, we mostly follow [60]. We use the notation $\mathcal{K}$ (or $\mathcal{Q}$) for an arbitrary class of structures. If τ is a vocabulary, $\mathcal{K}(\tau)$ is the class of structures over τ that are in $\mathcal{K}$.

Definition 19 (*Simple Unary Generalized Quantifier*)
A *simple unary generalized quantifier* is a class $\mathcal{Q}$ of structure over the vocabulary consisting of a unary relation symbol P, such that $\mathcal{Q}$ is closed under isomorphism, i.e. if $\mathcal{U} = \langle \mathcal{U}, \mathcal{P}^{\mathcal{U}} \rangle$ is a structure in $\mathcal{Q}$ and $\mathcal{U}' = \langle \mathcal{U}', \mathcal{P}^{\mathcal{U}'} \rangle$ is a structure that is isomorphic to $\mathcal{U}$, then $\mathcal{U}'$ is also in $\mathcal{Q}$.

The *existential* quantifier is the class of all structures $\mathcal{U} = \langle \mathcal{U}, \mathcal{P}^{\mathcal{U}} \rangle$ with P^U being a non-empty subset of U, while the *universal* quantifier consists of all structures of the form $\mathcal{U} = \langle \mathcal{U}, \mathcal{U} \rangle$. Numerous natural examples of simple unary generalized quantifiers on class of finite structures arise from properties that are not *FOL* definable on finite structures, such as "there is an even number of elements", "there are at least $log(n)$ many elements"... In particular, the quantifier "there is an even number of elements" can be viewed as the class: $\mathcal{Q}_{even} = \{\langle U, P^U \rangle : U$ is a finite set, $P^U \subseteq U$, and $|P^U|$ is even$\}$. We may expend Definition 19 to the *n-ary generalized quantifier*.

Definition 20 (*Lindström Quantifiers*)
Let us $(n_1, n_2, ..., n_\ell)$ be a sequence of positive integers. A *Lindström Quantifier of type* $(n_1, n_2, ..., n_\ell)$ is in a class $\mathcal{Q}$ of structure over the vocabulary consisting of relation symbols $(P_1, P_2, ..., P_\ell)$ such that P_i is n_i-ary for $1 \leq i \leq \ell$ and $\mathcal{Q}$ is closed under isomorphisms.

One of the most known examples of non-simple quantifiers is the *equicardinality or Härtig quantifier I*. This is a Lindström Quantifier of type (1, 1), which comprises all structures $\mathcal{U} = \langle \mathcal{U}, \mathcal{X}, \mathcal{Y} \rangle$ when $|X| = |Y|$. Another example is the *Rescher* quantifier whose mean is *more*.

Another way to extend *FOL* is to allow countable disjunctions and conjunctions:

Definition 21 (*Infinitary Logics*)

- $L_{\omega_1\omega}$ is the logic, which allows countable disjunctions and conjunctions;
- $L^k_{\omega_1\omega}$ is the logic, which allows countable disjunctions and conjunctions, but has only a total of k distinct variables;
- $L^k_{\infty\omega}$, $k \geq 1$ is the logic, which allows infinite disjunctions and conjunctions, but has only a total of k distinct variables;
- $L^\omega_{\infty\omega} = \bigcup L^k_{\infty\omega}$.

We assume that only variables involved are $v_0, \ldots, v_{k-1}$.

Now, we introduce the syntax and the semantics of the logic $L^k_{\infty\omega}$ that contains simple unary generalized quantifiers.

Definition 22 Let $\mathcal{Q} = \{\mathcal{Q}_i : i \in I\}$ be a family of simple unary generalized quantifiers and let k be a positive integer. The infinitary logic $L^k_{\infty\omega}(\mathcal{Q})$ with k variables and the generalized quantifiers $\mathcal{Q}$ has the following syntax (for any vocabulary τ):

- the variables of $L^k_{\infty\omega}(\mathcal{Q})$ are v_1, ..., v_k;
- $L^k_{\infty\omega}(\mathcal{Q})$contains all *FOL* formulae over τ with variables among v_1, ..., v_k;
- if φ is a formulae of $L^k_{\infty\omega}(\mathcal{Q})$, then so is $\neg\varphi$;
- if Ψ is a set of formulae of $L^k_{\infty\omega}(\mathcal{Q})$, then $\bigvee \Psi$ and $\bigwedge \Psi$ are also formulae of $L^k_{\infty\omega}(\mathcal{Q})$;
- if φ is a formulae of $L^k_{\infty\omega}(\mathcal{Q})$, then each of the expressions $\exists v_j\varphi$, $\forall v_j\varphi$, $\mathcal{Q}_i v_j\varphi$ is also a formulae of $L^k_{\infty\omega}(\mathcal{Q})$ for every j such that $1 \leq j \leq k$ and for every $i \in I$.

The semantic of $L^k_{\infty\omega}(\mathcal{Q})$ is defined by induction on the construction of the formulae. So, $\bigvee \Psi$ is interpreted as a disjunction over all formulae in Ψ and $\bigwedge \Psi$ is interpreted as a conjunction. Finally, if $\mathcal{U}$ is the structure having U as its universe and $\varphi(v_j, \bar{y})$ is a formulae of $L^k_{\infty\omega}(\mathcal{Q})$ with free variables among the variables of v_j and the variables in the sequence $\bar{y}$, and $\bar{u}$ is a sequence of elements from the universe of $\mathcal{U}$, then: $U, \bar{u} \models Q_i v_j\varphi(v_j, \bar{y})$ iff the structure $\langle U, \{a : U, a, \bar{u} \models \varphi(v_j, \bar{y})\}\rangle$ is in the quantifier $\mathcal{Q}_i$.

