Modeling and Simulation in Science, Engineering and Technology

Nicola Bellomo
Livio Gibelli
Editors

Crowd Dynamics, Volume 3

Modeling and Social Applications in the Time of COVID-19

Birkhäuser

Birkhäuser

Modeling and Simulation in Science, Engineering and Technology

More information about this series at https://link.springer.com/bookseries/4960

Nicola Bellomo • Livio Gibelli
Editors

Crowd Dynamics, Volume 3

Modeling and Social Applications in the Time of COVID-19

Birkhäuser

Editors
Nicola Bellomo
Politecnico di Torino, Torino, Italy

University of Granada
Granada, Spain

Livio Gibelli
Institute for Multiscale Thermofluids
University of Edinburgh
Edinburgh, UK

ISSN 2164-3679 ISSN 2164-3725 (electronic)
Modeling and Simulation in Science, Engineering and Technology
ISBN 978-3-030-91645-9 ISBN 978-3-030-91646-6 (eBook)
https://doi.org/10.1007/978-3-030-91646-6

Mathematics Subject Classification: 34H05, 35Q93, 90B20, 91B99, 93A30

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Crowd dynamics has been receiving a constantly growing attention in the last two decades. Besides the theoretical interest, a realistic modeling of pedestrian behavior may lead to relevant societal benefits, for example, improved design of buildings, aircraft, and ships, with respect to their safety in the event of an emergency evacuation and/or optimized management of a crowd during gathering events.

Modeling and simulation of human crowds pose a formidable challenge due to the multidisciplinary approach that is involved. Indeed, the formulation of computational models relies on "hard" natural sciences, but "soft" social sciences are also needed, especially if one attempts to capture pedestrian behavior in crisis situations, like a rapid evacuation due to incidents or when the crowd includes groups of activists that confront each other.

The recent outbreak of the Covid-19 pandemic has spurred the interest for another application of crowd dynamics, namely the prevention of the spreading of contagious diseases. This target may be achieved by coupling a contagion model to a model of human crowds which takes into account how social distancing as well as the awareness to the risk of contagion affect pedestrian behavior.

This edited book comprises nine chapters with contributions from leading experts in the field, and aims at presenting the state of the art, challenges, and future research perspectives in the area of modeling and simulation of human crowds as well as at providing practical guidelines for crowd management. As done in the previous two books of this series, the topics are covered from different perspectives, thus providing a comprehensive overview on the work carried out in this challenging research area.

While this edited book does not cover all the possible topics, we think that it fosters a deeper understanding of pedestrian dynamics and may help foreseeing future research directions.

Torino, Italy — Nicola Bellomo
Edinburgh, UK — Livio Gibelli

September 2021

Contents

Behavioral Human Crowds: Recent Results and New Research Frontiers

Nicola Bellomo and Livio Gibelli

Abstract This editorial chapter provides an introduction to the contents of this edited book and a general critical analysis which looks ahead to research perspectives. The presentation is organized in three parts. In the first part some key research topics are selected based not only on their theoretical interest but also on the potential impact that may have on the society well-being. The second part outlines the contents of the following chapters in light of the aforementioned key topics as well as of the preceding edited books (N. Bellomo and L. Gibelli, Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems, Birkhäuser, New York, 2018; L. Gibelli, Crowd Dynamics, Volume 2 - Theory, Models, and Applications, Birkhäuser, New York, 2020). The last part speculates on promising future research directions.

1 Introduction

Crowd dynamics has attracted enormous attention in recent years not only for its theoretical interest but also for the potential societal benefits. As an example, computational models of crowd movement can lead to more efficient transportation planning, a key driver of sustainability, thus reducing the cost of transportation, pollution, and improving the population's quality of life. Furthermore, these tools can contribute to city security and safety in that pedestrians/vehicles may be used as sensors to identify threats that compromise the safety of persons and infrastructures (e.g., in natural disasters or acts of terrorism).

N. Bellomo
Politecnico di Torino, Torino, Italy

University of Granada, Granada, Spain IMATI CNR, Pavia, Italy
e-mail: bellomo@imati.cnr.it

L. Gibelli (✉)
Institute for Multiscale Thermofluids, University of Edinburgh, UK
e-mail: livio.gibelli@ed.ac.uk

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_1

It is becoming increasingly clear that modeling and simulation of crowds is a challenging interdisciplinary research field which requires contributions from different disciplines, ranging from technology, which is needed to detect the main features of crowds, to mathematics and computational sciences, which allow one to derive models and to simulate pedestrians' dynamics, respectively.

Human psychology can also significantly contribute to crowd modeling [42] especially if one attempts to capture the pedestrians' behavior in crisis situations, like a rapid evacuation due to incidents or when the crowd includes groups of activists that confront with each other [28]. In all these applications, the pedestrian's dynamics is strongly influenced by social interactions [1, 27, 34] which contribute to spread out unusual behaviors through the crowd.

Including social interactions in the modeling and simulation of crowds requires to adopt a behavioral perspective, namely one must include those aspects that may explain when and why individuals behave as they do. The key concepts of behavioral dynamics have been recently reviewed in [33] by examining more than 400 articles. The author sharply discusses how individual behaviors are modified by interactions within a crowd and how walkers adapt their walking strategy to the collective dynamics.

This editorial chapter presents an overview of this edited book which addresses various aspects of modeling, simulations, and control of the dynamics of human crowds. Key references are the preceding Volumes 1 [9] and 2 [31] which provided important contributions to this research area and offered insightful suggestions for future studies.

We start by briefly presenting three key topics which will probably form the focus of future research activities. These topics are selected according to our own bias and, although they do not encompass all the current open problems, their discussion paves the way to a deeper understanding of the contents of this edited book and may help in foreseeing the future outlooks in this challenging research field.

Key Topic 1: As clearly shown by the preceding two edited volume, most of the mathematical models of crowd dynamics have been mainly focused on safety problems, starting from the pioneering papers by Helbing [35, 36]. These include modeling crowd dynamics in complex environment [21, 30, 40] and crowd control somehow related to safety [2, 3, 8, 17], to cite a few. Although, this are still crucial applications, new research trends are arising. This raises the question: How far can the modeling and simulation of crowd dynamics contribute to different fields of interest for our society?

Key Topic 2: In many instances, there is a complex interplay between the dynamics of a crowd and additional phenomena that occur within the crowd itself. A specific example consists of crowd with emotional contagion [12, 29] as the emotions spreading may significantly affect the pedestrians behavior and, eventually, the whole crowd. An even more timely application, motivated by the recent outbreak of Covid-19 pandemic, is the spreading of contagious diseases through a crowd. This raises the question: How can the complex interplay between crowd dynamics and emotional/contagious diseases be modeled?

Key Topic 3: The literature on crowd modeling indicates that the derivation of models is generally carried out at one of the three usual representation scales, i.e., the micro-scale (individual-based models), the meso-scale (kinetic models), and the macro-scale (hydrodynamical models). On the other hand, the human crowd dynamics shows multiscale features that, by definition, cannot be fully capture adopting a single representation scale as discussed in detail in [6]. This raises the question: How can a multiscale vision be developed?

The contents of this edited book are mainly focused on the first two key topics, albeit some contributions are also given to the third one, as all these three topics are closely interconnected. The common feature pervading all chapters is the focus on pedestrians' communications which is a crucial aspect of crowds' psychology [42].

The rest of this chapter is organized as follows. Section 2 outlines the contents of the chapters that comprise this edited book. Section 3 proposes some perspective speculations for carrying out a research activity with significant impact.

2 On the Contents of the Edited Book

The chapters of this edited book deal with front-edge research topics in the modeling, simulation, and control of crowd dynamics. Some contributions refer to new frontiers of crowd modeling, specifically the complex interplay between virus spreading and crowd dynamics with or without awareness of the contagion issue. These chapters closely refer to Key Topics 1 and 2.

More specifically, the first five chapters focus on methods for collecting empirical data about pedestrians' behavior, and on the different approaches for managing crowds. Although, these chapters provide important contributions, there are still many problems left open, as we will discuss in Sect. 3 referring specifically to the multiscale approach.

Chapter "Generalized Solutions to Opinion Dynamics Models with Discontinuities" is devoted to the modeling of social dynamics, specifically opinion formation. This topic is closely related to one of the key problems posed in [1], namely how pedestrians interact with their neighbors, and adjust their walking strategy accordingly. This chapter considers state dependent interactions with a fixed number of nearest neighbors and provide a sharp discussion on how to deal with discontinuous interactions between agents.

Chapter "Crowd Behaviour Understanding Using Computer Vision and Statistical Mechanics Principles" discusses the possibility of classifying crowds via thermodynamics-like parameters, such as energy and entropy, and to measure these parameters by means of computer vision techniques. This study is part of a larger research program that aims at systematically detect behavior anomalies or abnormality within a crowd.

Chapter "Applications of Crowd Dynamic Models: Feature Analysis and Process Optimization" focuses on feature analysis and process optimization techniques for evacuation management. It is clearly shown that the literature in this field still suffers from many limitations that hinder practical applications. For example, there is little scope for error in managing crowd evacuation, and few emergencies can be dealt with using the same approach. This chapter delineates specific tools for tackling this key safety problem.

Chapter "Optimized Leaders Strategies for Crowd Evacuation in Unknown Environments with Multiple Exits" investigates the role of leaders in controlling the dynamics of a crowd. The follower-leader dynamics is initially described at the microscopic scale by an agent-based model, and, subsequently, a mean-field type model is derived to approximate a crowd composed of many followers. A meta-heuristic approach is used to optimize the leaders' walking strategy based on specific objectives (i.e., minimization of evacuation time, maximization of evacuated pedestrians, optimal use of exits), and it is shown that leaders may effectively guide pedestrians to safely egress unknown environments.

The last four chapters focus on different aspects of crowd dynamics closely related to the contagion problem which is a hot topic since the Covid-19 pandemic outbreak [12]. These include physical distancing in crowds and virus transmission, crowd dynamics in the presence awareness of the risk of contagion, epidemiological models adapted to modeling the dynamics over networks, and description of heterogeneous populations within the agent-based modeling framework.

Chapter "The Impact of Physical Distancing on the Evacuation of Crowds" first reviews the implications of physical distancing on crowd dynamics both in normal conditions and emergencies. Then, it provides a detailed assessment of expected changes on the crowd evacuation behavior due to awareness of the contagion risk, including changes in the fundamental walking speed/density and flow/density relationships.

Chapter "A Kinetic Theory Approach to Model Crowd Dynamics with Disease Contagion" presents some perspective ideas on how to extend a kinetic-type model for crowd dynamics to account for an infectious disease spreading. The authors refer to recent developments of the mathematical theory of active particles [13] where the modeling approach is based on kinetic theory methods, and theoretical tools of game theory are used to model the interactions involving walkers. The model is tested on a problem involving a small crowd walking through a corridor, but an application to realistic scenarios is also presented, namely the passengers behavior in one terminal of Hobby Airport in Houston.

Chapter "Toward a Quantitative Reduction of the SIR Epidemiological Model" develops a modeling approach of SIR-type epidemiological models in the context of dynamic networks, where model reduction and coarse-graining techniques are applied to reduce the computational complexity of systems with a large number of degrees of freedom. The gap between the approximate and the original observable quantities is studied by analytic and computational methods. This approach appears suitable to model the dynamics of large-scale highly interacting inhomogeneous human crowds to gain a more fundamental understanding of viruses spreading.

Chapter "An Agent-Based Model of COVID-19 Diffusion to Plan and Evaluate Intervention Policies" shows how agent-based methods, based on a sharp selection of behavioral and interaction rules, can describe the Covid-19 pandemic spreading. The model uses data of Piedmont, an Italian region, albeit a straightforward calibration is possible using data of other geographic areas as well. The authors show how heterogeneous interacting populations can be simulated in a broad variety of realistic scenarios. Finally, the chapter introduces strategic planning of vaccinations using genetic algorithms. The hints delivered by this chapter, which closes the edited book, once properly interpreted at a technical level, can become a worthwhile legacy to scientists active in crowd modeling.

3 Research Perspectives

The contents of this edited book clearly show that Key Topics 1 and 2 are object of constant interest. Many interesting results have already been obtained and soon more will follow as many scientists are actively working in the field as reported in [33].

Key Topic 1 can be considered a relatively mature research field, and many important contributions have been given in the past decades. However, there are still many open questions and much room for modeling improvements. As an example, the development of discrete velocity models at the mesoscopic scale is a topic which certainly deserves further attention. In fact, in many instances, the crowd does not include a number of pedestrians high enough to justify the assumption of continuity of the probability distribution over the micro-state which, instead, it is required by the standard kinetic theory approach [10].

Discrete models are based on the idea of partitioning the velocity space into a finite number of sub-domains so that the representation of the system is delivered by the number of individual entities in each domain. This approach has been applied to the modeling of vehicular traffic by two different modeling strategies, i.e., using a fixed grid [26], and a grid depending on the local density [25]. A model with discrete velocity directions has been proposed in [11], where the speed is related to local density by heuristic models interpreting empirical data. Similar modeling strategies can be envisioned for crowd dynamics.

Key Topic 2 represents a rapidly expanding research area. Not by chance, half of this edited book is exactly about the interplay between crowd dynamics and contagion spreading. However, note that Key Topic 2 embraces a broader range of applications, ranging from panic arising in emergency situations [34, 43] to emergence of violence during protests [28].

Most of the models available in the literature considers the emotional-social state as constant parameter equally shared by all pedestrians. On the other hand, recent studies consider the emotional-social state as a dynamical variable, based on both continuous [14, 15, 44] and discrete [37, 38] velocity models. We do think that the development of these latter models is a promising research direction.

Compared to others, the Key Topic 3 is relatively less developed but, hopefully, many contributions will be given in the coming years. Indeed, the derivation of


(a) (b)

Fig. 1 The outbreak of the Covid-19 pandemic is not (**a**) a black swan but a consequence of a (**b**) strongly interconnected world

macroscopic (hydrodynamic) models from the underlying description at the micro-scale (i.e., individual-based or kinetic) remains the cornerstone for fully capture the complexity features of the crowd dynamics.

A few pioneering papers have already been devoted to this topic, e.g., [18]. A micro-macro derivation of models in unbounded domains has been developed in [10] referring specifically to the discrete velocity model proposed in [11]. The analytic problem presents conceptual difficulties somehow related to the sixth Hilbert problem, see [19, 20]. A first step of the quest towards the micro-macro derivation consists in deriving models at each scale by the same principles and by using parameters corresponding to the same physical dynamics [6].

As witnessed by the vast literature [33], including this series of edited books, crowd dynamics remains an incredibly active research field. The specific feature of this volume consists in bringing to the attention of researchers the great new variety of physical scenarios which is added to classical topics such as collecting empirical data [23, 24, 32, 41], studying analytic problems [4, 7, 22, 39], and developing more realistic computational models [16]. Indeed, most of the contributions of this book were motivated by the ongoing Covid-19 pandemic which has further increased the fragility of our already fragile planet [5]. No space to reason about a black swan. Arguably, modeling, simulations, and artificial intelligence can address research activity in crowd dynamics to support the decision making of crisis managers. Then, scientists are asked to contribute achieving this challenging objective (Fig. 1).

References

1. G. Ajmone Marsan, N. Bellomo, and L. Gibelli, Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics, *Mathematical Models and Methods in Applied Sciences*, **26**, 1051–1093, (2016).

2. G. Albi, M. Bongini, E. Cristiano, and D. Kalise, Invisible control of self-organizing agents leaving unknown environments, *SIAM Journal of Applied Mathematics*, **76(4)**, 1683–1710, (2016).
3. G.Albi, E. Cristiani, L. Pareschi, and D. Peri, Mathematical models and methods for crowd dynamics control, Chapter 8 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 159–198, (2020).
4. B. Andreianov, C. Donatello, U. Razafison, and M. D. Rosini, One-dimansional conservation laws with nonlocal point constraints on the flux, in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 103–135, (2018).
5. B. Avishai, The pandemic isn't a black swan but a portent of a more fragile global system, *The New Yorker*, April 21, (2020). https://www.newyorker.com/news/daily-comment/the-pandemic-isnt-a-black-swan-but-a-portent-of-a-more-fragile-global-system
6. B. Aylaj, N. Bellomo, L. Gibelli, and A. Reali, On a unified multiscale vision of behavioral crowds, *Mathematical Models and Methods in Applied Sciences*, **30(1)**, 1–22, (2020).
7. R. Bailo, J.A. Carrillo, and P. Degond, Pedestrian models based on rational behaviors, Chapter 9 in **Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems**, *Series: Modelling Simulations Science Engineering Technology*, Birkhäuser, New York, 259–292, (2018).
8. M.K. Banda, M. Herty, and T. Trimborn, Recent developments in controlled crowd dynamics, Chapter 7 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 133–158, (2020).
9. N. Bellomo and L. Gibelli, **Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, (2018).
10. N. Bellomo and A. Bellouquid, On multiscale models of pedestrian crowds from mesoscopic to macroscopic, *Communication in Mathematical Sciences*, **13(7)**, 1649–1664, (2015).
11. N. Bellomo, A. Bellouquid, and D. Knopoff, From the micro-scale to collective crowd dynamics, *Multiscale Modelling Simulations*, **11**, 943–963, (2013).
12. N. Bellomo, R. Bingham, M. Chaplain, G. Dosi, G. Forni, D. Knopoff, J. Lowengrub, R. Twarock, and M.E. Virgillito, A multi-scale model of virus pandemic: Heterogeneous interactive entities in a globally connected world, *Mathematical Models and Methods in Applied Sciences*, **30**, 1591–1651, (2020).
13. N. Bellomo, D. Burini, G. Dosi, L. Gibelli, D. A. Knopoff, N. Outada, P. Terna, and M.E. Virgillito, What is life? A perspective of the mathematical kinetic theory of active particles, *Mathematical Models and Methods in Applied Sciences*, **31**, 1821–1866 (2021) https://doi.org/10.1142/S0218202521500408.
14. N.Bellomo, L. Gibelli, and N. Outada, On the interplay between behavioral dynamics and social interactions in human crowds, *Kinetic and Related Models*, **12**, 397–409, (2019).
15. A.L. Bertozzi, J. Rosado, M.B. Short, and L. Wang, Contagion shocks in one dimension, *Journal Statistical Physics*, **158**, 647–664, (2015).
16. R. Borsche, A. Klar, and F. Schneider, Numerical methods for mean-field and moment models for pedestrian flow, Chapter 7 in Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 167–210, (2018).
17. A. Borzì , The Fokker-Planck framework in the modelling of pedestrians' motion, Chapter 6 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 111–132, (2020).
18. M. Burger, P. Markowich, J.F. Pietschmann, Continuous limit of a crowd motion and herding model: analysis and numerical simulations, *Kinetic Related Models*, **4(4)**, 1025–1047, (2011).
19. D. Burini and N. Chouhad, Hilbert method toward a multiscale analysis from kinetic to macroscopic models for active particles, *Mathematical Models and Methods in Applied Sciences*, **27(7)**, 1327–1353, (2017).

20. D. Burini and N. Chouhad, A Multiscale view of nonlinear diffusion in biology: From cells to tissues, *Mathematical Models and Methods in Applied Sciences*, **29(4)**, 791–823, (2019).
21. M. Colangeli, A. Muntean, O. Richardson, and T. Thieu, Modelling interactions between active and passive agents moving through heterogeneous environments, in **Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 211–258, (2018).
22. R. M. Colombo, M. Lecureux–Mercier, and M. Garavello, Crowd dynamics through conservation laws, Chapter 9 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, (2020).
23. A. Corbetta, A. Mountean, and K. Vafayi, Parameter estimation of social forces in pedestrian dynamics models via probabilistic method, *Mathematical Biosciences Engineering*, **12**, 337–356, (2015).
24. A. Corbetta, L. Schilders, and F. Toschi, High-statistics modeling of complex pedestrian avoidance scenarios, Chapter 4 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, (2018).
25. V. Coscia, M. Delitala, and P. Frasca, On the mathematical theory of vehicular traffic flow models II. Discrete velocity kinetic models, Int. J. Non-linear Mech. 42 (2007) 411–421.
26. M. Delitala and A. Tosin, Mathematical modelling of vehicular traffic: A discrete kinetic theory approach, *Mathematical Models and Methods in Applied Sciences*, **17**, 901–932, (2007).
27. E. Cristiani, B. Piccoli, and A. Tosin, **Multiscale Modeling of Pedestrian Dynamics**, Springer, (2014).
28. J.-M. Epstein, Modeling civil violence: An agent based computational approach, *Proceeding National Academy Sciences*, **99**, 7243–7250, (2002).
29. E. Estrada, COVID-19 and SARS-CoV-2. Modeling the present, looking at the future, *Physics Reports*, **869**, 1–51, (2020).
30. Hideki Fujii, Hideaki Uchida, Tomonori Yamada and Shinobu Yoshimura, Mixed Traffic Simulation of Cars and Pedestrians for Transportation Policy Assessment, Chapter 2 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, (2020).
31. Crowd Dynamics, Volume 2 - Theory, Models, and Applications, L. Gibelli, Ed., *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 1–10, (2020).
32. S. Göttlich and S. Knapp, Artificial Neural Networks for the Estimation of Pedestrian Interaction Forces, Chapter 3 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, (2020).
33. M. Haghani, The knowledge domain of crowd dynamics: Anatomy of the field, pioneering studies, temporal trends, influential entities and outside-domain impact, *Physica A*, **2021**, paper n. 126145, (2021).
34. M. Haghani and M. Sarvi, Social dynamics in emergency evacuations: Disentangling crowds attraction and repulsion effects, *Physica A*, **475**, 24–34, (2017).
35. D. Helbing D. and A. Johansson, Pedestrian crowd and evacuation dynamics, *Encyclopedia of Complexity and System Science*, Springer, 6476–6495, (2009).
36. D. Helbing, A. Johansson, and H.-Z. Al-Abideen, Dynamics of crowd disasters: An empirical study, *Physical Review E*, **75**, paper no. 046109, (2007).
37. D. Kim and A. Quaini, A kinetic theory approach to model pedestrian dynamics in bounded domains with obstacles. *Kinetic and Related Models*, **12**, 1273–1296, (2019).
38. D. Kim and A. Quaini, Coupling kinetic theory approaches for pedestrian dynamics and disease contagion in a confined environment, *Mathematical Models and Methods in Applied Sciences*, **30**, 1893–1915, (2020).
39. B. Piccoli and F. Rossi, Measure-theoretic models for crowd dynamics, in **Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 137–166,(2018).

40. E. Ronchi and D. Nilsson, Pedestrian Movement in Smoke: Theory, Data and Modelling Approaches, Chapter 3 in **Crowd Dynamics, Volume 1 - Theory, Models, and Safety Problems**, *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 37–62. (2018).
41. A. Schadschneider, M. Chraibi, A. Seyfried, A. Tordeux, and J. Zhang, Pedestrian dynamics: from empirical results to modeling, Chapter 4 in *Crowd Dynamics, Volume 1*, *Series: Modelling Simulations Science Engineering Technology*, Birkhäuser, New York, 63–102, (2018).
42. A. Templeton and F. Neville, Modelling collective behaviour: Insights and applications from crowd psychology, Chapter 5 in **Crowd Dynamics, Volume 2 - Theory, Models, and Applications**, L. Gibelli, Ed., *Modeling and Simulation in Science, Engineering, and Technology*, Birkhäuser, New York, 55–82, (2020).
43. H. Vermuyten, J. Belien, L. De Boeck, G. Reniers, and T. Wauters, A review of optimisation models for pedestrian evacuation and design problems, *Safety Science*, **87**, 167–178, (2016).
44. L. Wang, M.B. Short, and A.L. Bertozzi, Efficient numerical methods for multiscale crowd dynamics with emotional contagion, *Mathematical Models and Methods in Applied Sciences*, **27**, 205–230, (2017).

Generalized Solutions to Opinion Dynamics Models with Discontinuities

Francesca Ceragioli, Paolo Frasca, Benedetto Piccoli, and Francesco Rossi

Abstract Social dynamics models may present discontinuities in the right-hand side of the dynamics for multiple reasons, including topology changes and quantization. Several concepts of generalized solutions for discontinuous equations are available in the literature and are useful to analyze these models. In this chapter, we study Caratheodory and Krasovsky generalized solutions for discontinuous models of opinion dynamics with state dependent interactions. We consider two definitions of "bounded confidence" interactions, which we, respectively, call metric and topological: in the former, individuals interact if their opinions are closer than a threshold; in the latter, individuals interact with a fixed number of nearest neighbors. We compare the dynamics produced by the two kinds of interactions in terms of existence, uniqueness, and asymptotic behavior of different types of solutions.

1 Introduction and Summary of Results

In the last decades, researchers from many different fields explored the behavior of large systems of active particles or agents. The latter entities, also called self-propelled, intelligent, or greedy, are endowed with the capability of decision making and, usually, of altering the energy or other (otherwise conserved) quantities of the system. Examples include dynamics of opinions in social networks, animal groups, networked robots, pedestrian dynamics, and language evolution. Their

F. Ceragioli
DISMA, Politecnico di Torino, Torino, Italy

P. Frasca
Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, GIPSA-lab, Grenoble, France

B. Piccoli (✉)
Department of Mathematical Sciences, Rutgers University-Camden, Camden, NJ, USA
e-mail: piccoli@camden.rutgers.edu

F. Rossi
Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, Padova, Italy

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_2

dynamics is written as an Ordinary Differential Equation (ODE in the following) in large dimension. In order to cope with this large dimension, various mean-field, kinetic and hydrodynamic limit descriptions were studied in the literature, see [1, 2, 8, 13, 14, 22, 25, 37] and the references therein. The interaction among agents may be restricted to specific regions due to the physical aspects of the modeled phenomenon, giving rise to discontinuities. Even more, due to modeling choices these discontinuities may appear naturally at multiple scales, see for instance [6, 23].

One of the main phenomena of active particles is *self-organization* of the whole system, stemming from simple interaction rules at the particle level. Such interaction rules are often motivated by relationships among agents; thus, corresponding evolutions are referred to as *social dynamics* [5, 40, 41]. The most common self-organized configurations are: consensus [36], i.e. all agents reaching a common state; alignment, i.e. agents reaching consensus on a subset of the state variables (e.g. speed) [12]; and clustering, i.e. agents grouping in a small number of well-separated states [32, 35].

The description of social dynamics may require continuous or discrete quantities, and the corresponding ODE may have either continuous or discontinuous vector fields. As a matter of fact, there are multiple situations where discrete variables and discontinuities arise. A partial list includes:

– the presence of threshold effects caused by physical, communication, or psychological barriers [30];
– the presence of quantities taking values in discrete sets, when a finite number of choices is given (e.g. whether and which product to buy) or when communication takes place by means of a finite set of symbols [16, 19, 34];
– the presence of a pattern of allowed/forbidden interactions, such as can be encoded in a graph [5, 7].

The latter case includes all situations where physical or cognitive constraints limit interactions to agents that are close to each other, either spatially or behaviorally. When models are defined in discrete time, discontinuities of the right-hand side pose little mathematical difficulties and are easily managed or simply ignored. Instead, *discontinuities give rise to technical difficulties in continuous time*, as the study of ODEs is deeply based on the notions of continuity and differentiability. We notice that the problem of dealing with discontinuities is not limited to the case of ODE models, but it also occurs at other scales for learning dynamics in crowds combined with loss of symmetry features, see [6].

Even though discontinuities of some models can be avoided by defining suitable smoothed counterparts, which feature continuous approximations of discontinuous functions, the connections between continuous and discontinuous variants are not trivial [15]. Most importantly, discontinuities cannot always be avoided. This is the case when the agents are allowed a finite number of choices/actions/interactions. Another, classical, example of unavoidable discontinuity in ODEs comes from control theory. Whereas stabilizability of a system with inputs by means of a continuous feedback law implies asymptotic controllability by means of an open-loop control, the converse does not hold if only continuous feedback laws are considered. But if discontinuous feedback laws are allowed, then asymptotic controllability does imply stabilizability [3, 20]. This example shows not only that discontinuities cannot be avoided, but they may be helpful. This result and, more generally, all results on discontinuous ODEs depend on the notion of generalized solutions adopted: it is interesting to analyze and compare different notions in order to deduce common features and differences. This is one of the objectives of the present paper: we will make use of the main concepts of solutions that have been defined in mathematical analysis and in control theory, and in particular we will discuss *classical, Caratheodory, Filippov, and Krasovsky solutions*. We recall the precise definitions of these solutions in Sect. 2.1 below.

We now describe more precisely the two opinion dynamics models that we analyze in the present chapter. The fact that an individual influences those he communicates with can be taken as a principle when describing evolution of opinions. The most basic model which describes in a mathematical framework this principle is usually referred to as DeGroot's model (despite having earlier origin in French [29]). Its main feature is that in a group of individuals that communicate among them, consensus is achieved. On the other hand, everyone's experience suggests that consensus is not always achieved among individuals. For this reason, many researchers have proposed more complex models, aiming to describe agreement and disagreement at the same time: see [40, 41] for a comprehensive discussion. Crucially, several of these more complex models feature discontinuities of the right-hand side.

Here we consider a general model, which incorporates the well-known Hegselmann-Krause model [30]. The basic idea of Hegselmann and Krause is that trust towards others has some limitations. In their work, they assume that an individual is influenced by others only if opinions are not too far from one another. Here, we describe the fact that one's confidence towards others is limited, by describing in two different ways the set of neighbors of an individual. In the first setting, interactions among individuals follow Hegselmann and Krause's rule: one's neighbors are individuals whose opinions do not differ too much. We call this kind of interactions *metric interactions*. In the second setting, we assume that an individual follows only a fixed number of neighbors, the ones whose opinions are the nearest to his own. We call this kind of interactions *topological interactions*. Topological interactions can be motivated by the notion of Dunbar number [7, 26] that indicates a cognitive limit in the number of significant relationships among individuals. This concept is particularly meaningful in the contemporary world, where potential contacts and available information seem to be unlimited.

1.1 Mathematical Models and Main Results

In the mathematical description of the bounded confidence models, we start by considering a set $V = \{1, \dots, N\}$ of N agents (also called individuals) with states $x_i \in \mathbb{R}^n$ (e.g. position, opinion, speed). Each agent $i \in V$ interacts with other agents belonging to a subset of neighbors $N_i(x) \subseteq V$. The subset of neighbors $N_i(x)$ depends on the state and induces a graph $G(x)$ of interactions among the agents: V is the set of nodes and (i, j) is an edge if $j \in N_i(x)$. We denote the set of edges by $E(x)$. The dynamics can be written in the following form:

$$\dot{x}_i = \sum_{j \in N_i(x)} a(\|x_j - x_i\|)(x_j - x_i). \tag{1}$$

The function $a : [0, +\infty[\to [0, +\infty[$ satisfies the following hypotheses:

- a is Lipschitz continuous;
- $a(r) > 0$ for $r > 0$;
- a is not decreasing.

The function a represents the strength of interactions among agents. A more general model could be written with interaction functions a_{ij} that depend on the pair of neighbors. Most results stated in this article remain valid in this more general setting, provided that interactions are symmetric ($a_{ij} = a_{ji}$). Depending on how neighbors $N_i(x)$ are chosen, one obtains different bounded confidence models. From now on, we will use the notations N_i^m, N_i^t for the set of neighbors for the metric and topological versions that we make explicit below.

In the *metric bounded confidence model* agent i's neighbors are those whose state is not too far from his own, namely

$$N_i^m(x) = \{j \in V : \|x_j - x_i\| < 1\}.$$

This choice implies that interactions between agent i and j are symmetric, i.e. agent i is influenced by agent j if and only if agent j is influenced by agent i. We then write the metric bounded confidence dynamics as follows:

$$\dot{x}_i = \sum_{j \in N_i^m(x)} a(\|x_j - x_i\|)(x_j - x_i). \tag{2}$$

As already mentioned, the first and best known version of the metric bounded confidence model is Hegselmann-Krause's [10, 30, 44], which corresponds to $a \equiv 1$ and was originally written in discrete time with states $x_i \in \mathbb{R}$. Its continuous-time counterpart was first studied in [11].

The *topological bounded confidence model* is obtained when agent i interacts only with a fixed number κ of neighbors, where $1 \le \kappa \le N - 1$. More precisely, for every agent $i \in V$, her neighborhood $N_i^t(x)$ is defined in the following way. The elements of $V \setminus \{i\}$ are ordered by increasing values of $\|x_j - x_i\|$; then, the first κ elements of the list (i.e. those with smallest distance from i) form the set $N_i^t(x)$ of current neighbors of i. Should a tie between two or more agents arise, priority is given to agents with lower index. We then write

$$\dot{x}_i = \sum_{j \in N_i^t(x)} a(\|x_j - x_i\|)(x_j - x_i). \tag{3}$$

For topological interactions, agent i could be influenced by agent j without agent j being influenced by agent i, namely interactions are not symmetric. This fact is a major difference between the metric and topological bounded confidence models. This model was first pointed out in [5], while several other models of opinion dynamics and collective motion have considered topological interactions in different forms: see [21, 42] and the references therein.

For both models, we have the following crucial observation: the right-hand side of (2)–(3) is a discontinuous function. For this reason, one needs to carefully select a concept of solution to such discontinuous ODE. In our opinion, this aspect has been overlooked in the extensive literature about bounded confidence models, with some exceptions such as [9, 11, 15]. Here, we will consider mainly Caratheodory and Krasovsky solutions. Definitions and a brief discussion on different notions of solutions can be found in Sect. 2.1. The first result about solutions of (1) will be the following.

Theorem 1 (Existence and Uniqueness) *Consider the bounded confidence models, either in the metric case* (2) *or topological case* (3)*. Then, there exists a solution (global in time) for every initial condition in the Krasovsky sense. Uniqueness of solutions does not hold in general but holds for almost every initial datum. Moreover, the same result holds for Caratheodory solution, both in the metric case and in the topological case for* $\kappa = 1$.

The full proof of the result for (2) was given in [38], which extended partial results from [11, 15, 17]. Here we prove the corresponding claims for (3) in Sect. 3.

After solving the questions about existence and uniqueness, we focus on some properties of such solutions that have been explored in the rich literature about social dynamics models. We want to recall some of them. In the next definitions $x(t) = (x_1(t), \dots, x_N(t))$ will denote a solution of an unspecified type.

P1) **Average preservation.** $x_{ave}(t) = \frac{1}{N}\sum_i x_i(t)$ is invariant along trajectories.

P2) **Contractivity of the support.** For all $T^2 \geq T^1 \geq 0$, it holds

$$\overline{co}\left(\left\{x_1(T^1), x_2(T^1), \dots, x_N(T^1)\right\}\right) \supseteq \overline{co}\left(\left\{x_1(T^2), x_2(T^2), \dots, x_N(T^2)\right\}\right),$$

where $\overline{co}$ is the closed convex hull of the values in the brackets (defined in (4) below).

P3) **Convergence to cluster points.** Every solution $x(t)$ converges for $t \to +\infty$ to a cluster point, namely to a point $x^\infty = (x_1^\infty, \dots, x_N^\infty) \in \mathbb{R}^{nN}$, $x_i^\infty \in \mathbb{R}^n$, such that for every $i \in V$, for every $j \in N_i(x^\infty)$ it holds $x_i^\infty = x_j^\infty$. Every set of agents with coincident states is said to be a cluster.

Properties P1), P2), P3) will be discussed for both metric and topological models. Many examples will show the richness of possible behaviors, depending on the

Table 1 This table summarizes, for the reader's convenience, where in paper the main properties of the bounded confidence models are proved or disproved

	P1	P2	P3
Metric Caratheodory	Yes Proposition 5	Yes Proposition 6	Yes Proposition 11
Metric Krasovsky	Yes Proposition 5	Yes Proposition 6	Yes Proposition 11
Topological Caratheodory	No Example 5	Yes Proposition 6	No Example 7
Topological Krasovsky	No Example 5	Yes Proposition 6	No Example 7
Topological Caratheodory $\kappa = 1$	No Example 5	Yes Proposition 6	Yes Proposition 11
Topological Krasovsky $\kappa = 1$	No Example 5	Yes Proposition 6	No Example 8

chosen notion of solution. Indeed, the following theorem summarizes the results that we prove in the next sections: the proof scheme is summarized in Table 1.

Theorem 2 (Properties of Solutions)

(i) ***Metric bounded confidence model.*** *Caratheodory and Krasovsky solutions to* (2) *satisfy properties P1)-P2)-P3).*
(ii) ***Topological bounded confidence model.*** *Caratheodory and Krasovsky solutions to* (3) *satisfy property P2) and may not satisfy properties P1) and P3).*
(iii) ***Topological bounded confidence model with one neighbor.*** *In addition to the previous case, Caratheodory solutions to* (3) *with* $\kappa = 1$ *satisfy property P3).*

Some additional facts are easy to observe.

Remark 1 (Structure of Cluster Points, Metric Case) Note that for the metric bounded confidence model, different values assumed by the components of a cluster point x^∞ are at a distance greater than or equal to one. Actually, they can be at distance precisely one, as shown by Example 1 below.

Example 1 (Clusters at Distance 1) Consider the system (2) with $n = 2$, $N = 3$ and initial condition $\overline{x} = \Big((0, 0), (1, \frac{1}{3}), (1, -\frac{1}{3})\Big)$. There is a unique Krasovsky (thus also Caratheodory) solution starting at $\overline{x}$ which converges to the cluster point $x^\infty = ((0, 0), (1, 0), (1, 0))$. Note that the distance between the first two agents in x^∞ is precisely one.

Remark 2 (Non-Exclusive Dependence of the Asymptotic State on the Initial Data) A desirable property for solutions of any system is that the asymptotic state depends on the initial datum only. Bounded confidence models fail to have this property, because different solutions starting from the same initial condition can have different asymptotic states. This is the case for Caratheodory (and a fortiori also for Krasovsky) solutions, as shown by Example 5 below.