We may also enrich the expressive power of *FOL* by allowing quantification over relation symbols. Second Order Logic (*SOL*) is like *FOL*, but allows also variables and quantification over relation variables of various but fixed arities. Monadic Second Order Logic (*MSOL*) is the sublogic of *SOL* where relation variables are restricted to be unary. The meaning function of formulae is explained for arbitrary τ–structures, where τ is the vocabulary, i.e. a finite set of relation and constant symbols. Fixed Point Logic (*LFP*) can be viewed as a fragment of *SOL*, where the second order variables only occur positively and in the fixed point construction. Similarly *MLFP* corresponds to the case where the arity of the relation variables is restricted to 1. The semantics of the fixed point is given by the least fixed point, which does always exist because of the positivity assumption on the set variable. The Logic *LFP* is defined similarly with operators *k-LFP* for every $k \in N$ which bind $2k$ variables. On ordered structures *LFP* expresses exactly the polynomially recognizable classes of finite structures. Without order, every formula in *LFP* has a polynomial model

checker. For transition systems, *MLFP* corresponds exactly to μ-calculus, cf. [2, 36, 115].

The logic *MTC* (Monadic Transitive Closure) is defined inductively, like *FOL*. For a thorough discussion of it, cf. [52]. Atomic formulae are as usual. The inductive clauses include closure under the boolean operations, existential and universal quantification and one more clause:

Syntax If $\phi(x, y, \bar{u})$ is a *MTC*–formula with x, y and $\bar{u} = u_1, \ldots, u_n$ being its free variables, s, t are terms, then $MTCx, y, s, t\phi(x, y, \bar{u})$ is a *MTC*–formula with x, y bound and $\bar{u}$ free.

Semantics The formula $MTCx, y, s, t\phi(x, y, \bar{u})$ holds in a structure $\mathcal{U}$ under an assignment of variables z if $s_z, t_z \in TrCl(\phi^{\mathcal{U}})$.

The logic *TC* is defined similarly with operators k-*TC* for every $k \in N$ which bind $2k$ variables. For more detailed exposition, we refer to [1, 10, 33, 44, 45].

In [44], E. Grädel introduced a generalization of Ehrenfeucht–Fraïssé Games for *TC*. As we need this game in the proof of Theorem 4, we give Grädel's definition and Theorem 11 concerning Pebble Games for *TC* verbatim as in [44].

Definition 23 (*Ehrenfeucht–Fraïssé Games for TC*)
Suppose we have two structures $\mathcal{U}$ and $\mathcal{V}$ of the same vocabulary σ. Let $c_1, \ldots, c_s$ and $d_1, \ldots, d_s$ be the interpretation of the constants of σ in $\mathcal{U}$ and $\mathcal{V}$, respectively. The k–pebble game on the pair $(\mathcal{U}, \mathcal{V})$ is played by Players I and II as follows: There are k pairs $(u_1, v_1), \ldots, (u_k, v_k)$ of pebbles. Each round of the game consists of either an $\exists$–move, $\forall$–move or *TC*–move:

$\exists$*–move*: *Player I places a yet unused pebbles u_i on an element of $\mathcal{U}$. Player II answers by putting the corresponding pebble v_i on $\mathcal{V}$.*

$\forall$*–move*: *Similarly but with "reversed board": Player I places v_i on $\mathcal{V}$. Player II responds with u_i on $\mathcal{U}$.*

TC–move: *Suppose that r pairs of pebbles are already on the board. For some $l \leq (k - r)/2$, Player I selects a sequence $\bar{x}_0, \ldots, \bar{x}_m$ of l–tuples in $\mathcal{U}$ such that $\bar{x}_0$ and $\bar{x}_m$ consist only of sets of constants and already pebbled elements. Player II indicates a similar sequences (not necessary of the same length) of l–tuples $\bar{y}_0, \ldots, \bar{y}_n$ in $\mathcal{V}$ where $\bar{y}_0 = f(\bar{x}_0), \ldots, \bar{y}_n = f(\bar{x}_m)$.*
Player I then selects some $i \leq n$ and places $2l$ (yet unused) pebbles on $\bar{y}_i$ and $\bar{y}_{i+1}$. Player II selects some a $j \leq m$ and places the corresponding pebbles on $\bar{x}_j$ and $\bar{x}_{j+1}$.

$\neg$*TC–move*: *is like TC–move, but with structures $\mathcal{A}$ and $\mathcal{V}$ interchanged.*

When all pebbles are placed, Player I wins if the pebbles determine a local isomorphism from $\mathcal{U}$ to $\mathcal{V}$. More precisely: Let $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ be the elements carrying the pebbles $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$. If the mapping f with $f(a_i) = b_i$ for $i = 1, \ldots, k$ and $f(c_i) = d_i$ for $i = 1, \ldots, s$ is an isomorphism between the substructures of $\mathcal{A}$ and $\mathcal{V}$ that are generated by the pebbles elements and the constants, then Player II wins; otherwise Player I wins.

Theorem 11 (Grädel [44])
For all structures $\mathcal{U}$ and $\mathcal{V}$ and all $k \in N$, the following are equivalent:

1. *Player II has a winning strategy for the TC–game with k pebbles on $(\mathcal{U}, \mathcal{V})$.*
2. $\mathcal{U} \equiv^{k}_{TC} \mathcal{V}$.

11.1 Complexity of Computation for Extensions of First Order Logic

Computation for *MSOL* sits fully in the polynomial hierarchy. For the complexity of *FOL* and *MSOL*, see [41].