Complete proofs of the properties of the metric bounded confidence model can be found in [38]. Here we recall the main ideas in order to compare metric and topological cases. In fact, the topological bounded confidence model turns out to be rather different and more complex to characterize. A key reason, already pointed out, is that interactions are not symmetric. As a consequence, even a characterization of equilibria is not evident. Here, we construct a Lyapunov function and prove convergence to cluster points in the case $\kappa = 1$ only. In this case, we can also characterize the configuration of the network induced by (3) at any time, which is a directed pseudo-forest with a cycle of length 2 in each connected component (Proposition 12). In the general case, counterexamples show that convergence to cluster points cannot be expected (Examples 7–8).

This picture shows the theoretical interest of these models and the long way to go to fully understand them.

2 Generalized Solutions: Definitions and Basic Facts

In this article, we denote by λ^m the Lebesgue measure on $\mathbb{R}^m$. For $x \in \mathbb{R}^m$, $B(x, r)$ is the ball of radius $r > 0$ centered at x and $B(r) = B(0, r)$ is the ball centered at the origin. The Euclidean norm in $\mathbb{R}^m$ is denoted by $\|\cdot\|$. Given an embedded manifold $M \subset \mathbb{R}^m$, the symbol ∂M denotes the topological boundary. Given $A \subset \mathbb{R}^m$, we denote by $int(A)$ its interior, by $\overline{A}$ its closure and we set

$$co(A) = \left\{ \sum_{i=1}^{\ell} \alpha_i x_i : \ell \in \mathbb{N}, \alpha_i \in [0, 1], \sum_{i=1}^{\ell} \alpha_i = 1, x_i \in A \right\} \tag{4}$$

the convex hull of A, and denote by $\overline{co}(A)$ its closure.

We denote by $AC([0, T], \mathbb{R}^m)$ the space of absolutely continuous functions on a time interval $[0, T]$. Recall that every absolutely continuous function is differentiable for almost every time, i.e. except for times on a set of zero Lebesgue measure. We also introduce the following:

Definition 1 (Stratified Set) A set $\Gamma \subset \mathbb{R}^m$, $\Gamma = \cup_{i=1}^{m_\Gamma} M_i$, with $m_\Gamma \in \mathbb{N} \cup \{+\infty\}$ and M_i being $\mathcal{C}^1$ embedded manifold of dimension $n_i \leq m$, is stratified if:

i) The family M_i is locally finite: given a compact K, it holds $K \cap M_i \neq \emptyset$ only for finite many i.
ii) for $i \neq j$ it holds $M_i \cap M_j = \emptyset$, and if $M_i \cap \partial M_j \neq \emptyset$, then $M_i \subset \partial M_j$ and $n_i < n_j$.

We call $\max_i n_i$ the dimension of the stratified set Γ.

Remark 3 For simplicity we used the definition of topological stratification, even if the examples we consider admit Whitney stratification. We refer the reader to [33, 39, 43] for a discussion of the different concepts and the role played for discontinuous ODEs and optimal feedback control.

An autonomous ODE is written as:

$$\dot{x}(t) = g(x(t)), \tag{5}$$

where $x \in \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^m$ is a measurable and locally bounded function (defined at every point). The different concepts of solution will be discussed in the next Sect. 2.1.

A multifunction on $\mathbb{R}^m$ is a function $H : \mathbb{R}^m \to \mathcal{P}(\mathbb{R}^m)$, with $\mathcal{P}(\mathbb{R}^m)$ being the powerset of $\mathbb{R}^m$, i.e. the set of subsets of $\mathbb{R}^m$. Given a multifunction H, one can consider the differential inclusion:

$$\dot{x}(t) \in H(x(t)). \tag{6}$$

A solution is an absolutely continuous function $x(\cdot)$ which satisfies (6) for almost every t.

We define the Hausdorff distance d_H on the powerset of $\mathbb{R}^m$ as follows: given $x \in \mathbb{R}^m$ and $A, B \subset \mathbb{R}^m$ we set $d(x, A) = \inf\{d(x, y) : y \in A\}$ and $d_H(A, B) = \sup\{d(x, A), d(y, B) : x \in B, y \in A\}$. A multifunction H is continuous if it is continuous for the Hausdorff distance, while H is upper semicontinuous at x if for every $\epsilon > 0$ there exists $\delta > 0$ such that $H(y) \subset H(x) + B(\varepsilon)$ for every y with $\|x - y\| < \delta$.

A continuous multifunction H is also upper semicontinuous. It is well known that if H is upper semicontinuous with compact convex values, then the corresponding differential inclusion (6) admits solutions (locally in time) for every initial condition, see [4]. More precisely, we state the following fact.

Proposition 1 *Assume that the multifunction H in (6) is upper semicontinuous and, for every $x \in \mathbb{R}^m$, $H(x)$ is a nonempty, compact, and convex subset of $\mathbb{R}^m$. Then, for every initial condition x_0 there exists a local solution to (6).*

2.1 Solutions to Discontinuous Ordinary Differential Equations

Given the ODE (5) with g discontinuous, it is convenient to define the associated Krasovsky multifunction, defined as:

$$K(x) = \bigcap_{\delta > 0} \overline{co}\{g(y) : y \in (x + B_\delta)\}. \tag{7}$$

Similarly, the Filippov multifunction is defined as:

$$F(x) = \bigcap_{\delta > 0} \bigcap_{\lambda^m(N) = 0} \overline{co}\{g(y) : y \in (x + B_\delta \setminus N)\}. \tag{8}$$

Many definitions of solutions for (5) are then available, most of which coincide when g is sufficiently regular (e.g. locally Lipschitz). We summarize in the following definition the concepts we are considering in the rest of the paper.

Definition 2 (Notions of Solution) Given the ODE (5) and $T > 0$ we define the following:

1. A **classical solution** is a differentiable function $x : [0, T] \to \mathbb{R}^m$ that satisfies (5) at every time $t \in (0, T)$. At 0 and at T the equation must be satisfied with one-sided derivatives.
2. A **Caratheodory solution** is an absolutely continuous function $x : [0, T] \to \mathbb{R}^m$ which satisfies (5) at almost every time $t \in [0, T]$.
3. A **Krasovsky solution** is an absolutely continuous function $x : [0, T] \to \mathbb{R}^m$, which satisfies:

$$\dot{x} \in K(x(t))$$

   for almost every time $t \in [0, T]$, with K given by (7).
4. A **Filippov solution** is an absolutely continuous function $x : [0, T] \to \mathbb{R}^m$, which satisfies:

$$\dot{x} \in F(x(t))$$

   for almost every time $t \in [0, T]$, with F given by (8).

We denote the sets of classical, Caratheodory, Filippov, and Krasovsky solutions with $\mathcal{C}l$, $\mathcal{C}a$, $\mathcal{F}$, and $\mathcal{K}$, respectively.

The concept of classical solution is not used for discontinuous ODEs, because of general lack of existence. In the following examples we show that both models may not admit a classical solution for some initial condition. In these examples and later in this paper, it will be convenient to denote the vector fields defined by the right-hand sides of the metric model (2) and the topological model (3) by f^m and f^t, respectively.

Example 2 (Non-Existence of Classical Solutions, Metric) Let $N = 3$, $n = 1$, $a \equiv 1$ and consider point $\overline{x} = (-\frac{1}{2}, 0, \frac{1}{2})$. Let f^m be the vector field defined by the right-hand side of (2). We have $f^m(\overline{x}) = (\frac{1}{2}, 0, -\frac{1}{2})$, in fact agents 1 and 3 do not communicate in this configuration. As soon as $t > 0$ agents 1 and 3 start communicating as $|x_3(t) - x_1(t)| < 1$ for $t > 0$. Then we have that $\lim_{t\to 0^+} f^m(x(t)) = (\frac{3}{2}, 0, -\frac{3}{2})$ which is different from $f^m(\overline{x})$. This proves that a classical solution issuing from $\overline{x}$ does not exist. If we take the initial condition $(-\frac{2}{3}, 0, \frac{2}{3})$, a classical solution exists until the state $\overline{x}$ is reached but cannot be continued up to $+\infty$.

Example 3 (Non-Existence of Classical Solutions, Topological) Consider (3) with $N = 4$, $n = 2$, $\kappa = 1$, $a \equiv 1$ and the initial condition $\bar{x} = ((-1, 0), (0, 0), (1, 0), (1 - \epsilon, \sqrt{1 - \epsilon^2}))$ with $0 < \epsilon < \frac{1}{2}$. Then $N_1^t(\bar{x}) = \{2\}$, $N_2^t(\bar{x}) = \{1\}$, $N_3^t(\bar{x}) = \{2\}$, $N_4^t(\bar{x}) = \{3\}$. Therefore $\dot{x}_3 - \dot{x}_2 = 0$ and $(\dot{x}_4 - \dot{x}_3) \cdot (x_4 - x_3) < 0$, thus for all positive times it holds $N_3^t(x(t)) = \{4\}$ and there is no classical solution.

Caratheodory solutions are among the ones commonly used, as they are equivalent to solutions in the integral form:

$$x(t) = x(0) + \int_0^t g(x(s))\, ds.$$

Existence theorems for Caratheodory solutions are far from trivial, as we will see in Sect. 2.1.

The concepts of Filippov and Krasovsky solutions are often used to deal with general discontinuous ODEs. They have the advantage of being based on the well-developed theory of differential inclusions (6), see [4, 27]. In particular, we have the following proposition, see [4].

Proposition 2 (Local Existence) *Consider an ODE (5) with g measurable and locally bounded. Then the corresponding Krasovsky and Filippov multifunctions K and F defined by (7) and (8), respectively, are upper semicontinuous with nonempty, compact and convex values. Thus, the differential inclusions $\dot{x} \in K(x)$ and $\dot{x} \in F(x)$ admit local solutions for every initial condition.*

Among solutions a special role is played by *equilibrium solutions*, whose notion should of course be adapted to the chosen concept of solution. More precisely, we give the following definition.

Definition 3 We call $\overline{x} \in \mathbb{R}^m$ an equilibrium with respect to classical (respectively, Caratheodory, Krasovsky, Filippov) solutions, if the function $\phi(t) = \overline{x}$ is a classical (respectively, Caratheodory, Krasovsky, Filippov) solution.

We remark that $\overline{x}$ is an equilibrium with respect to classical and Caratheodory solutions if and only if $f(\overline{x}) = 0$. This fact implies that Caratheodory and classical equilibria coincide. Instead, $\overline{x}$ is an equilibrium with respect to Krasovsky (respectively, Filippov) solutions if and only if $0 \in K(\overline{x})$ (respectively, $0 \in F(\overline{x})$).

2.2 Inclusions Between Sets of Solutions

In this section, we study the inclusions between the different concepts of solutions introduced above. We first recall the standard inclusions between solutions that do not depend on the specific structure of (1). The proof is omitted, as it directly follows from definitions.

Proposition 3 (Solution Sets) *The following inclusions among sets of solutions hold true: $\mathcal{C}l \subseteq \mathcal{C}a \subseteq \mathcal{K}$ and $\mathcal{C}l \subseteq \mathcal{F} \subseteq \mathcal{K}$.*

We now prove that in the specific case of dynamics (2) and (3) the sets of Filippov and Krasovsky solutions actually coincide. In view of this result, in the rest of this paper we will no longer distinguish between Krasovsky and Filippov solutions and we will simply refer to them as to Krasovsky solutions. The proof is based on the following fact.

Lemma 1 (Lemma 2.8 in [31]) *Let* $f : \mathbb{R}^m \to \mathbb{R}^m$ *be such that:*

(i) there exist $M_\alpha \subseteq \mathbb{R}^m$, $\alpha \in \mathcal{A}$*, such that* $\cup_\alpha M_\alpha = \mathbb{R}^m$*,* $M_\alpha \cap M_\beta = \emptyset$ *for all* $\alpha, \beta \in \mathcal{A}$*,* $\alpha \neq \beta$*, and* $M_\alpha \subseteq \overline{\mathrm{int}(M_\alpha)}$ *for all* $\alpha \in \mathcal{A}$*,*
(ii) there exist $f_\alpha : \mathbb{R}^m \to \mathbb{R}^m$ *continuous such that* $f(x) = f_\alpha(x)$ *for all* $x \in M_\alpha$ *and for all* $\alpha \in \mathcal{A}$*.*

Then $\mathcal{K} = \mathcal{F}$ *for* (5)*.*

Proposition 4 (Krasovsky and Filippov Solutions Coincide) *For the metric model* (2) *and the topological model* (3)*, it holds* $\mathcal{K} = \mathcal{F}$*.*

Proof The system (1) can be written in standard form (5) by setting $m = nN$, $x = (x_1, \dots, x_N) \in \mathbb{R}^{nN}$, $f = (f_1, \dots, f_N)$ with $f_i : \mathbb{R}^n \to \mathbb{R}^n$ given by the right-hand side of (1).

We start considering the *metric bounded confidence model* (2). Given $i, j \in V$, $i \neq j$, we define the subset of $\mathbb{R}^{nN}$:

$$\Delta^m_{ij} = \{(x_1, \dots, x_N) \in \mathbb{R}^{nN} : \|x_i - x_j\| = 1\}, \tag{9}$$

and the union of such subsets as:

$$\Delta^m = \cup_{i,j:i\neq j} \Delta^m_{ij}. \tag{10}$$

The set Δ^m is the set of points at which the right-hand side of (2) fails to be continuous. $\mathbb{R}^{nN}$ is the disjoint union of $p = 2^{\binom{N}{2}}$ sets such that f^m restricted to each of them is continuous. We can enumerate these sets by starting with $M_1 = \{x \in \mathbb{R}^{nm} : \|x_i - x_j\| < 1\, \forall i, j \in V, i \neq j\}$, $M_2 = \{x \in \mathbb{R}^{nm} : \|x_i - x_j\| < 1\, \forall i, j \in V$ except for $\|x_1 - x_N\| \leq 1\}$ and finishing with $M_p = \{x \in \mathbb{R}^{nm} : \|x_i - x_j\| \geq 1\, \forall i, j \in V\}$. Since $M_1, \dots, M_p$ and f^m satisfy the assumptions of Lemma 1, then $\mathcal{K} = \mathcal{F}$ for (2).

An analogous argument can be repeated for the *topological bounded confidence model* (3). In this case we denote by

$$\Delta^t_{ijh} = \{(x_1, \dots, x_N) \in \mathbb{R}^{nN} : \|x_j - x_i\| = \|x_h - x_i\|\} \tag{11}$$

and by

$$\Delta^t = \cup_{i,j,h:i\neq j\neq h\neq i} \Delta^t_{ijh}. \tag{12}$$

Remark that the right-hand side of (3) is discontinuous on a subset of Δ^t. Also in this case, $\mathbb{R}^{nN}$ is the disjoint union of a finite number of sets delimited by the Δ^t_{ijh}'s such that f^t restricted to each of them is continuous. □

The next examples show that the inclusions $\mathcal{C}l \subseteq \mathcal{C}a \subseteq \mathcal{K}$ are proper for both dynamics. It also shows that Caratheodory and Krasovsky solutions starting from the same initial condition may converge to different equilibria.

Example 4 (Proper Inclusions Between Solution Sets, Metric) Consider (2) with $N = 3, n = 1$, $a \equiv 1$ and the initial condition $\overline{x} = (-\frac{1}{3}, 0, 1)$. Note that $\overline{x}$ is a discontinuity point of $f^m(x)$ as $x_3 - x_2 = 1$. We have $f^m(\overline{x}) = (\frac{1}{3}, -\frac{1}{3}, 0)$. In fact agent 2 and 3 do not communicate. There exists a unique classical solution $x_{\mathcal{C}}(t) = (-\frac{1}{6} - \frac{1}{6}e^{-2t}, -\frac{1}{6} + \frac{1}{6}e^{-2t}, 1)$ which converges to the point $(-\frac{1}{6}, -\frac{1}{6}, 1)$.

If we consider Caratheodory solutions, we note that there exists one more solution, that starts following the limit value of the vector field as $x_3 - x_2 \to 1^-$, namely $f^{m-}(\overline{x}) = (\frac{1}{3}, \frac{2}{3}, -1)$: this Caratheodory solution behaves as if agents 2 and 3 communicate. Its expression is $x_{\mathcal{C}a}(t) = (\frac{1}{9}e^{-3t} - \frac{2}{3}e^{-t} + \frac{2}{9}, -\frac{2}{9}e^{-3t} + \frac{2}{9}, \frac{1}{9}e^{-3t} + \frac{2}{3}e^{-t} + \frac{2}{9})$ and it converges to $(\frac{2}{9}, \frac{2}{9}, \frac{2}{9})$.

We finally consider Krasovsky solutions. Besides the ones already obtained there exists a solution that slides on the discontinuity plane $\pi : x_3 - x_2 = 1$. In fact admissible directions $\tilde{f}^m$ at the points of π belong to the set

$$Kf(x) = \{\alpha(x_2 - x_1, 1 + x_1 - x_2, -1) + (1 - \alpha)(x_2 - x_1, x_1 - x_2, 0) \; : \; \alpha \in [0, 1]\}.$$

Since the normal vector to π is $v_\perp = (0, -1, 1)$, we have that $v_\perp \cdot \dot{x} = -2\alpha + x_2 - x_1$ is equal to zero if $\alpha = \frac{1}{2}(x_2 - x_1)$. Namely, the Krasovsky solution corresponding to this α does not exit the discontinuity plane but slides on it. In fact the sliding solution keeps x_3 and x_2 at distance 1 as $\dot{x}_3 - \dot{x}_2 = 0$. The solution can stay on the discontinuity for arbitrarily long time: if it remains there forever, then it converges to the point $(-\frac{1}{9}, -\frac{1}{9}, \frac{8}{9})$. Other Krasovsky solutions may exit π at arbitrary times T in two different ways: either agents 2 and 3 influence each other and the solution converges to $(\frac{2}{9}, \frac{2}{9}, \frac{2}{9})$, or they stop interacting at all and the solution converges to $(x^*, x^*, x_3(T))$ with $x^* = \frac{1}{3} - \frac{x_3(T)}{2}$.

Example 5 (Proper Inclusions Between Solution Sets, Topological) We consider (3) with $N = 3$, $n = 1$, $\kappa = 1$, $a \equiv 1$ and the initial condition $\overline{x} = (0, -1, 1)$. We remark that the vector field defined by the right-hand side of (3) is discontinuous at $\overline{x}$ as it belongs to the plane $\pi : (x_1 - x_2) - (x_3 - x_2) = 0$. In order to have the equation satisfied at $t = 0$ classical solutions must satisfy the equations

$$\begin{cases} \dot{x}_1 = x_2 - x_1 \\ \dot{x}_2 = x_1 - x_2 \\ \dot{x}_3 = x_1 - x_3. \end{cases}$$

There is a unique classical solution starting from $\overline{x}$ and it converges to the point $\left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$. In fact $\dot{x}_1 + \dot{x}_2 = 0$, then $x_1(t) + x_2(t) = \bar{x}_1 + \bar{x}_2 = \lim_{t \to +\infty}[x_1(t) + x_2(t)] = -1$.