More precisely, the complexity of computation (in the size of the structure) of Second Order Logic expressible properties can be described as follows. The class *NP* of non-deterministic polynomial-time problems is the set of properties, which are expressible by Existential Second Order Logic on finite structures, cf. [37]. Computation for *SOL* definable properties is in the polynomial hierarchy, cf. [34]. Moreover, for every level of the polynomial hierarchy there is a problem, expressible in *SOL*, that belongs to this class. The same fact hold for *MSOL*, too, as observed in [69].

Computation for properties, definable in Fixed Point Logic, is polynomial, cf. [115]. *CTL** is a superset of Computational Tree Logic and Linear Temporal Logic. All the problems, which are expressible by *CTL**, can be computed in polynomial time, cf. [36]. The relation between *FOL* with generalized quantifiers and computations with oracles is investigated in [70]. Most properties, which appear in real-life applications, are stronger than *FOL* but weaker than *MSOL*, and their computational complexity is polynomial.

12 Weighted Monadic Second Order Logic

In this section, we follow [29] almost verbatim. Let $\mathbf{N} = \{1, 2, \dots\}$ be the set of natural numbers and let $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$. A ranked alphabet is a pair $(\Sigma, \mathbf{rk}_\Sigma)$ consisting of a finite alphabet Σ and a mapping $\mathbf{rk}_\Sigma : \Sigma \to \mathbf{N}_0$, which assigns to each symbol of Σ its rank. By $\Sigma^{(m)}$ we denote the set of all symbols with rank $m \in \mathbf{N}_0$ and $a^{(m)}$ denotes that $a \in \Sigma^{(m)}$. Let $\max_\Sigma = \max\{\mathbf{rk}_\Sigma(a) | a \in \Sigma\}$, the maximal rank of Σ. Let $\mathbf{N}^*$ be the set of all finite words over $\mathbf{N}$. A *tree domain* $\mathcal{B}$ is a finite, non-empty subset of $\mathbf{N}^*$ such that for all $u \in \mathbf{N}^*$ and $i \in \mathbf{N}$, $u.i$ is the prefix of u of length i. Moreover $u.i \in \mathcal{B}$ implies $u.1, \dots, u.(i-1) \in \mathcal{B}$, $u.1, \dots, u.(i-1)$ is called *immediate prefix* of $u.i$; the immediate prefix of $u.1$ is the empty set ϵ. Note that the tree domain of $\mathcal{B}$ is prefix-closed. A *tree over a set L (of labels)* is a mapping $t : \mathcal{B} \to L$, such that $dom(t) = \mathcal{B}$ is a tree domain, $im(t)$ is the image of t. The elements of $dom(t)$ are called *positions* of t and $t(u)$ is called *label* of t at $u \in dom(t)$. The set of all trees over L is denoted by T_L.

Definition 24 (*Tree Valuation Monoid*)
A tv-monoid is a quadruple $\mathcal{D} = (D, +, Val, \mathbf{0})$ such that $(D, +, \mathbf{0})$ is a commutative monoid and $Val : T_D \to D$ is a function with $Val(d) = d$ for every tree $d \in T_D$ and $Val(t) = \mathbf{0}$, whenever $\mathbf{0} \in im(t)$ for $t \in T_D$.

Val is called a *(tree) valuation function*.

Definition 25 (*Product Tree Valuation Monoid*)
A ptv-monoid $\mathcal{D} = (D, +, Val, \diamond, \mathbf{0}, \mathbf{1})$ consists of a tree valuation monoid, a constant $\mathbf{1} \in D$ with $Val(t) = \mathbf{1}$, whenever $im(t) = \{\mathbf{1}\}$ for $t \in T_D$, and an operation $\diamond : D^2 \to D$ with $\mathbf{0} \diamond d = d \diamond \mathbf{0} = \mathbf{0}$ and $\mathbf{1} \diamond d = d \diamond \mathbf{1} = \mathbf{1}$.

Note that the operation $\diamond$, in general, has to be neither commutative nor associative.

12.1 Definition of WMSOL

Given a ptv-monoid D, the syntax of $WMSOL$ over D is defined by the following way:
Boolean formulae:

- $label_a(x)$ and $edge_i(x, y)$ for $a \in \Sigma$ and $1 \leq i \leq \max_\Sigma$;
- $x \in X, \neg\beta_1, \beta_1 \wedge \beta_2, \forall x \beta_1, \forall X \beta_1$ for first order variable x and second order variable X.

Weighted formulae:

- d for $d \in D$;
- β for boolean formula β;
- $\phi_1 \vee \phi_2$, $\phi_1 \wedge \phi_2$, $\exists x \phi_1$, $\forall x \phi_1$, $\exists X \phi_1$, $\forall X \phi_1$.

The set $free(\phi)$ of free variables occurring in ϕ is defined as usual. Semantics of $WMSOL$ valuates trees by elements of D. There is no change in semantics of boolean formulae. $\mathbf{0}$ defines the semantics of the truth value "false". $\mathbf{1}$ defines the semantics of the truth value "true". The monoid operation "+" is used to define semantics of disjunction and existential quantifier. The monoid Val function is used to define the semantics of the first order universal quantification. If, for example, we use the max-plus-semiring the semantical interpretation of $\forall x \phi$ is the sum of all weights (rewards or time) defined by ϕ for all different positions x. More precisely, for a $(\mathcal{V}, t)$-assignment that maps $\tilde{\sigma} : \mathcal{V} \to dom(t) \cup PS(dom(t))$, with $\tilde{\sigma}(x) \in dom(t)$ and $\tilde{\sigma}(X) \subseteq dom(t)$, and $s \in T_{\Sigma_\mathcal{V}}$. The formal definition of the semantics of $WMSOL$ see in [29].

12.2 Expressive Power of Weighted Monadic Second Order Logic

WMSOL and its fragments have a considerable expressive power. In [81], the coincidence of recognizable trace series with those, which are definable by restricted formulae from a weighted logics over traces, was proved. In [40], a notion of a *WMSOL* logics over pictures was introduced, weighted 2-dimensional on-line tessellation automata (*W2OTA*) was defined and it was proved that for commutative semirings, the class of picture series defined by sentences of the weighted logics coincides with the family of picture series that are computable by *W2OTA*. In [77], quantitative models for texts were investigated, an algebraic notion of recognizability was defined and it was shown that recognizable text series coincide with text series definable in weighted logics. Nested words are a model for recursive programs. In [78], quantitative extensions of nested word series were considered and it was shown that regular nested word series coincide with series definable in weighted logics. Moreover, lots of optimization problems and counting problems are expressible in *WMSOL*, cf. [9, 30–32, 76]. The logic may be used in order to describe data mining problems, decision making, planning and scheduling.

13 Proof of Theorem 4 for MTC^m

In this section, we give our original proof of the following part of Theorem 4:

Theorem 12 *Let $\mathcal{I}$ be an index structure. Then $DJ - PP(MTC^m, MTC^m)$ and $FShu - PP(MTC^m, MTC^m)$ hold.*

Proof We use pebble game for *MTC* as introduced in [44].
∃–move: If Player I puts pebble u on some element a of structure $\mathcal{A}_i$ for some $i \in I$, Player II now places her pebble v on b of structure $\mathcal{B}_i$ using the winning strategy of the components.
∀–move is similar.
MTC–**move**: If Player I selects a sequence $x_0,\ldots,x_m$ in $\mathcal{A}$, we divide the sequence into segments $x_{m_{0,0}},\ldots,x_{m_{0,1}},x_{m_{1,0}},\ldots,x_{m_{1,1}},\ldots,x_{m_{p,0}},\ldots,x_{m_{p,1}}$ with $m_{0,0} = 0, m_{p,1} = m$ and such that each subsequence $x_{m_{q,0}},\ldots,x_{m_{q,1}}$ lies in the same component $\mathcal{A}_{i_q}$.
Player II now constructs her sequence $y_0,\ldots,y_n$ in $\mathcal{B}$ segment-wise as follows: She uses two auxiliary pebbles U_0, U_1 and V_0, V_1 on each structure. For the segment $x_{m_{q,0}},\ldots,x_{m_{q,1}}$ she puts U_0 on $x_{m_{q,0}}$ and U_1 on $x_{m_{q,1}}$. Using the winning strategy on components i_q she places the pebbles V_0 and V_1 on elements $y_{m_{q,0}}$ and $y_{m_{q,1}}$ and chooses the intermediate elements according to the winning strategy of the *MTC*–move. The auxiliary pebbles are reused after every segment.
If Player I now pebbles two neighboring elements in the sequence $y_0,\ldots,y_n$ in $\mathcal{B}$, two cases can occur:

1. If both pebbled elements are in the same component, Player II just follows her winning strategy on the corresponding component on $\mathcal{A}$.
2. If the two pebbled elements are in different components, then she plays accordingly.

It is now easy to verify that this is indeed a winning strategy. The auxiliary pebbles are only used temporarily to mark the beginning and the end of the segments.

Note that the index set is part of the structures $\mathcal{A}$ and $\mathcal{B}$, and is treated itself like a component. However, in this component Player II copies faithfully the moves of player I.

References

1. Abiteboul, S., Hull, R., Vianu, V.: Foundations of Databases. Addison-Wesley (1995)
2. Arnold, A., Niwiński, D.: Fixed point characterization of weak monadic logic definable sets of trees. In: A. M.Nivat (ed.) Tree Automata and Languages, pp. 159–188. Elsevier Science Publishers B.V. (1992)
3. Ashenhurst, R.: The decomposition of switching functions, vol. 29, pp. 74–116. Annals Computation Laboratory, Harvard University (1959)
4. Bagirov, A., Ordin, B., Ozturk, G., Xavier, A.: An incremental clustering algorithm based on hyperbolic smoothing. Comp. Opt. Appl. **61**(1), 219–241 (2015). http://dx.doi.org/10.1007/s10589-014-9711-7
5. Belardinelli, F., Lomuscio, A.: Quantified epistemic logics for reasoning about knowledge in multi-agent systems. Artif. Intell. **173**(9), 982–1013 (2009)
6. Benedikt, M., Koch, C.: From XQuery to relational logics. ACM Trans. Database Syst. **34**(4), 25:1–25:48 (2009)
7. Boerkoel Jr. J.C., Durfee, E.: Distributed reasoning for multiagent simple temporal problems. J. Artif. Intell. Res. **47**, 95–156 (2013)
8. Boerkoel, J., Planken, L., Wilcox, R., Shah, J.: Distributed algorithms for incrementally maintaining multiagent simple temporal networks. In: Proceedings of the 23rd International Conference on Automated Planning and Scheduling(ICAPS-13), pp. 11–19. AAAI Press (2013). http://www.st.ewi.tudelft.nl/~planken/papers/icaps13.pdf
9. Bollig, B., Gastin, P.: Weighted versus probabilistic logics. In: Diekert, V., Nowotka, D. (eds.) DLT 2009, LNCS R5583, pp. 18–38. Springer, Berlin Heidelberg (2009)
10. Bosse, U.: An Ehrenfeucht–Fraïssé game for fixed point logic and stratified fixed point logic. In: CSL'92. Lecture Notes in Computer Science, vol. 702, pp. 100–114. Springer (1993)
11. Bosse, U.: Ehrenfeucht–Fraïssé games for fixed point logic. Ph.D. thesis, Department of Mathematics, University of Freiburg, Germany (1995)
12. Burgess, J.: Basic Tense Logic, vol. 2, chap. 2, pp. 89–133. D. Reidel Publishing Company (1984)
13. Cao, T., Creasy, P.: Fuzzy types: a framework for handling uncertainty about types of objects. Int. J. Approx. Reason **25**(3), 217–253 (2000)
14. Casanova, G., Pralet, C., Lesire, C.: Managing dynamic multi-agent simple temporal network. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '15, pp. 1171–1179. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2015)
15. Chang, C., Keisler, H.: Model Theory, 3rd edn. Studies in Logic, vol. 73. North–Holland (1990)