We now observe that there is a Caratheodory solution that does not satisfy the equations at $t = 0$ but does for $t > 0$, namely the solution of the equations

$$\begin{cases} \dot{x}_1 = x_3 - x_1 \\ \dot{x}_2 = x_1 - x_2 \\ \dot{x}_3 = x_1 - x_3. \end{cases}$$

This Caratheodory solution converges to the equilibrium point $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$.

Finally we have a Krasovsky solution starting at $\overline{x}$ that slides on the plane π. In fact, if we denote by $f^{t-}(x)$ and $f^{t+}(x)$ its limit values as x approaches the plane π from the negative and positive sides, respectively, we have $f^{t-}(x) = (x_2 - x_1, x_1 - x_2, x_1 - x_3)$ and $f^{t+}(x) = (x_3 - x_1, x_1 - x_2, x_1 - x_3)$. We can then compute the Krasovsky set-valued map on π:

$$Kf(x) = \{(\alpha x_2 + (1-\alpha)x_3 - x_1, x_1 - x_2, x_1 - x_3), \alpha \in [0, 1]\}.$$

By posing $\alpha = \frac{1}{2}$ we obtain the admissible direction $f^t_{1/2}(x) = (0, x_1 - x_2, x_1 - x_3)$ which is parallel to π. This implies that there is a Krasovsky solution starting from $\overline{x}$ and sliding on π, namely $x_1(t) = 0$, $x_2(t) = -e^{-t}$, $x_3(t) = e^{-t}$, which converges to the origin.

2.3 P1) Average Preservation

In this section, we discuss property P1) that is preservation of the average value of the agents. We prove that P1) is satisfied for metric interaction models, while this is not the case for topological interactions. This is one more consequence of the lack of symmetry of topological interactions.

Proposition 5 (Average Preservation) *Caratheodory and Krasovsky solution of* (2) *have property P1).*

The proof of Proposition 5 can be found in [15] in the case $n = 1$ and in [38] in the general case. The same property does not hold for solutions of the topological bounded confidence model, by the following example.

Example 6 (Example 5, Continued) Let $x(t)$ be the unique classical solution starting from the point $(0, -1, 1)$. Observe that $x(t)$ is such that $x_{ave}(0) = 0$, but the limit of $x(t)$ for $t \to +\infty$ is $(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$, so that $\lim_{t \to +\infty} x_{ave}(t) = -\frac{1}{2}$.

2.4 P2) Contractivity of the Support

In this section, we prove that the support of solutions (in any sense given above) is weakly contractive. This is a well-known property of Caratheodory solutions for bounded confidence models, see e.g. [11]. The proof of such property for Krasovsky solutions for the metric model (2) on the real line can be found in [15, Prop. 3.iii].

We will give a general proof for Krasovsky solutions in any dimension, both for the metric (2) and topological models (3). The proof is similar to the one provided in [38], thus we provide a sketch only.

Proposition 6 (Contractivity of the Support) *Let* $x(t) = (x_1(t), x_2(t), \ldots, x_N(t))$ *be a solution to either* (2) *or* (3)*, in any of the senses given in Definition* 2*. Assume* $a : [0, +\infty[\to [0, +\infty[$ *continuous and* $0 \le T^1 < T^2$*, then*

$$\overline{co}\left(\left\{x_1(T^1), x_2(T^1), \ldots, x_N(T^1)\right\}\right) \supseteq \overline{co}\left(\left\{x_1(T^2), x_2(T^2), \ldots, x_N(T^2)\right\}\right). \tag{13}$$

Proof Let $x(\cdot)$ be a Krasovsky solution. Define $X(t) := \overline{co}\,(\{x_1(t), x_2(t), \ldots, x_N(t)\})$ and

$$A(T^1) := \left\{T \in (T^1, +\infty) \text{ s.t. } X(T^1) \not\supseteq X(T)\right\}.$$

We claim that $A(T^1)$ is empty, which implies (13). Otherwise, by contradiction, we can define $T^3 = \inf A(T^1) \ge T^1$. Following the same argument as in [38], we can prove the following:

Claim a) It holds either $\inf(A(T^1)) = T^1$ or $\inf(A(T^3)) = T^3$.

Without loss of generality, we can assume $T^1 = 0$ or $T^3 = 0$, thus $\inf(A(0)) = 0$. Let $t_k \searrow 0$ be such that $x_i(t_k) \notin X(0)$ for a fixed $i \in V$, then by continuity of the trajectory it holds $\bar{x}_i := x_i(0) \in \partial X(0)$, where ∂ indicated the topological boundary. Since $X(0)$ is a convex polyhedron, it is supported by a finite number of hyperplanes at $\bar{x}_i$, thus, by possibly passing to subsequences, we can find a unitary vector ν such that $(x_i(t_k) - \bar{x}_i) \cdot \nu > 0$ for all t_k. Moreover, it holds $(x_j(0) - \bar{x}_i) \cdot \nu \le 0$ for all $j \in V$.

Now, define $\phi_j(x) = (x_j - \bar{x}_i) \cdot \nu$, and $\phi(x) := \max_{j \in V} \phi_j(x)$. Observe that $\phi(x(0)) \le 0$ and $\phi(x(t_k)) > 0$. We can apply Danskin Theorem [24] to ϕ, thus even though ϕ may be not differentiable, it admits directional derivatives. Denote by $h = (h_1, \ldots, h_N)$ the displacement, then by applying Danskin formula, the directional derivative D_h along h is given by

$$D_h\phi(x) = \max_{j \in A_i(x)} \sum_{k=1}^{N} h_k \cdot \nabla_{x_k}\phi_j(x) = \max_{j \in A_i(x)} h_j \cdot \nabla_{x_j}\phi_j(x) = \max_{j \in A_i(x)} h_j \cdot \nu,$$

where $A_i(x)$ is the set of indexes $j \neq i$ realizing the maximum in the definition of $\phi(x)$. Since $D_h\phi(x)$ is always defined and $\dot{x}(t)$ exists for almost every $t \in (0, T)$, then also $\dot{\phi}(x(t))$ exists for almost every $t \in (0, T)$. Moreover, by direct computation, we get:

$$\begin{aligned}\lim_{\tau\to 0} \frac{\phi(x(t+\tau)) - \phi(x(t))}{\tau} &= \lim_{\tau\to 0} \frac{\phi(x(t) + \tau\dot{x}(t) + o(\tau)) - \phi(x(t))}{\tau} \\ &= \lim_{\tau\to 0} \frac{\phi(x(t)) + \tau D_{\dot{x}(t)}\phi(x(t)) + o(\tau) - \phi(x(t))}{\tau} \\ &= D_{\dot{x}(t)}\phi(x(t)),\end{aligned} \tag{14}$$

thus it holds:

$$\dot{\phi}(x(t)) = \max_{j\in A_i(x(t))} \dot{\phi}_j(t) = \max_{j\in A_i(x(t))} \dot{x}_j(t) \cdot \nu. \tag{15}$$

Now, if $x(\cdot)$ is differentiable at t, and $j \in A_i(x(t))$, then for every $k \neq j$ we have

$$\begin{aligned}(x_k(t) - x_j(t)) \cdot \nu &= (x_k(t) - \bar{x}_i) \cdot \nu + (\bar{x}_i - x_j(t)) \cdot \nu \\ &= \phi_k(x(t)) - \phi_j(x(t)) \le \phi(t) - \phi(t) = 0.\end{aligned} \tag{16}$$

Since $x(\cdot)$ is a Krasovsky solution, there exist $b_{jk} \geq 0$ such that $\dot{x}_j = \sum_{k=1}^{N} b_{jk} a(\|x_k - x_j\|)(x_k - x_j)$. Substituting this expression in (15), we get:

$$\dot{\phi}(x(t)) = \max_{j \text{ s.t. } \phi(t)=\phi_j(t)} \sum_{k=1}^{N} b_{jk} a(\|x_k - x_j\|)(x_k - x_j) \cdot \nu \le 0.$$

This contradicts the fact that $\phi(x(0)) = 0$ and $\phi(t_k) > 0$. Thus (13) holds, for the Krasovsky solution $x(\cdot)$. Since the proof holds for every Krasovsky solution, by recalling the inclusions of Sect. 2.2, the statement holds for any definition of solution. □

3 Existence and Uniqueness of Solutions

In this section, we study existence and uniqueness of solutions, both for the metric and the topological models.

3.1 Existence of Solutions

Proposition 7 (Existence of Krasovsky Solutions) *For any initial condition, equations* (2) *and* (3) *admit a Krasovsky solution defined on* $[0, +\infty)$.

Proof For both (2) and (3) the right-hand side is locally bounded. The local existence of Filippov solutions then follows from Proposition 2. By Proposition 4, the sets of Krasovsky and Filippov solutions coincide, then local Krasovsky solutions also exist. Proposition 6 guarantees that solutions are bounded, then they can be continued on $[0, +\infty)$ by standard arguments. □

In general, Krasovsky solutions are not unique, as already shown in Example 5. In the following proposition we state the existence of Caratheodory solutions for both metric and topological bounded confidence models. The proof for the metric model in the case $n = 1$ was first given in [11] and then generalized to any n in [38]. The proof for the topological case with $\kappa = 1$ is new. We conjecture that the result holds for $\kappa > 1$ as well, although with a more involved argument that we avoid to develop here.

Proposition 8 (Existence of Caratheodory Solutions)

(i) ***Metric bounded confidence.*** *For any initial condition, equation* (2) *admits a Caratheodory solution defined on* $[0, +\infty)$.
(ii) ***Topological bounded confidence.*** *If* $\kappa = 1$*, then any initial condition* (3) *admits a Caratheodory solution defined on* $[0, +\infty)$.

Proof We only consider the topological bounded confidence model with $\kappa = 1$. We build a Caratheodory solution as follows. For each initial datum $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_N)$, we construct an oriented graph G for which there exists $T > 0$ and a curve defined on $[0, T]$ having G as connectivity graph. For each index $i \in V$ there exists one and only one index that we denote with $\Gamma(i)$, such that $(i, \Gamma(i)) \in G$. This implies that $\dot{x}_i = a(\|x_{\Gamma(i)} - x_i\|)(x_{\Gamma(i)} - x_i)$ for the whole time interval $[0, T]$. We then need to prove that the corresponding trajectory $(x_1(t), \ldots, x_N(t))$ is indeed a Caratheodory solution for (3). Remark that one might aim to choose $\Gamma(i)$ as the single element of $N_i^t(\bar{x})$, that is the minimal index (in the lexicographic order) among nearest neighbors of $\bar{x}_i$. In our proof, this is not the case, as one might choose $\Gamma(i)$ not to be the nearest neighbor with minimal index at the initial time, but to be the nearest neighbor for all $t \in (0, T)$. We then conclude the proof by piecing together Caratheodory solutions on time intervals $[0, T_1]$, $[T_1, T_2]$, ... to build a solution on $[0, +\infty)$.

Let $\bar{x}$ be an initial configuration, that is fixed from now on. We build an oriented graph G recursively by selecting, for each index $i \in V$, a unique index $\Gamma(i)$ such that $(i, \Gamma(i)) \in G$. First define:

$$A_i := \operatorname{argmin}_{j \neq i}(\|\bar{x}_j - \bar{x}_i\|),$$

which is the set of agents realizing the minimal distance to $\bar{x}_i$.

The graph G is constructed using the following algorithm:

Step 1) For all i such that $N_i^t = \{j\}$, define $\Gamma(i) := j$.
Step 2) WHILE there exists a pair of indexes i, j such that $i \notin G$, $j \notin G$ and $j \in A_i, i \in A_j$
DO: define $\Gamma(i) := j$ and $\Gamma(j) := i$.
Step 3) IF there exists $i \notin G$ such that $A_i = \{j_1, \ldots, j_l\}$ and $j_1, \ldots, j_l \in G$
DO: choose

$$\Gamma(i) \in \text{argmin}_{j \in \{j_1, \ldots, j_l\}} \psi_i(j) \tag{17}$$

where

$$\psi_i(l) := (x_l - x_i) \cdot \big(a(\|x_{\Gamma(l)} - x_l\|)(x_{\Gamma(l)} - x_l) - a(\|x_l - x_i\|)(x_l - x_i)\big) \tag{18}$$

Step 4) IF for all $i \in V$ it holds $i \in G$, STOP.
ELSE: Go to Step 3.

Observe that the number of edges of G is increased at each step and is bounded by N. Thus, there exists a limit graph G', reached after a finite number of steps. We now prove the following claim:

$$(\mathbf{C}) \qquad \text{for all } i \in V \quad \text{it holds} \quad i \in G'.$$

To prove **(C)**, assume by contradiction that there exists $i \notin G'$ and, by possibly relabeling indexes, $i = 1 \notin G'$. By definition of G', Step 3 does not add edges to G', in particular no edge to agent $i = 1$; thus A_1 contains at least one index that we relabel as 2, such that $2 \notin G'$. If $1 \notin A_2$, we can find another index, relabeled as 3, such that $3 \notin G'$ and so on. Finally there exists $k \notin G'$ such that $A_k \cap \{1, 2, \ldots, k-1\} \neq \emptyset$. Possibly reducing the sequence and changing the initial element, we assume $1 \in A_k$.

Now, if there exist $i, i+1 \in \{1, 2, \ldots, k\}$ such that $i+1 \in A_i$ and $i \in A_{i+1}$, then we are in contradiction with Step 2. Therefore, we can assume that for all $i \in \{1, 2, \ldots, k-1\}$ it holds $i \notin A_{i+1}$. Since $i+2 \in A_{i+1}$ by construction, we have

$$\|x_i - x_{i+1}\| > \|x_{i+1} - x_{i+2}\| \quad \text{for all } i \in \{1, 2, \ldots, k-2\}. \tag{19}$$

Using (19) and recalling $1 \in A_k$ we get

$$\|x_1 - x_2\| > \|x_2 - x_3\| > \ldots > \|x_{k-1} - x_k\| > \|x_k - x_1\|.$$

This implies $2 \notin A_1$, achieving a contradiction. This concludes the proof of claim **(C)**.

Let $x(\cdot)$ be the curve satisfying $\dot{x}_i = a(\|x_{\Gamma(i)} - x_i\|)(x_{\Gamma(i)} - x_i)$, i.e. with dynamics associated with G, with initial condition $\bar{x}$. We first show that there exists a time $T > 0$ such that $\Gamma(i) \in \text{argmin}_{j \neq i}(\|x_j(t) - x_i(t)\|)$ for all $t \in [0, T]$.

More precisely, for each $i \in V$, and $k \in V \setminus \{i, \Gamma(i)\}$ we show that there exists $T_{ik} > 0$ such that $\|x_i(t) - x_k(t)\| \geq \|x_i(t) - x_{\Gamma(i)}(t)\|$ on $[0, T_{ik}]$. Then it will be sufficient to define $T = \min_{ik} T_{ik}$ (with the convention that the minimum is $+\infty$ if all $T_{ik} = +\infty$).

Now, fix $i \in V$, and $k \in V \setminus \{i, \Gamma(i)\}$. Notice that if $\|\bar{x}_i - \bar{x}_k\| > \|\bar{x}_i - \bar{x}_{\Gamma(i)}\|$, then by continuity there exists $T_{ik} > 0$ such that $\|x_i(t)-x_k(t)\| > \|x_i(t)-x_{\Gamma(i)}(t)\|$ for all $t \in [0, T_{ik}]$. Therefore, from now on, we assume

$$\|\bar{x}_i - \bar{x}_k\| \leq \|\bar{x}_i - \bar{x}_{\Gamma(i)}\|.$$

By definition of A_i, this inequality is indeed an equality, otherwise $\Gamma(i) \notin A_i$. We distinguish two sub-cases:

Case 1) $\bar{x}_k = \bar{x}_{\Gamma(i)}$.
Case 2) $\bar{x}_k \neq \bar{x}_{\Gamma(i)}$.

In **Case 1)** there exists a (possibly empty) set of indexes $L := \{l_1, \ldots, l_r\}$ such that $\bar{x}_k = \bar{x}_{\Gamma(i)} = \bar{x}_{l_m}$, hence $A_k = L \cup \{\Gamma(i)\}$. From claim **(C)**, for each $l \in \{k, \Gamma(i)\} \cup L$ the neighbor $\Gamma(l)$ is well-defined, thus $\bar{x}_l = \bar{x}_{\Gamma(l)}$ by the condition of minimal distance. This in turn implies $\dot{x}_l \equiv 0$, and similarly for all other indexes in A_k. Since all indexes $l \in \{k, \Gamma(i)\} \cup L$ satisfy such property, it holds $A_l \subseteq \operatorname{argmin}_j(\|x_j(t) - x_l(t)\|)$. Observe moreover that the dynamics does not allow for merging particles, thus the inclusion is indeed an equality. This in turn means that $x_k(t) = x_{\Gamma(i)}(t)$ for all $t > 0$, hence we can choose $T_{ik} = +\infty$.

Consider now **Case 2)**. Observe that $\bar{x}_k \neq \bar{x}_i$, otherwise $\|\bar{x}_{\Gamma(i)} - \bar{x}_i\| \leq \|\bar{x}_k - \bar{x}_i\| = 0$, which gives $\bar{x}_{\Gamma(i)} = \bar{x}_i = \bar{x}_k$. We now prove that there exists $T_{ik} > 0$ such that

$$\|x_i(t) - x_{\Gamma(i)}(t)\| < \|x_i(t) - x_k(t)\| \quad \text{for all } t \in (0, T_{ik}]. \tag{20}$$

Consider the function

$$\phi_{ik}(t) := \tfrac{1}{2}\|x_{\Gamma(i)}(t) - x_i(t)\|^2 - \tfrac{1}{2}\|x_k(t) - x_i(t)\|^2.$$

Since $\phi_{ik}(0) = 0$, to prove (20) it is enough to show $\phi'_{ik}(0) < 0$. Set $j = \Gamma(i)$, then from (18), we get

$$\phi'_{ik}(0) = (\bar{x}_j - \bar{x}_i) \cdot (\dot{x}_j(0) - \dot{x}_i(0)) - (\bar{x}_k - \bar{x}_i) \cdot (\dot{x}_k(0) - \dot{x}_i(0)) = A_{ijk} - B_{ijk}$$

with

$$\begin{aligned} A_{ijk} &= \psi_i(j) - \psi_i(k) \\ &= (\bar{x}_j - \bar{x}_i) \cdot (a(\|\bar{x}_{\Gamma(j)} - \bar{x}_j\|)(\bar{x}_{\Gamma(j)} - \bar{x}_j) - (\bar{x}_j - \bar{x}_i) \cdot a(\|\bar{x}_j - \bar{x}_i\|)(\bar{x}_j - \bar{x}_i) \\ &\quad -(\bar{x}_k - \bar{x}_i) \cdot a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)(\bar{x}_{\Gamma(k)} - \bar{x}_k) + (\bar{x}_k - \bar{x}_i) \cdot (a(\|\bar{x}_k - \bar{x}_i\|)(\bar{x}_k - \bar{x}_i)), \\ B_{ijk} &= (\bar{x}_k - \bar{x}_i) \cdot (a(\|\bar{x}_k - \bar{x}_i\|)(\bar{x}_k - \bar{x}_i)) - (\bar{x}_k - \bar{x}_i) \cdot (a(\|\bar{x}_j - \bar{x}_i\|)(\bar{x}_j - \bar{x}_i))). \end{aligned}$$

The idea of the decomposition is that the last term in A_{ijk} would correspond to $(\bar{x}_k - \bar{x}_i)\dot{x}_i(0)$ by choosing k as the neighbor of i. Thus, B_{ijk} is the corrector given by the actual choice $j = \Gamma(i)$. We now show $A_{ijk} \leq 0$ and $B_{ijk} > 0$, which implies $\phi'_{ik}(0) < 0$.