16. Courcelle, B., Makowsky, J., Rotics, U.: Linear time solvable optimization problems on graph of bounded clique width, extended abstract. In: J. Hromkovic, O. Sykora (eds.) Graph Theoretic Concepts in Computer Science. Lecture Notes in Computer Science, vol. 1517, pp. 1–16. Springer (1998)
17. Courcelle, B.: The monadic second-order logic of graphs ix: Machines and their behaviours. Theor. Comput. Sci. **151**(1), 125–162 (1995). (Selected Papers of the Workshop on Topology and Completion in Semantics)
18. Courcelle, B.: The monadic second-order logic of graphs VIII: orientations. Ann. Pure Appl. Logic **72**, 103–143 (1995)
19. Courcelle, B., Walukiewicz, I.: Monadic second-order logic, graphs and unfoldings of transition systems. Ann. Pure Appl. Logic **92**, 35–62 (1995)
20. Culler, D., Karp, R., Patterson, D., Sahay, A., Schauser, K., Santos, E., Subramonian, R., von Eicken, T.: LogP: towards a realistic model of parallel computation. In: PPOPP'93 Proceedings of the fourth ACM SIGPLAN Symposium on Principles and practice of parallel programming, vol. 28(7), pp. 1–12 (1993)
21. Curtis, H.: A new approach to the design of switching circuits. Van Nostrand (1962)
22. Cvrček, D.: Authorization model for strongly distributed information systems. Ph.D. thesis, Faculty of Electrical Engineering and Computer Science, Brno University of Technology, Czech Republic (2000)
23. Cyriac, A., Gastin, P.: Reasoning about distributed systems: WYSIWYG (invited talk). In: 34th International Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2014, December 15–17, 2014, New Delhi, India, pp. 11–30 (2014). http://dx.doi.org/10.4230/LIPIcs.FSTTCS.2014.11
24. Dawar, A., Hellat, L.: The expressive power of finitely many genaralized quantifiers: Technical Report CSR 24–93. University of Wales, University College of Swansea, U.K, Computer Science Department (1993)
25. Dibangoye, J., Mouaddib, A.I., Chai-draa, B.: Point-based incremental pruning heuristic for solving finite-horizon DEC-POMDPs. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, AAMAS'09, vol. 1, pp. 569–576. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2009). http://dl.acm.org/citation.cfm?id=1558013.1558092
26. Diekert, V., Rozenberg, G.: The Book of Traces. World Scientific (1995). https://books.google.ca/books?id=vNFLOE2pjuAC
27. D'Inverno, M., Luck, M., Georgeff, M., Kinny, D., Wooldridge, M.: The dMARS architecture: a specification of the distributed multi-agent reasoning system. Auton. Agents Multi-Agent Syst **9**(1–2), 5–53 (2004)
28. Doherty, P., Driankov, D.: Nonmonotonicity, fuzziness, and multi-values. In: Lowen, R., Roubens, M. (eds.) Fuzzy Logic, pp. 3–15. Kluwer Academic Publishers, Dordrecht (1993)
29. Droste, M., Götze, D., Märcker, S., Meinecke, I.: Weighted tree automata over valuation monoids and their characterization by weighted logics. In: Kuich, W., Rahonis, G. (eds.) Bozapalidis Festschrift. LNCS 7020, pp. 30–55. Springer, Berlin Heidelberg (2011)
30. Droste, M., Meinecke, I.: Describing average-and longtime-behavior by weighted MSO logics. In: P. Hliněný, A. Kučera (eds.) MFCS 2010. LNCS 6281, pp. 537–548. Springer, Berlin Heidelberg (2010)
31. Droste, M., Gastin, P.: Weighted automata and weighted logics. Theor. Comput. Sci. **380**, 69–86 (2007)
32. Droste, M., Vogler, H.: Weighted logics for unranked tree automata. Theory Comput. Syst. **48**(1), 23–47 (2009)
33. Ebbinghaus, H., Flum, J., Thomas, W.: Mathematical Logic, 2nd edn. Undergraduate Texts in Mathematics. Springer (1994)
34. Ebbinghaus, H., Flum, J.: Finite Model Theory. Perspectives in Mathematical Logic. Springer (1995)
35. Ehrenfeucht, A.: An application of games to the completeness problem for formalized theories. Fundamenta Mathematicae **49**, 129–141 (1961)

36. Emerson, E.: Temporal and modal logic. In: J. van Leeuwen (ed.) Handbook of Theoretical Computer Science, vol. 2, chap. 16, pp. 995–1072. Elsevier Science Publishers (1990)
37. Fagin, R.: Generalized first-order spectra and polynomial time recognizable sets. In: R. Karp (ed.) Complexity of Computation. American Mathematical Society Proceeding, vol. 7, pp. 27–41. Society for Industrial and Applied Mathematics (1974)
38. Feferman, S., Vaught, R.: The first order properties of products of algebraic systems. Fundamenta Mathematicae **47**, 57–103 (1959)
39. Ferber, J.: Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence, 1st edn. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1999)
40. Fichtner, I.: Weighted picture automata and weighted logics. In: STACS 2006, pp. 313–324. Springer (2006)
41. Frick, M., Grohe, M.: The complexity of first-order and monadic second-order logic revisited. Ann. Pure Appl. Logic **130**(1), 3–31 (2004)
42. Gandhi, A., Khoussainov, B., Liu, J.: Finite automata over structures. In: M. Agrawal, A. Li, S. Cooper (eds.) Theory and Applications of Models of Computation: 9th Annual Conference, TAMC 2012, 2012. Proceedings, pp. 373–384. Springer, Heidelberg (2012)
43. Grädel, E., Siebertz, S.: Dynamic definability. In: 15th International Conference on Database Theory, ICDT'12, pp. 236–248. Berlin, Germany, March 26–29 (2012). https://doi.org/10.1145/2274576.2274601
44. Grädel, E.: On transitive closure logic. In: E. Börger, G. Jäger, H.K. Büning, M. Richter (eds.) Computer Science Logic. Lecture Notes in Computer Science, vol. 626, pp. 149–163. Springer (1992)
45. Grohe, M.: The structure of fixed point logics. Ph.D. thesis, Department of Mathematics, University of Freiburg, Germany (1994)
46. Gurevich, Y.: Modest theory of short chains. I. J. Symb. Logic **44**, 481–490 (1979)
47. Harel, D.: Dynamic Logic, vol. 2, chap. 10, pp. 497–604. Springer Netherlands, Dordrecht (1984)
48. Hasemann, J.M.: Planning, behaviours, decomposition, and monitoring using graph grammars and fuzzy logic. In: Proceedings of the Second International Conference on Artificial Intelligence Planning Systems, University of Chicago, Chicago, Illinois, USA, June 13–15, 1994, pp. 275–280 (1994)
49. Hilbert, D., Bernays, P.: Grundlagen der Mathematik, I. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungn, vol. 40, 2nd edn. Springer, Heidelberg (1970)
50. Immerman, N.: Descriptive complexity. Graduate texts in computer science. Springer (1999). https://doi.org/10.1007/978-1-4612-0539-5
51. Immerman, N.: Relational queries computable in polynomial time. In: STOC'82, pp. 147–152. ACM (1982)
52. Immerman, N.: Languages that capture complexity classes. SIAM J. Comput. **16**(4), 760–778 (1987)
53. Immerman, N.: Expressibility and parallel complexity. SIAM J. Comput. **18**, 625–638 (1989)
54. Jou, J.Y., Liu, C.N.: An efficient functional coverage test for HDL descriptions at RTL. In: International Conference Computer Design, pp. 325–327 (1999)
55. Kandel, A., Davis, H.: The First Fuzzy Decade: (bibliography on Fuzzy Sets and Their Applications). Computer Science Report. New Mexico Institute of Mining and Technology (1976)
56. Kandel, A.: On the decomposition of fuzzy functions. IEEE Trans. Comput. **25**(11), 1124–1130 (1976)
57. Keren, D., Sagy, G., Abboud, A., Ben-David, D., Schuster, A., Sharfman, I., Deligiannakis, A.: Geometric monitoring of heterogeneous streams. IEEE Trans. Knowl. Data Eng. **26**(8), 1890–1903 (2014)
58. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall Inc, Upper Saddle River, NJ, USA (1995)
59. Koenig, S., Likhachev, M.: D*Lite. In: Eighteenth National Conference on Artificial Intelligence, pp. 476–483. American Association for Artificial Intelligence, Menlo Park, CA, USA (2002). http://dl.acm.org/citation.cfm?id=777092.777167