Consider first A_{ijk}. The index $j = \Gamma(i)$ was not chosen in Step 1 since $k \neq j$ and $k, j \in A_i$. If $j = \Gamma(i)$ was chosen in Step 2, then $\Gamma(j) = i$, and

$$A_{ijk} \leq -2\|\bar{x}_j - \bar{x}_i\|^2 a(\|\bar{x}_j - \bar{x}_i\|) + \|\bar{x}_k - \bar{x}_i\| \cdot \|\bar{x}_{\Gamma(k)} - \bar{x}_k\| a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)$$
$$+\|\bar{x}_k - \bar{x}_i\| \cdot \|\bar{x}_k - \bar{x}_i\| a(\|\bar{x}_k - \bar{x}_i\|).$$

By definition of $\Gamma(k)$ we have $\|\bar{x}_{\Gamma(k)} - \bar{x}_k\| \leq \|\bar{x}_i - \bar{x}_k\|$, and, since we are in **Case 2**, it holds $\|\bar{x}_j - \bar{x}_i\| = \|\bar{x}_k - \bar{x}_i\|$. Recalling that $a(r)$ is non-decreasing, we get $A_{ijk} \leq 0$. Assume now that $j = \Gamma(i)$ was chosen in Step 3, then, by construction $j \in \text{argmin}_\ell \psi_i(\ell)$, thus $0 \geq \psi_i(j) - \psi_i(k) = A_{ijk}$.

We now prove $B_{ijk} > 0$. Observe that $j, k \in A_i$, hence both $\bar{x}_j$ and $\bar{x}_k$ lay on the same circle centered at $\bar{x}_i$ on the plane containing $\bar{x}_i, \bar{x}_j$ and $\bar{x}_k$. Thus $a(\|\bar{x}_k - \bar{x}_i\|) = a(\|\bar{x}_j - \bar{x}_i\|)$. Since $\bar{x}_k \neq \bar{x}_j$ (**Case 2**) and $\bar{x}_k \neq \bar{x}_i$, the amplitude α of the angle $\widehat{\bar{x}_j \bar{x}_k \bar{x}_i}$, on the plane containing $\bar{x}_i, \bar{x}_j$ and $\bar{x}_k$, belongs to $(-\pi/2, \pi/2)$, thus

$$B_{ijk} = (\bar{x}_k - \bar{x}_i) \cdot a(\|\bar{x}_j - \bar{x}_i\|)(\bar{x}_k - \bar{x}_j)$$
$$= a(\|\bar{x}_j - \bar{x}_i\|)\|\bar{x}_k - \bar{x}_i\| \, \|\bar{x}_k - \bar{x}_j\| \cos(\alpha) > 0.$$

From $A_{ijk} \leq 0$ and $B_{ijk} > 0$ we get $\phi'_{ik}(0) < 0$ and we are done.

We now show that $x(\cdot)$ is a Caratheodory solution of (3). If $\Gamma(i) = N_i^t(x(t))$ for all times $t \in (0, T)$, then we are done. Otherwise, there exists $t \in (0, T)$ such that $\Gamma(i) \neq k = N_i^t(x(t))$, thus $\|x_{\Gamma(i)}(t) - x_i(t)\| \geq \|x_k(t) - x_i(t)\|$. Recalling the definition of T_{ik}, we deduce that **Case 1)** holds, thus $\Gamma(i) \neq k = N_i^t(x(t))$ and $x_{\Gamma(i)}(t) = x_k(t)$, i.e. the indexes k and $\Gamma(i)$ are different but the agents' positions coincide. As a consequence, we have

$$\dot{x}_i = a(\|x_{\Gamma(i)} - x_i\|)(x_{\Gamma(i)} - x_i) = a(\|x_k - x_i\|)(x_k - x_i),$$

and (3) is satisfied.

We now prove that the trajectory can be prolonged to $[0, +\infty)$. If $T = +\infty$, then we are done. Otherwise, observe that the trajectory $x(\cdot)$ is compact, due to contractivity of the support proved in Proposition 6, thus we can use transfinite induction as follows. Since $\dot{x}(\cdot)$ is uniformly bounded, $x(\cdot)$ is a uniformly Lipschitz function of time with Lipschitz constant $L := \max_{ij} a(\|\bar{x}_i - \bar{x}_j\|)$, and $x(T)$ is well-defined. We can apply the same algorithm at time T, and find $T_1 > 0$ such that the trajectory is well-defined on $[T, T_1]$. If $T_1 = +\infty$, we are done; otherwise define $T_2 < T_3 < \ldots$ in the same way and extend the trajectory to $[T_i, T_{i+1}]$. If $T_i = +\infty$

for some i or $\lim_{i\to+\infty} T_i = +\infty$, then we are done. Assume, by contradiction that $T^* = \lim_{i\to+\infty} T_i < +\infty$. Then $x(T^*)$ is well-defined and using the algorithm we can extend the trajectory beyond T^*.

Using for T_i the same argument as for T, we have that $x(\cdot)$ is a Caratheodory solution to (3) on each interval (T_l, T_{l+1}). Since $\{T_i\}$ is a countable set, we are done. □

3.2 Uniqueness for Almost Every Initial Condition

In this section, we study uniqueness of solutions. Examples 4 and 5 show that Caratheodory solutions are not unique in general (thus neither Krasovsky). Nevertheless, uniqueness of Krasovsky (and then also Caratheodory) solutions holds for almost all initial condition, both for metric and topological models. For the metric case, the result was already given in [38, Prop. 6.2].

We then focus on uniqueness of Krasovsky solutions for almost every initial datum for (3). We first set

$$I = \{(i, j, k) : i \neq j, i \neq k, j \neq k\}$$

and define

$$\mathcal{M} = \cup_{ijk\in I}\mathcal{M}_{ijk}, \quad \mathcal{M}_{ijk} = \{x : \|x_i - x_j\| = \|x_j - x_k\|\}.$$

Notice that $\mathcal{M}$ contains (in general strictly contains) the set where the right-hand side of (3) is discontinuous.

The main reason for uniqueness is that Krasovsky solution cannot enter the manifolds $\mathcal{M}$ and slide on it, except possibly on a set of codimension two. We first show this fact for the case $\kappa = 1$ for simplicity.

Given $(i, j, k) \in I$, consider the functions

$$\theta_{ijk}(x) = \|x_j - x_i\|^2 - \|x_k - x_i\|^2 \tag{21}$$

and denote by π_{ijk} the subset of the manifold $\mathcal{M}_{ijk}$ where the right-hand side f^t of (3) is discontinuous and where $\theta_{jvw}(x)\theta_{khu}(x)$ is different from zero for all v, w, h, u (that is, where the only discontinuity is due to j or k). We want to prove that π_{ijk} cannot be attractive with respect to Krasovsky solutions. Fix $\bar{x}$ a discontinuity point for f^t, thus $\bar{x} \in \pi_{ijk}$ for some $(i, j, k) \in I$ and either $j \in N_i^t(\bar{x})$ or $k \in N_i^t(\bar{x})$. We denote by $f^{t+}(\bar{x})$ and $f^{t-}(\bar{x})$ the limit values of $f^t(x)$ as $x \to \bar{x}$ and $\theta_{ijk}(x) > 0$ (the neighbor of i is then k) and $\theta_{ijk}(x) < 0$ (the neighbor of i is then j), respectively. We denote by $\Gamma(j)$ the neighbor of j at $\bar{x}$ and by $\Gamma(k)$ the neighbor of k at $\bar{x}$. We first denote by γ the angle between the vectors $\bar{x}_j - \bar{x}_i$ and $\bar{x}_k - \bar{x}_i$ and $l = \|\bar{x}_j - \bar{x}_i\| = \|\bar{x}_j - \bar{x}_i\|$. If $l = 0$, the angle is not uniquely defined, but this plays no role in the following. Let us compute the two quantities $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x})$ and $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x})$. We have:

$$
\begin{aligned}
&\nabla\theta_{ijk}(\bar{x}) \cdot f^{+t}(\bar{x}) \\
&= (\bar{x}_j - \bar{x}_i) \cdot [a(\|\bar{x}_{\Gamma(j)} - \bar{x}_j\|)(\bar{x}_{\Gamma(j)} - \bar{x}_j) - a(\|\bar{x}_k - \bar{x}_i\|)(\bar{x}_k - \bar{x}_i)] \\
&-(\bar{x}_k - \bar{x}_i) \cdot [a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)(\bar{x}_{\Gamma(k)} - \bar{x}_k) - a(\|\bar{x}_k - \bar{x}_i\|)(\bar{x}_k - \bar{x}_i)] \\
&= a(\|\bar{x}_{\Gamma(j)} - \bar{x}_j\|)(\bar{x}_j - \bar{x}_i) \cdot (\bar{x}_{\Gamma(j)} - \bar{x}_j) - a(l)l^2\cos(\gamma) \\
&-a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)(\bar{x}_k - \bar{x}_i) \cdot (\bar{x}_{\Gamma(k)} - \bar{x}_k) + a(l)l^2
\end{aligned}
$$

and

$$
\begin{aligned}
\nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) &= (\bar{x}_j - \bar{x}_i) \cdot [a(\|\bar{x}_{\Gamma(j)} - \bar{x}_j\|)(\bar{x}_{\Gamma(j)} - \bar{x}_j) - a(\|\bar{x}_j - \bar{x}_i\|)(\bar{x}_j - \bar{x}_i)] \\
&-(\bar{x}_k - \bar{x}_i) \cdot [a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)(\bar{x}_{\Gamma(k)} - \bar{x}_k) - a(\|\bar{x}_j - \bar{x}_i\|)(\bar{x}_j - \bar{x}_i)] \\
&= a(\|\bar{x}_{\Gamma(j)} - \bar{x}_j\|)(\bar{x}_j - \bar{x}_i) \cdot (\bar{x}_{\Gamma(j)} - \bar{x}_j) - a(l)l^2 \\
&-a(\|\bar{x}_{\Gamma(k)} - \bar{x}_k\|)(\bar{x}_k - \bar{x}_i) \cdot (\bar{x}_{\Gamma(k)} - \bar{x}_k) + a(l)l^2\cos(\gamma).
\end{aligned}
$$

We have:

$$
\begin{aligned}
\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) - \nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) &= -a(l)l^2\cos(\gamma) \\
&\quad +a(l)l^2 + a(l)l^2 - a(l)l^2\cos(\gamma) \\
&= 2a(l)l^2(1 - \cos(\gamma)).
\end{aligned}
$$

First of all, we remark that $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) - \nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) \geq 0$. Moreover $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) - \nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) = 0$ if and only if $\gamma = 0$, so that $f^{t+}(\bar{x})$ and $f^{t-}(\bar{x})$ are parallel. In this case, π_{ijk} is crossed by Krasovsky solutions, unless $f^{t+}(\bar{x})$ is tangent to the manifold, but this may occur only on a set of codimension at least two. Let us then consider the case $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) - \nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) > 0$ and analyze different possibilities.

- Case $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) > 0$. If also $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) > 0$, then Krasovsky solutions cross π_{ijk}. If $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) \leq 0$, then Krasovsky solutions can either leave π_{ijk} or slide on it.
- Case $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t+}(\bar{x}) \leq 0$. Then $\nabla\theta_{ijk}(\bar{x}) \cdot f^{t-}(\bar{x}) < 0$ and Krasovsky solutions cross π_{ijk}.

We conclude that solutions, which originate from outside π_{ijk} and reach it, must cross it. Therefore, uniqueness can only fail for (sliding) solutions that originate inside π_{ijk}. After this informal argument for $\kappa = 1$, we proceed to give a complete proof for any κ, thereby completing the proof of Theorem 1.

Proposition 9 (Uniqueness from Almost Any Initial Datum, Topological) *The set of initial data from which there exist more than one Krasovsky solutions for* (3) *has zero Lebesgue measure in* $\mathbb{R}^{nN}$.

Proof Fix an initial condition $\bar{x}$ and let $X_{\bar{x}}$ be the set of solutions $x(\cdot)$ to (3) such that $x(0) = \bar{x}$ defined on $[0, T(x(\cdot))[$, with $0 < T(x(\cdot)) \leq +\infty$. Define

$$t_U = \inf\{t : \exists x(\cdot), y(\cdot) \in X_{\bar{x}}, t \leq \min\{T(x(\cdot)), T(y(\cdot))\}, x(t) \neq y(t)\}, \tag{22}$$

and

$$\mathcal{A} = \{\bar{x} \in \mathbb{R}^{nN} \setminus \mathcal{M} : t_U < +\infty\}. \tag{23}$$

Notice that $\mathcal{M}$ is a stratified set of codimension 1, thus of zero Lebesgue measure in $\mathbb{R}^{nN}$. Therefore the statement is equivalent to prove that $\mathcal{A}$ has zero Lebesgue measure. For $\bar{x} \in \mathcal{A}$, we define:

$$\tilde{t} = \inf\{t : \exists x(\cdot) \in X_{\bar{x}}, x(t) \in \mathcal{M}\}. \tag{24}$$

Since (3) is Lipschitz continuous on $\mathbb{R}^{nN} \setminus \mathcal{M}$, there exists a unique solution in $X_{\bar{x}}$ at least until reaching $\mathcal{M}$, thus $\tilde{x} = x(\tilde{t}) \in \mathcal{M}$ depends only on $\bar{x}$. Now define the set of indexes

$$J = \{(i, j, k, i', j', k') : (i, j, k) \neq (i', j', k')\},$$

and the stratified sets:

$$\mathcal{M}_{ijki'j'k'} = \mathcal{M}_{ijk} \cap \mathcal{M}_{i'j'k'}.$$

We now analyze the dynamics on $\mathcal{M} \setminus \left(\cup_{(i,j,k,i',j',k') \in J} \mathcal{M}_{ijki'j'k'}\right)$ to identify a stratified set of codimension two out of which trajectories cross $\mathcal{M}$ transversally. Consider now $x \in \mathcal{M}_{ijk}$, assume (i, j, k) is the unique index for which $x \in \mathcal{M}_{ijk}$. We also assume $\|x_i - x_j\| = \max_{\ell \in N_i^t(x)} \|x_i - x_\ell\|$, i.e. j is among the farthest $\kappa \geq 1$ neighbors of i, otherwise N_i^t is constant in a neighbor of x and uniqueness holds. Since (i, j, k) is the unique index for which $x \in \mathcal{M}_{ijk}$, we indeed have that $\max_{\ell \in N_i^t(x)} \|x_i - x_\ell\|$ is achieved exactly for indexes j and k. Now, set $P_i = N_i^t(x) \setminus \{j, k\}$, $P_j = N_j^t(x)$, $P_k = N_k^t(x)$. Define the following:

$$f_m(x) = \sum_{\ell \in P_m} a(\|x_\ell - x_m\|)(x_\ell - x_m) \tag{25}$$

for $m = i, j, k$. Then a Krasovsky solution $y(\cdot)$ with $y(0) = x$, if differentiable at 0, satisfies:

$$\dot{y}_i(0) = f_i(x) + \alpha\, a(\|x_j - x_i\|)(x_j - x_i) + (1-\alpha)\, a(\|x_k - x_i\|)(x_k - x_i),$$

for some $\alpha \in [0, 1]$, and:

$$\dot{y}_j(0) = f_j(x), \quad \dot{y}_k(0) = f_k(x).$$

Recall the definition of the function θ_{ijk} given in (21). If θ_{ijk} computed along $y(\cdot)$ is differentiable at 0, then:

$$\begin{aligned}\dot{\theta}_{ijk}(0) = C(x) &+ 2\alpha\, a(\|x_i - x_j\|)\,(x_j - x_i)\cdot(x_k - x_j)\\ &+2(1-\alpha)\, a(\|x_i - x_j\|)\,(x_k - x_i)\cdot(x_k - x_j),\end{aligned} \tag{26}$$

where we used $\|x_i - x_j\|^2 = \|x_i - x_k\|^2$ and

$$C(x) = 2(f_i - f_j)\cdot(x_i - x_j) - 2(f_i - f_k)\cdot(x_i - x_k). \tag{27}$$

Define the stratified sets

$$\begin{aligned}\widehat{\mathcal{M}}_{ijk} = \{x \in \mathcal{M}_{ijk} : C(x) &+ 2\, a(\|x_i - x_j\|)\,(x_j - x_i)\cdot(x_k - x_j) = 0,\\ \text{or } \; C(x) &+ 2\, a(\|x_i - x_j\|)\,(x_k - x_i)\cdot(x_k - x_j) = 0\},\end{aligned}$$

and finally

$$\widehat{\mathcal{M}} = \left(\cup_{ijk}\widehat{\mathcal{M}}_{ijk}\right) \bigcup \left(\cup_{(i,j,k,i',j',k')\in J}\mathcal{M}_{ijki'j'k'}\right).$$

Notice that $\widehat{\mathcal{M}}$ is of codimension 2, and we state the following claim:

Claim a) If $\tilde{x} \in \mathcal{M}\backslash\widehat{\mathcal{M}}$, then there exists $\epsilon > 0$ such that $x(t) \notin \mathcal{M}$ for $t \in]\tilde{t}, \tilde{t}+\epsilon[$, and $x \equiv y$ on $[0, \tilde{t}+\epsilon[$ for every $x(\cdot), y(\cdot) \in X_{\tilde{x}}$.

To prove Claim a), let $(i, j, k) \in I$ be the unique triplet such that $\tilde{x} \in \mathcal{M}_{ijk}$. Assume $j \in N_i^t(\tilde{x})$ or $k \in N_i^t(\tilde{x})$, otherwise the claim is obvious. The function θ_{ijk} computed along $x(\cdot)$ satisfies $\theta_{ijk}(\tilde{t}) = 0$, is twice continuously differentiable on $[0, \tilde{t}[$ with bounded derivatives, thus we can define $\tilde{\xi} := \lim_{t\to\tilde{t}-} \dot{\theta}_{ijk}(t)$.