60. Kolaitis, P.G., Väänänen, J.A.: Generalized quantifiers and pebble games on finite structures. In: LiCS'92, pp. 348–359. IEEE (1992)
61. Kreutzer, S., Riveros, C.: Quantitative monadic second-order logic. In: 28th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2013, New Orleans, LA, USA, June 25–28, 2013, pp. 113–122 (2013)
62. Labai, N., Makowsky, J.: Logics of finite Hankel rank. In: Fields of Logic and Computation II-Essays Dedicated to Yuri Gurevich on the Occasion of his 75th Birthday, pp. 237–252 (2015)
63. Labai, N., Makowsky, J.: Weighted automata and monadic second order logic. In: Proceedings Fourth International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2013, Borca di Cadore, Dolomites, Italy, 29–31th August 2013, pp. 122–135 (2013)
64. Lauer, M., Riedmiller, M.: An algorithm for distributed reinforcement learning in cooperative multi-agent systems. In: Proceedings of the Seventeenth International Conference on Machine Learning, pp. 535–542. Morgan Kaufmann (2000)
65. Lichtenstein, O., Pnueli, A., Zuck, L.: Logics of program. In: Lecture Notes in Computer Science, vol. 193, pp. 196–218. Springer (1985)
66. Lindström, P.: First order predicate logic with generalized quantifiers. Theoria **32**, 186–195 (1966)
67. Lindström, P.: On extensions of elementary logic. Theoria **35**, 1–11 (1969)
68. Łukasiewicz, J.: O logice trójwartościowej. Ruch Filozoficzny **5**, 170–171 (1920)
69. Makowsky, J., Pnueli, Y.: Arity vs. alternation in second order definability. In: LFCS'94. Lecture Notes in Computer Science, vol. 813, pp. 240–252. Springer (1994)
70. Makowsky, J., Pnueli, Y.: Oracles and quantifiers. In: Computer Science Logic, 7th Workshop, CSL'93, Swansea, United Kingdom, September 13–17, 1993, Selected Papers, pp. 189–222 (1993). https://doi.org/10.1007/BFb0049333
71. Makowsky, J., Ravve, E.: BCNF via attribute splitting. In: A. Düsterhöft, M. Klettke, K.D. Schewe (eds.) Conceptual Modelling and Its Theoretical Foundations-Essays Dedicated to Bernhard Thalheim on the Occasion of his 60th Birthday. Lecture Notes in Computer Science, vol. 7260, pp. 73–84. Springer (2012). https://doi.org/10.1007/978-3-642-28279-9
72. Makowsky, J., Ravve, E.: Incremental model checking for decomposable structures. In: Mathematical Foundations of Computer Science (MFCS'95). Lecture Notes in Computer Science, vol. 969, pp. 540–551. Springer (1995)
73. Makowsky, J.: Compactness, embeddings and definability. In: J. Barwise, S. Feferman (eds.) Model-Theoretic Logics, Perspectives in Mathematical Logic, chap. 18, pp. 645–716. Springer (1985)
74. Makowsky, J.: Some observations on uniform reduction for properties invariant on the range of definable relations. Fundamenta Mathematicae **99**, 199–203 (1978)
75. Makowsky, J.: Algorithmic uses of the Feferman-Vaught theorem. Ann. Pure Appl. Logic **126**(1–3), 159–213 (2004)
76. Mandrali, E., Rahonis, G.: Recognizable tree series with discounting. Acta Cybern. **19**(2), 411–439 (2009)
77. Mathissen, C.: Definable transductions and weighted logics for texts. In: Proceedings of the 11th International Conference on Developments in Language Theory, DLT'07, pp. 324–336. Springer, Berlin, Heidelberg (2007)
78. Mathissen, C.: Weighted logics for nested words and algebraic formal power series. In: Proceedings of the 35th International Colloquium on Automata, Languages and Programming, Part II, ICALP'08, pp. 221–232. Springer, Berlin, Heidelberg (2008)
79. Meer, K., Naif, A.: Generalized finite automata over real and complex numbers. In: T. Gopal, M. Agrawal, A. Li, S. Cooper (eds.) Theory and Applications of Models of Computation: 11th Annual Conference, TAMC 2014, Chennai, India, April 11–13, 2014. Proceedings, pp. 168–187. Springer International Publishing, Cham (2014)
80. Meer, K., Naif, A.: Periodic generalized automata over the reals. In: Language and Automata Theory and Applications-10th International Conference, LATA 2016, Prague, Czech Republic, March 14–18, 2016, Proceedings, pp. 168–180 (2016)

81. Meinecke, I.: Weighted logics for traces. In: Proceedings of the First International Computer Science Conference on Theory and Applications, CSR'06, pp. 235–246. Springer, Berlin, Heidelberg (2006)
82. Mogali, J., Smith, S., Rubinstein, Z.B.: Distributed decoupling of multiagent simple temporal problems. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16, pp. 408–415. AAAI Press (2016). http://dl.acm.org/citation.cfm?id=3060621.3060679
83. Mohammadian, M.: Supervised learning of fuzzy logic systems. In: Encyclopedia of Artificial Intelligence, pp. 1510–1517 (2009)
84. Monk, J.: Mathematical Logic. Graduate Texts in Mathematics. Springer (1976)
85. Monmege, B.: Spécification et vérification de propriétés quantitatives: Expressions, logiques et automates. Ph.D. thesis, Laboratoire Spécification et Vérification, École Normale Supérieure de Cachan, Cedex, France (2013)
86. Mostowsky, A.: On a generalization of quantifiers. Fundamenta Mathematicae **44**, 12–36 (1957)
87. Nola, A., Sanchez, E., Pedrycz, W., Sessa, S.: Fuzzy Relation Equations and Their Applications to Knowledge Engineering. Kluwer Academic Publishers, Norwell, MA, USA (1989)
88. Oral, T., Polat, F.: A multi-objective incremental path planning algorithm for mobile agents. In: Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, WI-IAT'12, vol. 02, pp. 401–408. IEEE Computer Society, Washington, DC, USA (2012). http://dx.doi.org/10.1109/WI-IAT.2012.143
89. Plastria, F., Bruyne, S.D., Carrizosa, E.: Dimensionality reduction for classification. In: Advanced Data Mining and Applications, 4th International Conference, ADMA 2008, Chengdu, China, October 8–10, 2008. Proceedings, pp. 411–418 (2008)
90. Post, E.: Introduction to a General Theory of Elementary Propositions. Columbia University (1920)
91. Rabin, M.: A simple method for undecidability proofs and some applications. In: Y.B. Hillel (ed.) Logic, Methodology and Philosophy of Science II, Studies in Logic, pp. 58–68. North Holland (1965)
92. Rabin, M.: Decidability of second order theories and automata on infinite trees. Trans. Am. Math. Soc. **141**, 1–35 (1969)
93. Ravve, E.V., Volkovich, Z., Weber, G.W.: A uniform approach to incremental automated reasoning on strongly distributed structures. In: G. Gottlob, G. Sutcliffe, A. Voronkov (eds.) GCAI 2015. Global Conference on Artificial Intelligence. EasyChair Proceedings in Computing, vol. 36, pp. 229–251. EasyChair (2015)
94. Ravve, E., Volkovich, Z., Weber, G.W.: Effective optimization with weighted automata on decomposable trees. Optim. J. **63**, 109–127 (2014). (Special Issue on Recent Advances in Continuous Optimization on the Occasion of the 25th European Conference on Operational Research (EURO XXV 2012))
95. Ravve, E., Volkovich, Z., Weber, G.W.: Reasoning on strongly distributed multi-agent systems. In: Proceedings of the 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, pp. 251–256 (2015)
96. Ravve, E., Volkovich, Z.: A systematic approach to computations on decomposable graphs. In: 15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2013, Timisoara, Romania, September 23–26, 2013, pp. 398–405 (2013)
97. Ravve, E., Volkovich, Z.: Four scenarios of effective computations on sum-like graphs. In: Proceedings of the The 9th International Multi-Conference on Computing in the Global Informationin Technology, pp. 1–8 (2014)
98. Ravve, E., Volkovich, Z.: Incremental reasoning on fuzzy strongly distributed systems (2016). (To appear in Proceedings of The Eleventh International Multi-Conference on Computing in the Global Information Technology)
99. Ravve, E.: Maintenance of queries under database changes: a unified logic based approach. In: Foundations of Information and Knowledge Systems-9th International Symposium, FoIKS 2016, Linz, Austria, March 7–11, 2016. Proceedings, pp. 191–208 (2016)

100. Ravve, E.: Model checking for various notions of products. Master's thesis, Thesis, Department of Computer Science, Technion–Israel Institute of Technology (1995)
101. Ravve, E.: Incremental computations over strongly distributed databases. Concurr Comput Pract. Exp. **28**(11), 3061–3076 (2016)
102. Ren, W., Cao, Y.: Distributed Coordination of Multi-agent Networks: Emergent Problems, Models, and Issues. Springer Publishing Company, Incorporated (2013)
103. Rossi, F., van Beek, P., Walsh, T. (eds.): Handbook of Constraint Programming. Elsevier Science Inc., New York, NY, USA (2006)
104. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 2 edn. Pearson Education (2003)
105. Ryzko, D., Rybinski, H.: Distributed default logic for multi-agent system. In: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Hong Kong, China, 18–22 December 2006, pp. 204–210 (2006)
106. Salamon, T.: Design of Agent-Based Models. Eva & Tomas Bruckner Publishing, Czech Republic (2011)
107. Sanchez, E.: Resolution of composite fuzzy relation equations. Inf. Control **30**(1), 38–48 (1976)
108. Savku, E., Azevedo, N., Weber, G.: Optimal control of stochastic hybrid models in the framework of regime switches. In: Pinto, A.A., Zilberman, D. (eds.) Modeling, Dynamics, Optimization and Bioeconomics II, pp. 371–387. Springer International Publishing, Cham (2017)
109. Schubert, E., Weiler, M., Kriegel, H.P.: Signitrend: scalable detection of emerging topics in textual streams by hashed significance thresholds. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'14, pp. 871–880. ACM, New York, NY, USA (2014)
110. Seibel, A., Schlattmann, J.: A generalized α-level decomposition concept for numerical fuzzy calculus. In: Proceedings of the 16th World Congress of the International Fuzzy Systems Association (IFSA), pp. 66–69 (2015)
111. Shamma, J.: Cooperative Control of Distributed Multi-Agent Systems. Wiley-Interscience, New York, NY, USA (2008)
112. Spaan, M.T.J., Oliehoek, F.A.: Tree-based solution methods for multiagent POMDPs with delayed communication. In: Proceedings of 24th Benelux Conference on Artificial Intelligence, pp. 319–320 (2012). (Extended abstract)
113. Tarski, A.: A model-theoretical result concerning infinitary logics. Not. Am. Math. Soc. **8**, 260–280 (1961)
114. Valiant, L.: A bridging model for parallel computation. Commun. ACM **33**(B), 103–111 (1990)
115. Vardi, M.: The complexity of relational query languages. In: STOC'82, pp. 137–146. ACM (1982)
116. Vrba, J.: General decomposition problem of fuzzy relations. Fuzzy Sets Syst. **54**(1), 69–79 (1993)
117. Weber, G.-W., Savku, E., Kalayci, B., Akdogan, E.: Stochastic Optimal Control of Impulsive Systems under Regime Switches and Paradigm Shifts, in Finance and Economics, Biology and "Human Sciences", Seminar at Institute of Computer Science, Poznan University of Technology, October 17 (2017)
118. Weiss, G. (ed.): Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press, Cambridge, MA, USA (1999)
119. Wooldridge, M.: An Introduction to MultiAgent Systems, 2nd edn. Wiley (2009)
120. Wooldridge, M., Jennings, N.: Intelligent agents: theory and practice. Knowl. Eng. Rev. **10**, 115–152 (1995)
121. Yang, Y., Wang, X.: On the convergence exponent of decomposable relations. Fuzzy Sets Syst. **151**(2), 403–419 (2005)
122. Yang, Y., Wang, X.: The general α-decomposition problem of fuzzy relations. Inf. Sci. **177**(22), 4922–4933 (2007)

123. Ying, H.: Structural decomposition of the general MIMO fuzzy systems. Int. J. Intell. Control Syst. **1**(3), 327–337 (1996)
124. Zadeh, L.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
125. Zhong, C.: A new approach to generate Fuzzy system. In: Proceeding of the IEEE Singapore International Symposium on Control Theory and Application, pp. 250–254 (1997)