First assume $\tilde{\xi} > 0$: then, there exists $\delta > 0$ such that both $\theta_{ijk}(t) < 0$ and $j \in N_i^t(x(t))$ on $]\tilde{t}-\delta, \tilde{t}[$. Then, possibly restricting $\delta > 0$, on $]\tilde{t}-\delta, \tilde{t}[$ we have $\dot{\theta}_{ijk}(t) = C(x(t)) + 2\, a(\|x_i(t) - x_j(t)\|)\,(x_j(t) - x_i(t))\cdot(x_k(t) - x_j(t)) > 0$ and $\tilde{\xi} = C(\tilde{x}) + 2\, a(\|\tilde{x}_i - \tilde{x}_j\|)\,(\tilde{x}_j - \tilde{x}_i)\cdot(\tilde{x}_k - \tilde{x}_j) > 0$. Since $x(\cdot)$ is a Krasovsky solution, for almost every time $\dot{\theta}_{ijk}(t)$ can be computed as in (26) for some $\alpha(t) \in [0, 1]$. From $(x_k - x_i) = (x_k - x_j) + (x_j - x_i)$, we get $(x_k - x_i)\cdot(x_k - x_j) = (x_j - x_i)\cdot(x_k - x_j) + \|x_k - x_j\|^2$, thus $\dot{\theta}_{ijk}(t) > 0$ for t sufficiently close to $\tilde{t}$. This implies that there exists $\epsilon > 0$ such that $\theta_{ijk} > 0$, $j \notin N_i^t(x(t))$ and $k \in N_i^t(x(t))$ for $t \in]\tilde{t}, \tilde{t}+\epsilon[$. Since $\tilde{x} \notin \widehat{\mathcal{M}}$, by possibly reducing ϵ, it holds $x(t) \notin \widehat{\mathcal{M}}$. In particular, all sets $N_\ell^t(x(t))$ are constant for $t \in]\tilde{t}, \tilde{t} + \epsilon[$. Thus we conclude that Claim a) holds.

The case $\tilde{\xi} < 0$ can be treated similarly, while the case $\tilde{\xi} = 0$ is excluded since $\tilde{x} \notin \widehat{\mathcal{M}}_{ijk}$.

Now set:

$$t_{\widehat{\mathcal{M}}} = \inf\{t : \exists x(\cdot) \in X_{\bar{x}}, x(t) \in \widehat{\mathcal{M}}\}, \tag{28}$$

Claim a) ensures $t_{\widehat{\mathcal{M}}} \leq t_U$, as uniqueness can be lost only when crossing $\widehat{\mathcal{M}}$. Therefore, if $\bar{x} \in \mathcal{A}$, then every $x(\cdot) \in X_{\bar{x}}$ is Lipschitz continuous and coincides (at least) up to $t_{\widehat{\mathcal{M}}}$. This implies that $H^{1+\epsilon}(\{x(t) : t \in [0, t_{\widehat{\mathcal{M}}}], x(\cdot) \in X_{\bar{x}}\}) = 0$ for every $\epsilon > 0$, where H^r is the Hausdorff measure of dimension r in $\mathbb{R}^{nN}$. Since $\widehat{\mathcal{M}}$ is of codimension 2, by Fubini Theorem, for $0 < \epsilon < 1$ we have:

$$H^{nN}(\mathcal{A}) \leq \int_{\widehat{\mathcal{M}}} \left(H^{1+\epsilon}(\{x(t) : t \in [0, t_{\widehat{\mathcal{M}}}], x(\cdot) \in X_{\bar{x}}\}) \right) dH^{nN-2+\epsilon}(\bar{x}) = 0.$$

The measure H^{nN} coincides with the Lebesgue measure on $\mathbb{R}^{nN}$, thus $\mathcal{A}$ has zero Lebesgue measure.

□

4 Asymptotic Behavior of Solutions

We now study convergence to cluster points, i.e. Property P3). We first need to investigate the relationships between cluster points and equilibria for Caratheodory and Krasovsky solutions to (2) and (3).

4.1 Equilibria and Cluster Points

We begin by recalling that cluster points are points $x^\infty = (x_1^\infty, \ldots, x_N^\infty)$, $x_i^\infty \in \mathbb{R}^n$, such that for every $i \in V$, for every $j \in N_i(x^\infty)$ it holds $x_i^\infty = x_j^\infty$. It is easy to prove that cluster points are Caratheodory equilibria for both (2) and (3). One can ask whether all equilibria are cluster points: here, the metric and topological models are completely different. For the metric bounded confidence model, all Krasovsky equilibria are indeed cluster points (see [38, Prop. 7.1]). Instead, for the topological bounded confidence model, there exist equilibria that are not cluster points, in the following cases:

- for $\kappa > 1$, for any kind of solutions, see Example 7;
- for Krasovsky solutions even with $\kappa = 1$, see Example 8.

Instead, Caratheodory equilibria for $\kappa = 1$ are all clusters, as proved in Proposition 10.

Example 7 (Non-Convergence to Clusters, Topological $\kappa \geq 2$) Let $N = 7$, $n = 1$, $\kappa = 2$, $a \equiv 1$ and consider the point $\overline{x}$ with $\overline{x}_2 = \overline{x}_4 = \overline{x}_5 = 0$, $\overline{x}_1 = \frac{1}{2}$, $\overline{x}_3 = \overline{x}_6 = \overline{x}_7 = 1$. It can be easily computed that $f^t(\overline{x}) = 0$ and therefore $\overline{x}$ is an equilibrium point with respect to classical, Caratheodory, and Krasovsky solution. Moreover, it is not a cluster point as there is just $1 < \kappa = 2$ index with value $\frac{1}{2}$. We remark that $\overline{x}$ is *not* locally attractive with respect to all types of solutions.

As equilibria correspond to constant solutions, this example also shows that solutions of the topological bounded confidence model do not satisfy property P3), in general.

In the previous example we fixed $\kappa = 2$. For $\kappa = 1$, we will show that all classical and Caratheodory equilibria are cluster points, as stated by the following proposition.

Proposition 10 (Caratheodory Equilibria, Topological $\kappa = 1$) *If $\kappa = 1$, Caratheodory equilibria of (3) are cluster points.*

Proof As $\kappa = 1$, cluster points are such that agents can be divided into groups of at least two agents with the same value. The i-th component of the vector field writes $f_i^t(x) = a(\|x_{\Gamma(i)} - x_i\|)(x_{\Gamma(i)} - x_i)$, where $\Gamma(i)$ is the (state dependent) neighbor of i. Note that $f_i^t(x)$ is null if and only if $x_{\Gamma(i)} = x_i$, as $a(r) > 0$ for $r > 0$. Then, $\overline{x}_i = \overline{x}_{\Gamma(i)}$ for all $i \in V$ and $\overline{x}$ is a cluster point. □

Krasovsky equilibria which are not cluster points appear even if $\kappa = 1$, as shown by the following example. Their existence implies that, even if $\kappa = 1$, Krasovsky solutions do not necessarily converge to cluster points.

Example 8 (Non-Convergence of Krasovsky to Clusters, Topological $\kappa = 1$) We consider (3) with $N = 5$, $n = 1$, $\kappa = 1$, $a \equiv 1$ and the initial condition $\overline{x} = (-1, 1, 0, 1, -1)$. Observe that $\overline{x}$ is a discontinuity point of the vector field f^t. Among the limit values of the vector field f^t at $\overline{x}$ there are $(0, 0, -1, 0, 0)$ and $(0, 0, 1, 0, 0)$. Then $0 \in Kf^t(\overline{x})$ and $\overline{x}$ is a Krasovsky equilibrium which is not a cluster point.

4.2 P3) Convergence to Cluster Points

The following proposition summarizes our results about convergence to cluster points. The result is the best possible one in terms of convergence to clusters, given the above counterexamples.

Proposition 11 (Convergence to Cluster Points)

(i) ***Metric bounded confidence.*** *For any Krasovsky solution of* (2)*, Property P3) holds.*

(ii) ***Topological bounded confidence.*** *If* $\kappa = 1$*, for any Caratheodory solution of* (2)*, Property P3) holds.*

The proof of (i) is given in [38, Prop. 7.1] and = is based on the observation that (2) can be written as a gradient flow as follows. Define

$$\Phi_{ij}(r) = \begin{cases} \int_0^r a(s)s\,ds & \text{for } r < 1 \\ \int_0^1 a(s)s\,ds & \text{for } r \geq 1 \end{cases}$$

and observe that, if $\|x_i - x_j\| \neq 1$, for every $i \neq j$, then

$$\dot{x}_i = -\sum_{j \neq i} \nabla \Phi_{ij}(\|x_i - x_j\|).$$

This suggests to define the candidate Lyapunov function

$$V(x) = \sum_{i,j \neq i} \Phi_{ij}(\|x_i - x_j\|),$$

which satisfies $\dot{V}(x(t)) \leq 0$ for a.e. time and allows (despite being nonsmooth and non-proper) to establish an ad-hoc convergence argument.

The proof of (ii) requires a slightly different reasoning. The special case of (piecewise) classical solutions in dimension $n = 1$ was proved in [18] by exploiting the special structure of its induced graph $G(x)$, which we describe next.

Proposition 12 (Directed Pseudo-Forest) *If* $\kappa = 1$*, then for all* n*, for every* $x \in \mathbb{R}^{nN}$*, the interaction graph* $G(x)$ *of* (3) *is the union of weakly connected components, such that each component contains exactly one circuit of length 2 and the two nodes of the circuit can be reached from all nodes of the component.*

Examples of weakly connected components can be found in Figs. 1 and 2.

Proof Let x be fixed and consider a connected component of $G(x)$, called G'. We first prove that G' has exactly one circuit of length 2. Let M be the number of nodes of the connected component. As any node has exactly one out-edge, the number of edges of the component is exactly M, then G' contains one circuit (this kind of graph is referred to as directed pseudo-forest). Furthermore, we can observe that the nodes of the circuit are reachable from any node in G'. As any node has an outgoing edge, starting from any node there exists an infinite walk. As the number of nodes is finite, it must contain a circuit. This means that the walk contains the nodes of the circuit.

We now prove that any circuit cannot have length greater than 2. Assume by contradiction that there is a circuit with length $p > 2$. Let $i_1, \ldots, i_p$ be its nodes and i_1 be the smallest index. Thanks to the definition of neighbor, it must hold

$$\|x_{i_1} - x_{i_2}\| \leq \|x_{i_1} - x_{i_2}\| \leq \ldots \leq \|x_{i_1} - x_{i_p}\|.$$

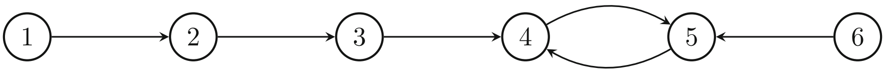


Fig. 1 Example of weakly connected component of graph $G(\overline{x})$, where $N = 6$, $n = 1$, $\kappa = 1$, $\overline{x} = (0, 10, 19, 27, 28, 30)$

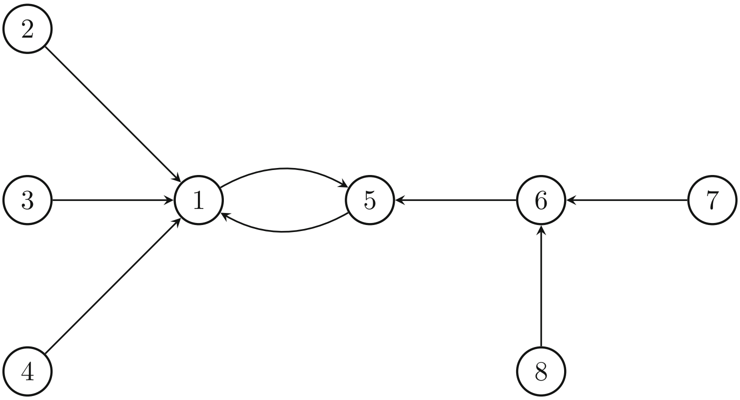


Fig. 2 Example of weakly connected component of graph $G(\bar{x})$ where $N = 8$, $n = 2$, $\kappa = 1$, $\bar{x}$ = ((0,0) (0,1),(-1,0), (0,-1), (1/2,0),(1,0),(1,1),(1,-1))

If $\|x_{i_1} - x_{i_2}\| < \|x_{i_1} - x_{i_p}\|$, then i_p is the neighbor of i_2 instead of i_1, contradiction. Then it must hold $\|x_{i_1} - x_{i_2}\| = \|x_{i_1} - x_{i_2}\| = \ldots = \|x_{i_1} - x_{i_p}\|$. In this case i_1 should be the neighbor of all nodes $i_3, i_4, \ldots, i_{p-1}$, as it is the smallest index. Finally G' has exactly one circuit: indeed, to connect two circuits there should be a node with out-degree at least 2 (Fig. 1). □

An interaction graph with the structure of $G(x)$, if kept static, would guarantee convergence to consensus for each connected component and, therefore, convergence to a cluster point. However, the graph $G(x(t))$ evolves with time in such a way that connected components can split and distinct connected components can merge. The latter phenomenon is illustrated in the following example (Fig. 2).

Example 9 (Merging Components in Caratheodory Solutions) Let $N = 4$, $n = 1$, $\kappa = 1$, $a \equiv 1$ and consider the initial condition $\overline{x} = (-1, 0, 1, 1)$. Consider the Caratheodory solution $x(t) = (1 - te^{-t} - 2e^{-t}, 1 - e^{-t}, 1, 1)$. Note that $x(0) = \overline{x}$ and $x(t)$ that satisfies (3) at all $t > 0$ but not at $t = 0$. The graph $G(x(0))$ has two connected components whose vertices are $\{1, 2\}$ and $\{3, 4\}$ whereas $G(x(t))$ is connected for all $t > 0$.

This counterexample prevents leveraging the topology of $G(x)$ to prove (ii) for Caratheodory solutions. We therefore resort to a Lyapunov-like argument, which is partly inspired by the one in [38] for metric interactions, but will be valid for topological interactions in the case of $\kappa = 1$ only.

We introduce the integral function

$$I(r) := \int_0^r a(s)s\,ds$$

and write the candidate Lyapunov function

$$W(x) := \sum_{i,j \in N_i^t(x)} I(\|x_j - x_i\|). \tag{29}$$

One might hope to write (3) as $\dot{x} = -\nabla W(x)$, like in the metric case. This is false, as one can easily observe that this expression entails interactions that are symmetric, while this is not the case for (3).

We will anyway be able to prove that $W(x)$ is a Lyapunov function for solutions to (3). This is only the case for $\kappa = 1$ and for Caratheodory solutions, but the proof is quite different from the case of metric bounded confidence (2), again due to the asymmetry of the interactions. Instead, $W(x)$ is not a Lyapunov function, neither for Caratheodory solutions with $\kappa > 1$ nor for Krasovsky solutions with $\kappa \geq 1$, as shown by Examples 10 and 11 below.

Proposition 13 (*W* **Is Lyapunov**) *Let $\kappa = 1$. Then, the function $W(x(t))$ is continuous and non-increasing for Caratheodory solutions.*

Proof The proof is based on rewriting $W(x) = \sum_{i=1}^N W_i(x)$ where

$$W_i(x) = \min_{j \neq i} I(\|x_i - x_j\|). \tag{30}$$

It is clear that both $I(r)$ and $x(t)$ are continuous. Then, both all $W_i(x(t))$ and their sum $W(x(t))$ are continuous too. The rest of the proof is based on Danskin theorem[1] [24] for $W_i(x)$. Similarly to the proof of Proposition 6, even though $W_i(x)$ can be non-differentiable, it admits directional derivative with respect to any direction. We apply it to our function, denoting the direction of displacement with $h = (h_1, \ldots, h_N)$, where each h_k is the n-dimensional direction of displacement of the position of the agent k. By applying Danskin formula, the directional derivative D_h along h is given by

$$\begin{aligned}
D_h W_i(x) &= \min_{j \in A_i(x)} \sum_{k=1}^N h_k \cdot \nabla_{x_k} I(\|x_i - x_j\|) = \min_{j \in A_i(x)} \sum_{k \in \{i,j\}} h_k \cdot \nabla_{x_k} I(\|x_i - x_j\|) \\
&= \min_{j \in A_i(x)} \big(h_i \cdot a(\|x_i - x_j\|)(x_i - x_j) - h_j \cdot a(\|x_i - x_k\|)(x_i - x_j)\big) \\
&= \min_{j \in A_i(x)} a(\|x_i - x_j\|)(h_i - h_j) \cdot (x_i - x_j),
\end{aligned}$$

[1] In Danskin notation, we have $F = F(x, j) = I(\|x_i - x_j\|)$ maximized with respect to $j \in V \setminus \{i\}$.

where $A_i(x)$ is the set of indexes $j \neq i$ realizing $\min I(\|x_i - x_j\|)$.

We now apply this formula to compute the time derivative $\dot{W}_i(x(t))$, whenever it exists. Following the same computations as for (14), we have $\dot{W}_i(x(t)) = D_{\dot{x}(t)} W_i(x(t))$. Since the time derivative $\dot{x}(t)$ exists for almost every time $t \in (0, T)$, this holds for $\dot{W}_i(x(t))$ too. We compute this derivative, by restricting ourselves to Caratheodory solutions, that satisfy $\dot{x}_i = a(\|x_i - x_k\|)(x_k - x_i)$ for almost every time, with $k \in N_i^t(x)$. Denote with $L := \|x_i - x_k\|$ and with l the unique element $l \in N_k^t(x)$. Observe that $l \in N_k^t(x)$ implies $\|x_l - x_k\| \leq \|x_i - x_k\| = L$ that in turn implies $a(\|x_l - x_k\|) \leq a(\|x_i - x_k\|) = a(L)$ and $(x_l - x_k) \cdot (x_i - x_k) \geq -L^2$. Since $k \in A_i(x)$, and using previous estimates, it holds

$$\begin{aligned}
D_{\dot{x}(t)} W_i(x(t)) &= \min_{j \in A_i(x)} a(\|x_i - x_j\|)(\dot{x}_i - \dot{x}_j) \cdot (x_i - x_j) \\
&\leq a(L)(\dot{x}_i - \dot{x}_k) \cdot (x_i - x_k) \\
&= a(L)(a(\|x_i - x_k\|)(x_k - x_i) \cdot (x_i - x_k) \\
&\quad -a(\|x_k - x_l\|)(x_l - x_k) \cdot (x_i - x_k)) \\
&\leq -a(L)^2 L^2 + a(L)^2 L^2 = 0.
\end{aligned} \tag{31}$$

Since $W_i(x(t))$ is continuous, this implies that each $W_i(x(t))$ is non-increasing.

Passing to $W(x(t))$, that is a finite sum of continuous and non-increasing functions, the proof follows. □

Example 10 ($W(x(t))$ Increasing if $\kappa > 1$) We now prove that $W(x)$ given in (29) is not a Lyapunov function for (3) in the case $\kappa > 1$ for Caratheodory solutions (hence for Krasovsky solutions too). Consider the following initial configuration of $N = 8$ agents on the real line with $\kappa = 2$ and $a(r) \equiv 1$: choose $\bar{x} = (-9, -9, -9, -2, 2, 9, 9, 9)$ and observe that the unique solution of (3) in the Krasovsky sense (that is even classical and Caratheodory) is given by

$$\begin{cases}
\dot{x}_1 = \dot{x}_2 = \dot{x}_3 = \dot{x}_6 = \dot{x}_7 = \dot{x}_8 = 0 \\
\dot{x}_4 = (x_5 - x_4) + (x_1 - x_4) \\
\dot{x}_5 = (x_4 - x_5) + (x_6 - x_5).
\end{cases} \tag{32}$$

By symmetries, it holds $x_4(t) = -x_5(t)$, thus $\dot{x}_5 = 9 - 3x_5$. It then holds $x_5(t) = 3 - e^{-3t}$. Notice that the topology does not change and the solution converges to the equilibrium point $x^\infty = (-9, -9, -9, -3, 3, 9, 9, 9)$ which is not a cluster point. A direct computation gives

$$\begin{aligned}
W(x(t)) &= \tfrac{1}{2}\left((x_1(t) - x_4(t))^2 + 2(x_4(t) - x_5(t))^2 + (x_5(t) - x_6(t))^2\right) \\
&= 4x_5^2(t) + (9 - x_5(t))^2,
\end{aligned} \tag{33}$$

that satisfies $\dot{W}(x(t)) = (10x_5(t) - 18)\dot{x}_5(t) > 0$ for all $t \in [0, +\infty)$.

Example 11 ($W(x(t))$ Increasing for Krasovsky Solutions) We now prove that $W(x)$ given in (29) is not a Lyapunov function for Krasovsky solutions with $\kappa = 1$. Consider the following initial configuration of $N = 5$ agents on the real line and $a(r) \equiv 1$: choose $\bar{x} = (-1 - y_0, -1 + y_0, 0, 1 - y_0, 1 + y_0)$ with $y_0 \in \left(0, \frac{1}{17}\right)$. Observe that one of the solutions of (3) in the Krasovsky sense is

$$x_1(t) = -1 - y(t),\ x_2(t) = -1 + y(t),\ x_3(t) = 0,\ x_4(t) = 1 - y(t),\ x_5(t) = 1 + y(t),$$

with $y(t) = \exp(-2t)y_0$. Indeed, for all $t > 0$ this solution satisfies

$$\begin{aligned}\dot{x} &= \frac{1}{2}(x_2 - x_1, x_1 - x_2, x_4, x_5 - x_4, x_4 - x_5)\\ &\quad + \frac{1}{2}(x_2 - x_1, x_1 - x_2, x_2, x_5 - x_4, x_4 - x_5)\\ &= (2y(t), -2y(t), 0, 2y(t), -2y(t)) .\end{aligned}$$

A direct computation gives

$$\begin{aligned}2W(x(t)) &= 2(x_1(t) - x_2(t))^2 + (x_3(t) - x_2(t))^2 + 2(x_4(t) - x_5(t))^2\\ &= 4(2y(t))^2 + (1 - y(t))^2 = 1 - 2y(t) + 17y(t)^2.\end{aligned}$$

Its derivative is $4y(t) - 17 \cdot 4y(t)^2$, that is positive for $y(t) \in \left(0, \frac{1}{17}\right)$. This holds whenever $y_0 \in \left(0, \frac{1}{17}\right)$. As a consequence, $W(x(t))$ is strictly increasing.

We are now ready to describe the structure of the limits of Caratheodory solutions of (3) with $\kappa = 1$ that are indeed clusters.

Proposition 14 (Convergence and Cluster Properties, Topological $\kappa = 1$) *Let $x_1(t), \ldots, x_N(t)$ be a Caratheodory solution of* (3) *with $\kappa = 1$. Then, the following clustering properties hold:*

- *each agent satisfies $\lim_{t\to+\infty} x_i(t) = x_i^\infty$ for some $x_i^\infty \in \mathbb{R}^n$;*
- *for each $i \in V$ there exists at least one $j \neq i$ such that $x_i^\infty = x_j^\infty$.*

This also implies that P3) holds and $W(x^\infty) = 0$.

Proof First recall that $x(t)$ is bounded, due to contractivity of the support proved in Proposition 6. This implies that $a(\|x_i(t) - x_j(t)\|)$ is bounded too, as $a(r)$ is Lipschitz continuous by hypothesis. This in turn implies that both $x(t)$ and $a(\|x_i(t) - x_j(t)\|)$ are Lipschitz continuous too. Boundedness also implies that the ω-limit is bounded.

Fix now any $x^* = (x_1, \ldots, x_N)$ in the ω-limit of $x(t)$. By definition, there exists a sequence $t_k \to +\infty$ such that $x(t_k) \to x^*$. Fix $\varepsilon > 0$ and $K = K_\varepsilon$ sufficiently large to have

$$\|((x_i(t_k) - x_j(t_k)) \cdot (x_l(t_k) - x_m(t_k))) - (x_i^* - x_j^*) \cdot (x_l^* - x_m^*)\| < 2\varepsilon \quad (34)$$

for all $i, j, l, m \in V$ and $k > K_\varepsilon$. Since trajectories are bounded and Lipschitz continuous, there exists a uniform $\delta > 0$ such that

$$\|((x_i(t_k+\tau) - x_j(t_k+\tau)) \cdot (x_l(t_k+\tau) - x_m(t_k+\tau))) - (x_i^* - x_j^*) \cdot (x_l^* - x_m^*)\| < \varepsilon \quad (35)$$

for all $\tau \in (-N^{3N}\delta, N^{3N}\delta)$.

Fix now $i = 1$, recall (30) and consider the derivative

$$D_{\dot{x}(t)} W_1(x(t)) = \min_{j \in A_1(x(t))} a(\|x_1(t) - x_j(t)\|)(\dot{x}_1(t) - \dot{x}_j(t)) \cdot (x_1(t) - x_j(t)), \quad (36)$$

whenever $\dot{x}(t)$ is well-defined, i.e. for almost every $t > 0$. Since the number of nearest neighbors of x_1 in $A_1(x)$ is $N-1$ at most, there exists at least one index j_1^k such that j_1^k is the minimizer in the right-hand side of (36) for all $\tau \in I_k^{1,a}$, where $I_k^{1,a} \subset (t_k - N^{3N}\delta, t_k + N^{3N}\delta)$ has Lebesgue measure $2N^{3N-1}\delta$. Since the number of possible j_1^k is finite, eventually passing to a subsequence in k, we assume that $j_1^k = j_1$ is constant. By recalling that Caratheodory solutions satisfy the dynamics (3) for almost every time, it holds

$$\begin{aligned} D_{\dot{x}(t)} W_1(x(t)) = a(\|x_1(t) - x_{j_1}(t)\|)[&a(\|x_1(t) - x_l(t)\|)(x_l(t) - x_1(t)) \\ &- a(\|x_{j_1}(t) - x_m(t)\|)(x_m(t) - x_{j_1}(t))] \cdot (x_1(t) - x_{j_1}(t)), \end{aligned}$$

for almost every $t \in I_k^{1,a}$, where $l \in N_1^t(x(t))$ and $m \in N_{j_1}^t(x(t))$. Again, since the number of possible neighbors of 1 in $N_1^t(x(t))$ is $N-1$ at most, then there exists l_1^k such that $l_1^k \in N_1^t(x(t))$ for all $t \in I_k^{1,b}$ where $I_k^{1,b} \subset I_k^{1,a}$ has Lebesgue measure $2N^{3N-2}\delta$. By passing to a subsequence in k, we can assume l_1 constant. With a similar argument, we can find m_1 and $I_k^{1,c} \subset I_k^{1,b}$ with Lebesgue measure $2N^{3N-3}\delta$ such that $m_1 \in N_{j_1}^t(x(t))$ for all $t \in I_k^{1,c}$.

We now choose the index 2 and define the corresponding indexes j_2, l_2, m_2 and sets $I^{2,c} \subset I^{2,b} \subset I_k^{2,a} \subset I_k^{1,c}$, each with Lebesgue measure being $1/N$ of the previous one. We then move to indexes $3, 4, \ldots N$, finally reaching $I_k := I_k^{N,c}$ with Lebesgue measure 2δ and such that, for each $i \in V$ there exists corresponding j_i, l_i, m_i such that for all $\tau \in I_k$ the following hold:

- the index j_i is the minimizer in the right-hand side of (36);
- the index l_i is the unique element of $N_i^t(x(\tau))$;
- the index m_i is the unique element of $N_{j_i}^t(x(\tau))$.

Fix now any $i \in V$ and the corresponding j_i, l_i, m_i defined above. We now prove that it holds

$$\mathcal{A}_i := a(\|x_i^* - x_{j_i}^*\|)[a(\|x_i^* - x_{l_i}^*\|)(x_{l_i}^* - x_i^*) \cdot (x_i^* - x_{j_i}^*) \quad (37)$$

$$-a(\|x_{j_i}^* - x_{m_i}^*\|)(x_{m_i}^* - x_{j_i}^*) \cdot (x_i^* - x_{j_i}^*)] = 0.$$

By contradiction, first assume that $\mathcal{A}_i > 0$: then, observe that (37) coupled with (35), implies that there exists $\bar{k}$ such that $D_{\dot{x}(\tau)} W_i(x(\tau)) > \mathcal{A}_i/2$ for every $\tau \in I_k$ with $k \geq \bar{k}$. Since the set of such τ has non-zero Lebesgue measure, this contradicts the fact that $W_i(x(t))$ is a non-increasing function.

Assume now $\mathcal{A}_i < 0$ and use the same reasoning to prove that $D_{\dot{x}(\tau)} W_i(x(\tau)) < -|\mathcal{A}_i|/2$ for every $\tau \in I_k$ with $k \geq \bar{k}$. Since for all times in $(t_k - N^{3N}\delta, t_k + N^{3N}\delta)$ we have $W_i(x(t))$ non-increasing, we can write

$$\begin{aligned} W_i(x(t_k + N^{3N}\delta)) &\leq W_i(x(t_k - N^{3N}\delta)) + \int_{I_k} d\tau\, D_{\dot{x}(\tau)} W_i(x(\tau)) \\ &\leq W_i(x(t_k - N^{3N}\delta)) - \delta|\mathcal{A}_i|. \end{aligned} \tag{38}$$

This implies $\lim_{t_k \to +\infty} W_i(x(t_k + N^{3N}\delta)) = -\infty$. This contradicts the fact that W_i is bounded from below. We have now proved (37).

We now prove that (37) ensures $W(x^*) = 0$. For each $i \in V$, recall the definition of corresponding indexes j_i, l_i, m_i given above. For $i = 1$, condition (37) implies one of the following cases:

- **Case 1A)** the index j_1 satisfies $\|x_1^* - x_{j_1}^*\| = 0$. This in turn implies that the only $j \in N_1^t(x^*)$ satisfies $\|x_1^* - x_j^*\| \leq \|x_1^* - x_{j_1}^*\| = 0$. This in turn implies $a(\|x_1^* - x_j^*\|) = 0$, i.e. $W_1(x^*) = 0$.
- **Case 1B)** the index j_1 satisfies $\|x_i^* - x_{j_1}^*\| \neq 0$. Observe that, by construction of j_1, l_1, it holds $\|x_i(t) - x_{j_1}(t)\| = \|x_i(t) - x_{l_1}(t)\|$, thus by continuity it holds $\|x_i^* - x_{l_1}^*\| = \|x_i^* - x_{j_1}^*\| \neq 0$. Moreover, the definition of j in (36) implies that the following estimate holds

$$\begin{aligned} &a(\|x_1(t_k) - x_{j_1}(t_k)\|)[a(\|x_1(t_k) - x_{l_1}(t_k)\|)(x_{l_1}(t_k) - x_1(t_k)) \\ &\quad - a(\|x_{j_1}(t_k) - x_{m_1}(t_k)\|)(x_{m_1}(t_k) - x_{j_1}(t_k))] \cdot (x_1(t_k) - x_{j_1}(t_k)) \leq \\ &a(\|x_1(t_k) - x_{l_1}(t_k)\|)[a(\|x_1(t_k) - x_{l_1}(t_k)\|)(x_{l_1}(t_k) - x_1(t_k)) \\ &\quad - a(\|x_{l_1}(t_k) - x_{l_{l_1}}(t_k)\|)(x_{l_{l_1}}(t_k) - x_{l_1}(t_k))] \cdot (x_1(t_k) - x_{l_1}(t_k)) \leq 0, \end{aligned} \tag{39}$$

where l_{l_1} is the unique index in $N_{l_1}^t(x(t_k))$. The last inequality can be proved as in (31). The left-hand side of (39) converges to (37) as $t_k \to +\infty$, thus it converges to zero. Then, the middle term converges to zero too, i.e.

$$-a(\|x_1^* - x_{l_1}^*\|)\|x_{l_1}^* - x_1^*\|^2 = a(\|x_{l_1}^* - x_{l_{l_1}}^*\|)(x_{l_{l_1}}^* - x_{l_1}^*) \cdot (x_1^* - x_{l_1}^*). \tag{40}$$

Here we used the fact that $a(\|x_1^* - x_{l_1}^*\|) \neq 0$. Since $\|x_1^* - x_{l_1}^*\| \geq \|x_{l_1}^* - x_{l_{l_1}}^*\|$ by construction of l_{l_1} and a is non-decreasing, the only possibility for (40) to hold is to have $\|x_{l_1}^* - x_{l_{l_1}}^*\| = \|x_1^* - x_{l_1}^*\| \neq 0$ and $(x_{l_{l_1}}^* - x_{l_1}^*) \cdot (x_1^* - x_{l_1}^*) = -\|x_{l_1}^* - x_1^*\|^2$, i.e.

$x_1, x_{l_1}, x_{l_{l_1}}$ being on the same line with x_{l_1} as middle point. This also implies that $1, l_1, l_{l_1}$ are all different indexes. We relabel l_1, l_{l_1} as indexes $2, l_2$, for simplicity of notation.

We then apply the same idea to index 2 (either coming from relabeling or not), and we have the following cases:

- **Case 2A)** The index j_2 satisfies $\|x_2^* - x_{j_2}^*\| = 0$. Since $j_2 \in A_2(x(t))$ for all $t \in I_k$ and l_2 is the unique element of $N_2^t(x(t))$, this implies

$$\|x_2(t) - x_{l_2}(t)\| = \|x_2(t) - x_{j_2}(t)\|,$$

thus $\|x_2^* - x_{l_2}^*\| = 0$ by continuity. This implies that **Case 1B)** is not compatible with **Case 2A)**: indeed, (40) implies $a(\|x_1^* - x_{l_1}^*\|)\|x_{l_1}^* - x_1^*\|^2 = 0$, thus $\|x_1^* - x_{l_1}^*\| = 0$ and, by continuity, it holds $\|x_1^* - x_{j_1}^*\| = 0$.
- **Case 2B)** The index j_2 satisfies $\|x_2^* - x_{j_2}^*\| \neq 0$. By following the reasoning of **Case 1B)**, we find that $2, l_2, l_{l_2}$ are aligned, with l_2 being the middle point. This implies that, if both **Case 1B)** and **Case 2B)** hold, then $1, 2, l_2, l_{l_2}$ are aligned, each with the same distance with respect to the previous one. Like in **Case 1B)**, we also have that the four indexes are all distinct. This also allows to relabel l_2, l_{l_2} as $3, l_3$, for simplicity of notation.

We now apply the same reasoning to all indexes $i = 3, \ldots, N$ either after relabeling (due to **Case** i**B**) or not. By incompatibility between **Case** i**B** and **Case** $(i+1)$**A**, we have the following structure: we first have i cases that are **Cases 1A-2A-...-**i**A**, then $N - i$ cases that are **Cases** $(i+1)$**B-...-**N**B**. We prove that $i = N$, by contradiction. Observe that **Cases** $(i+1)$**B-...-**N**B** force us to have agents $i+1, \ldots, N, l_N$ aligned on the same line, each with the same distance with respect to the previous one. Since the number of agents is N, the agent l_N is one among $1, \ldots, N$. By the alignment condition, it cannot be any of the agents $i+1, \ldots, N$, hence **Case** l_N**A** holds. By incompatibility of conditions described above, **Case** N**B** cannot hold. This raises a contradiction. As a consequence, for each $i \in V$ the **Case** i**A** is satisfied. This means that for each $i \in V$ there exists $j \neq i$ such that $x_i^* = x_j^*$. This also implies $W_i(x^*) = 0$. In particular, x^* satisfies the second statement of this proposition.

We are now left to prove that the ω-limit is reduced to a single point, i.e. that x^* given above is x^∞ in the first statement of this proposition. Define the following equivalence relation: $i \sim j$ if $x_i^* = x_j^*$. Observe that each class of equivalence $[i]_\sim$ is composed of at least two elements. We have two possibilities:

- There exists a single class of equivalence $[i]_\sim$. Then, for each $\varepsilon > 0$ there exists t_k such that $\|x_i(t_k) - x_i^*\| = \|x_i(t_k) - x_1^*\| < \varepsilon$. Since this holds for all indexes, then the support of the solution $x(t_k)$ is contained in $B(x_1^*, \varepsilon)$. Since the support is non-increasing, due to Proposition 6, then the solution $x(t)$ is contained in $B(x_1^*, \varepsilon)$ for all $t \geq t_k$. Since this condition holds for all $\varepsilon > 0$, then $x_i(t) \to x_1^* = x_i^*$.

- There exist at least two classes of equivalence $[i]_\sim \neq [j]_\sim$. Define the minimal distance between clusters as $5\lambda := \min_{i \not\sim j} \|x_i^* - x_j^*\|$ that satisfies $\lambda > 0$. By convergence of $x_i(t_k)$ to x_i^*, there exists $\bar{k}$ sufficiently large to have $\|x_i(t_{\bar{k}}) - x_i^*\| < \lambda$ for all $i \in V$. As a consequence, the following **cluster separation condition** holds:

$$\text{If } i \sim j, \text{ then it holds } \|x_i(t_{\bar{k}}) - x_j(t_{\bar{k}})\| < 2\lambda.$$
$$\text{If } i \not\sim j, \text{ then it holds } \|x_i(t_{\bar{k}}) - x_j(t_{\bar{k}})\| > 3\lambda.$$

  It is now easy to prove that, for all $t \geq t_{\bar{k}}$, the same cluster separation condition holds too, since interactions between agents of different clusters do not occur: the proof is similar to Proposition 6 and is omitted. As a consequence, each of the cluster acts as an independent system starting from $t_{\bar{k}}$. In particular, we can apply Proposition 6 to each cluster independently: similarly to the previous case, for each $\varepsilon > 0$ there exists $k \geq \bar{k}$ such that for each class of equivalence $[i]_\sim$ the support of $\{x_j(t) \text{ s.t. } j \in [i]_\sim\}$ is contained in $B(x_i^*, \varepsilon)$ for all $t \geq t_k$. By letting $\varepsilon \to 0$, we have $x_j(t) \to x_i^*$.

In both cases, we have proved that the ω-limit of $x(t)$ is reduced to x^*. Thus, by choosing $x^\infty = x^*$ we have that the statement is proved. □

5 Future Directions

In this paper we explored various concepts of solutions for discontinuous differential equations, motivated by social dynamics models. In particular, we focused on the so-called bounded confidence models, where each agent is interacting either with neighbors within a fixed distance (metric case) or with the κ closest ones (topological case). As per the concepts of solutions we focused on Caratheodory and Krasovsky, after proving that the set of Filippov solutions coincides with that of Krasovsky solutions for the considered models.

Existence of solutions and uniqueness for almost every initial datum are proved in Krasovsky and Caratheodory sense for both models. We also explored properties of solutions such as preservation of the average, contractivity of support, and convergence to cluster points. Contractivity of the support always holds true, the other properties hold for the metric case (and both concepts of solutions), while they fail for the topological case with the exception of convergence to cluster points that holds for Caratheodory solutions if $\kappa = 1$.

Future investigations may include:

- Exploring existence, uniqueness, and properties of trajectories for other concepts of solutions, such as limit of Euler or CLSS, stratified solutions, and others [38];
- Studying the implications of our results to approximated solutions produced by numerical schemes;

- Considering the topological-metric case, where each agent interacts with the closest κ neighbors if they are within a fixed distance;
- Extending the scope of our analysis to include dynamical models with other types of discontinuities, as those generated by quantization [17] or hybrid setting [28].

References

1. G. Ajmone Marsan, N. Bellomo, and L. Gibelli. Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics. *Mathematical Models and Methods in Applied Sciences*, 26(06):1051–1093, 2016.
2. G. Albi, N. Bellomo, L. Fermo, S.-Y. Ha, J. Kim, L. Pareschi, D. Poyato, and J. Soler. Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives. *Mathematical Models and Methods in Applied Sciences*, 29(10):1901–2005, 2019.
3. F. Ancona and A. Bressan. Patchy vector fields and asymptotic stabilization. *ESAIM. Control, optimisation and calculus of variations*, 4:445–471, 1999.
4. J.-P. Aubin and A. Cellina. *Differential inclusions*. Springer-Verlag, Berlin, 1984.
5. A. Aydoğdu, M. Caponigro, S. McQuade, B. Piccoli, N. Pouradier Duteil, F. Rossi, and E. Trélat. Interaction network, state space, and control in social dynamics. In N. Bellomo, P. Degond, and E. Tadmor, editors, *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 99–140. Springer, 2017.
6. B. Aylaj, N. Bellomo, L. Gibelli, and A. Reali. A unified multiscale vision of behavioral crowds. *Mathematical Models and Methods in Applied Sciences*, 30(01):1–22, 2020.
7. M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008.
8. A. L. Bertozzi, T. Laurent, and J. Rosado. L^p theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64(1):45–83, 2011.
9. V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. Existence and uniqueness of solutions for a continuous-time opinion dynamics model with state-dependent connectivity. Website supplement, 2009.
10. V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54(11):2586–2597, 2009.
11. V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. Continuous-time average-preserving opinion dynamics with opinion-dependent communications. *SIAM Journal on Control and Optimization*, 48(8):5214–5240, 2010.
12. M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse stabilization and control of alignment models. *Mathematical Models and Methods in Applied Sciences*, 25(03):521–564, 2015.
13. J. A. Carrillo, Y.-P. Choi, and M. Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, pages 1–46. Springer, Vienna, 2014.
14. J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 42(1):218–236, 2010.
15. F. Ceragioli and P. Frasca. Continuous and discontinuous opinion dynamics with bounded confidence. *Nonlinear Analysis: Real World Applications*, 13(3):1239–1251, 2012.
16. F. Ceragioli and P. Frasca. Consensus and disagreement: The role of quantized behaviors in opinion dynamics. *SIAM Journal on Control and Optimization*, 56(2):1058–1080, 2018.

17. F. Ceragioli and P. Frasca. Discontinuities, generalized solutions, and (dis) agreement in opinion dynamics. In *Control Subject to Computational and Communication Constraints*, pages 287–309. Springer, 2018.
18. F. Ceragioli, P. Frasca, and W. S. Rossi. Modeling Limited Attention in Opinion Dynamics by Topological Interactions. In: Lasaulce S., Mertikopoulos P., Orda A. (eds) Network Games, Control and Optimization. NETGCOOP 2021. *Communications in Computer and Information Science*, vol 1354. Springer, Cham. https://doi.org/10.1007/978-3-030-87473-5_24
19. N. R. Chowdhury, I. C. Morărescu, S. Martin, and S. Srikant. Continuous opinions and discrete actions in social networks: A multi-agent system approach. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 1739–1744, 2016.
20. F. H. Clarke, Y. S. Ledyaev, E. D. Sontag, and A. I. Subbotin. Asymptotic controllability implies feedback stabilization. *IEEE Transactions on Automatic Control*, 42(10):1394–1407, 1997.
21. E. Cristiani, P. Frasca, and B. Piccoli. Effects of anisotropic interactions on the structure of animal groups. *Journal of mathematical biology*, 62(4):569–588, 2011.
22. E. Cristiani, B. Piccoli, and A. Tosin. Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints. In G. Naldi, L. Pareschi, and G. Toscani, editors, *Mathematical modeling of collective behavior in socio-economic and life sciences*, Modeling and Simulation in Science, Engineering and Technology, pages 337–364. Birkhäuser, Boston, 2010.
23. E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*. MS&A. Springer International Publishing, 2014.
24. J. M. Danskin. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics*, 14(4):641–664, 1966.
25. P. Degond, G. Dimarco, and T. B. N. Mac. Hydrodynamics of the Kuramoto–Vicsek model of rotating self-propelled particles. *Mathematical Models and Methods in Applied Sciences*, 24(02):277–325, 2014.
26. R. Dunbar. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6):469–493, 1992.
27. A. Filippov. *Differential Equations with Discontinuous Right-hand Sides*, volume 18 of *Mathematics and its Applications*. Springer, Netherlands, 1988.
28. P. Frasca, S. Tarbouriech, and L. Zaccarian. Hybrid models of opinion dynamics with opinion-dependent connectivity. *Automatica*, 100:153–161, 2019.
29. J. R. P. French. A formal theory of social power. *Psychological review*, 63(3):181, 1956.
30. R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation*, 5(3), 2002.
31. O. Hájek. Discontinuous differential equations, i. *Journal of Differential Equations*, 32(2):149–170, 1979.
32. P.-E. Jabin and S. Motsch. Clustering and asymptotic behavior in opinion formation. *Journal of Differential Equations*, 257(11):4165 – 4187, 2014.
33. A. Marigo and B. Piccoli. Regular syntheses and solutions to discontinuous ODEs. *ESAIM: Control, Optimisation and Calculus of Variations*, 7:291–307, 2002.
34. A. C. Martins. Continuous opinions and discrete actions in opinion dynamics problems. *International Journal of Modern Physics C*, 19(04):617–624, 2008.
35. S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56(4):577–621, 2014.
36. R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.
37. B. Piccoli and F. Rossi. Measure-theoretic models for crowd dynamics. In *Crowd Dynamics, Volume 1*, pages 137–165. Springer, 2018.
38. B. Piccoli and F. Rossi. Generalized solutions to bounded-confidence models. *Mathematical Models and Methods in Applied Sciences*, 31(06): 1237–1276 (2021).
39. B. Piccoli and H. J. Sussmann. Regular synthesis and sufficiency conditions for optimality. *SIAM Journal on Control and Optimization*, 39(2):359–410, 2000.

40. A. V. Proskurnikov and R. Tempo. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43:65–79, 2017.
41. A. V. Proskurnikov and R. Tempo. A tutorial on modeling and analysis of dynamic social networks. Part II. *Annual Reviews in Control*, 45:166–190, 2018.
42. W. S. Rossi and P. Frasca. Opinion dynamics with topological gossiping: Asynchronous updates under limited attention. *IEEE Control Systems Letters*, 4(3):566–571, 2020.
43. H. Sussmann. Synthesis, presynthesis, sufficient conditions for optimality and subanalytic sets. In *Nonlinear Controllability and Optimal Control*, pages 1–19. CRC Press, New York, 1990.
44. Y. Yang, D. V. Dimarogonas, and X. Hu. Opinion consensus of modified Hegselmann–Krause models. *Automatica*, 50(2):622–627, 2014.

Crowd Behaviour Understanding Using Computer Vision and Statistical Mechanics Principles

Zoheir Sabeur and Banafshe Arbab-Zavar

Abstract Crowd behaviour understanding in computer science is a research discipline which has grown rapidly in recent years. Specifically, we are currently able to generate large and high-resolution observation data through crowd sensing in varieties of spatial environments. This has also given us the advantage to adopt computer vision methods for detecting human behaviour. In this study, we adopted statistical mechanics principles with analogies of entropy and kinetic energy in classical molecular gases to derive features which describe crowd motions. These are implicitly measured, as basis for understanding *behaviour*, using a holistic three-dimensional representation, of crowd features including *structure*, *energy* and *translation*. As a result, we measured those features using computer vision in the view of machine understanding crowd behaviour. *Usual behaviour* is established from our expected crowd motions in context of the specific recipient spaces of our experiments. The behaviour which does not fall within the expected usual behaviour measurement is considered as an *unusual behaviour*. This research work was initiated in 2013 under the eVACUATE project, while it is currently ongoing under the S4AllCities project since 2020.

1 Introduction

The most recent development in understanding crowd behaviour using machine intelligence over the last two decades has been boosted by the affordability and availability of smart sensing for observing humans in spatial environments. In particular, the deployment of smart CCTV cameras for monitoring the safety and security of citizens in public spaces has become the regulated norm of security modus operandi in the majority of modern cities around the world. But in order to reach scalability in the surveillance of large crowds' critical safety in public spaces,

Z. Sabeur (✉) · B. Arbab-Zavar
Department of Computing and Informatics, Faculty of Science and Technology, Bournemouth University, Dorset, England, UK
e-mail: zsabeur@bournemouth.ac.uk

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_3

one will require the support of computer machine intelligence for focusing only on those events where specifically detected behaviour requires attention from security practitioners and first responders. It is in this perspective that we specialised over the years, and since the launch of the international eVACUATE research project [1], on the automated detection of individuals or groups of humans and unusual behaviour in public spaces using artificial intelligence (AI). Specifically, we characterised crowd as a dynamic system whose thermodynamic-related parameters can be directly or indirectly measured using sensory observations including vision-based features. These are then used for detecting and understanding human behaviour while adapting statistical mechanics principles which relate to the state of dynamic order or disorder of the crowd system at multiple spatial scales. This was the foundation of our research work over the years, which is ongoing, regarding human behaviour understanding using AI.

1.1 Definition of Crowd

The definition of crowd, for understanding crowd behaviour, requires deep insights into the important selected features representing crowds. These should be viewed in context of a complex dynamic system. A crowd 'system' can be considered as a collection of loosely coordinated individuals who may share a common and temporarily bound interest. This covers spectators and people moving.

From a practical sense, there are some nuances to the simple definition presented above. Figure 1 shows two images of a pedestrian crossing at two different times. Initially there are two distinct crowds which each desire to cross to the other side of the road. This is shown by the two large red ellipses in Fig. 1a. After the situation has evolved, there are a number of different possibilities as shown in the red ellipses in Fig. 3b. These two figures show that the notion of crowd refers to a dynamically changing system which may potentially undergo some phase transitions, in this case crowd splitting or merging through time. The intention of each individual in the crowd is unpredictable instantaneously, but it could become understood, therefore possibly predictable over the time, as it is shown in Fig. 1.

1.1.1 Crowd Multiple Scales

There is also the notion of scales which needs to be considered when one tries to understand crowd behaviour. This is indeed needed to study dynamically complex systems. Namely, and for the case of crowd systems, the observed collective crowd behaviour is related to the inner dynamics of each behaviour of individuals within the crowd. It also includes the entailed learning processes between individuals and their abilities to share a collective intelligent crowd behaviour.

Fig. 1 Pedestrian crossing showing dynamic evolution of crowds at different times (**a**) and (**b**)

Fig. 2 Microscale view of crowd with single individuals

Crowd modelling challenges and the interpretation of empirical observation data go hand in hand with the multi-scaling perspectives. In fact, the methods used for understanding or modelling crowd behaviour employ multiple-scale perspectives, in order to generate suitable mathematical structure describing crowd dynamics at each individual spatial scale. The specific scales at which the crowd can be defined are as follows:

Microscale Crowd structure and behaviour is identified by analysing each composite individual behaviour in the crowd. The state of each of individual behaviour is computed using features such as position in space and velocity. These are understood as time-dependent variables (see Fig. 2 for illustration).

Mesoscale Crowd structure and behaviour is understood through the dynamics of distinct patterns representing clusters of individuals who may share similar behaviour. Namely, this is represented by spatial cluster or group positions, collective velocities and kinetic energies. In this case the crowd system is represented statistically through a distribution function over the mesoscale-based states. This is illustrated in Fig. 3.

Macroscale Crowd structure and behaviour is assimilated as a continuum (or blob), where its dynamic state is described by average quantities such as density, central position, velocity and kinetic energy. These features are time- and space-dependent variables. They are statistical averages of the microscale states of individuals. This is illustrated with two distinct crowds as shown in Fig. 4.

Fig. 3 Mesoscale view of a crowd with a group of individuals sharing behaviour

Fig. 4 Macroscale view of a crowd with two distinct crowds (blobs)

1.1.2 Crowd Behaviour Considerations

Crowd behaviour detection requires the measurement and understanding of the dynamic motion or mechanics of the crowd at multiple scales. This may be highly dependent on the unfolding of contextual events and the nature of the recipient spatial environment in which crowd evolves with time. Using a pedestrian crossing as an example (see Fig. 4), the expected behaviour of the crowd is that they intend to cross safely to the other side of the road. In such context of spatial environment, human behaviour may be considered unusual if the crowd at any considered macro-, meso- or microscale deviates from what is expected. Thus, with this example, unusual behaviour could be the whole crowd trying to cross with the presence of cars. Equally, groups within the crowd may want to do such thing or indeed a single individual. But note that the so-called unusual behaviour may sometimes not be 'compromising to safety', but it is just unexpected to occur given the context of rules concerning pedestrian crossings. In much of the literature, unusualness is also defined as a statistical deviation from what is happening overall (the so-called 'usual' behaviour).

Below is a set of considered definitions of pedestrian behaviour with typically observed features for measurements and understanding:

- *Pedestrians taking detours or moving in a different walking direction to the main crowd, with the intention of taking the fastest route, than that of the crowd, in* *order to reach their specific desired destination. This is also not the shortest route spatially.*

- *Pedestrians in usual circumstances keep to individual optimal speeds, the value of which is normally distributed around a mean of 1.34* m/s.
- *Pedestrians, when they can, usually keep a certain distance from one another, as well as from pavements, walls and other obstacles. This distance gets smaller, with increasing pedestrians speed and/or density.*
- *Pedestrians' speeds can considerably increase in context of a perceived situation which may lead to compromising their safety and security. Their individual motions will appear random to almost unpredictable.*
- *At sufficiently high crowd densities, the motion of pedestrians is observed to be similar to that of fluid flows. In this case, concepts and associated features of fluid flow turbulent diffusion and advection, i.e. dispersion, can be adopted.*
- *Crowd behaviour primarily including motions is not due to immediate neighbours' interactions but often distant ones too. This is often caused by so-called behaviour propagation in the crowd.*

The above-mentioned crowd behaviours are viewed under the framework of a complex dynamic system, while simulation models use specific features to reproduce them realistically. In other words, the following capabilities of human behaviour in context of crowd systems should be reached while they are embedded in performing numerical models:

- **Ability to express a strategy**: Humans are capable of developing specific strategies related to their organisational ability depending on their own state and on that of the entities in their immediate vicinity. These can be expressed without the application of any principle imposed by the outer environment.
- **Heterogeneity**: Crowds, irrespective of their types, can be assumed to be heterogeneously distributed. This includes, in addition to different walking capabilities, the possible presence of leaders and the individual level of experience or prior knowledge.
- **Interactions:** These can involve individuals within the crowd connected to their immediate neighbours but also distant ones. In fact, crowd systems can be assumed to communicate at various spatial scales and may possibly choose different interaction paths, depending on the circumstances and spatial boundary conditions in which they could be.

Recall that there are conditions of natural groupings or clustering in crowds. This was illustrated in Fig. 3, when we introduced crowd mesoscale behaviour earlier. In this case, one also introduced the concepts of 'seed' and seed behaviour understanding. The seed may be considered at the leading individual in a grouping or cluster at mesoscale. The seed's behaviour is basically the closest to the group aggregate. Thus, the seed can be considered to influence the rest of the group and beyond. It is understood that a seed may be the originator or the source of behaviour which not only has the potential to lead a cluster and influence its motion, therefore behaviour, but also the triggering mechanism for the propagation of such behaviour across the crowd system in the macroscale.

2 Crowd Behaviour Detection and Modelling

2.1 From Motions to Behaviour Understanding

As was discussed earlier, crowds can be viewed as complex systems, where their behaviour is determined by the inner dynamics of the system, including those respective to the macro-, meso- and microscales. To infer a full description of the system's behaviour from these dynamics remains a challenging task. It is also worth noting that such description of system dynamics is often not fully achieved, due to limited sensor observations for measurements and/or the constraints due to regulations for openly experimenting on the spatial environment with crowds of interest. Further, the full analysis of different scales and various types of dynamics within the crowd system will require the use of different mathematical models. In an attempt to handle the complexities and ambiguities of the realm of crowd behaviour detection, research efforts which deal with the problem of crowd behaviour often settle for answering specific questions about crowd behaviour instead of offering a full description of crowd behaviour as a complete theory. Thus, some specific questions on crowd behaviour are considered in this chapter. These are listed below:

- *Is the crowd behaving in an unusual manner?*
- *Is the crowd showing signs of specific behaviours (for example, panic)?*
- *Is the crowd changing behaviour due to the actions of an individual or group of individuals?*
- *Is this change in behaviour propagating in the crowd?*

2.2 Measurement of Crowd

The question of 'how to measure a crowd' can have many different answers. The answer may depend on the type of information required and the level of granularity of interest which needs to be adopted. Here, the inner workings and state of a whole crowd is investigated. As a result, a set of features are defined for the crowd. These features are chosen with the aim of characterising the state and type of crowds in terms of human behaviour. Nevertheless, we will assume that the crowd is homogeneous in type. In this, while the micro-level motions within the crowd are observed and measured, the defined properties and features describe the overall crowd as one type of crowd. An example of such homogeneous crowd type can be a competitive marathon, where the crowd is composed of single individuals who all share the same goal. The homogeneity assumption will also hold for cases such as a shopping mall wherein there may be small groups of people as well as individuals while one assumes that they have the same goal of shopping in such environment. The method proposed here is motivated by physical analogies of thermodynamics and statistical mechanics, where the macroscopic properties of matter are derived from microscopic properties and states of the underlying molecular systems.

2.2.1 Crowd Analogies to Physical Systems

Various physical analogies and modelling approaches have been used in crowd and traffic modelling. A physical model requires a hypothetical structure of controlled parameters which need to be fine-tuned for simulating crowd dynamics that is in accord with experimental observations. In this section, some of the more popular modelling analogies in this domain are reviewed and evaluated. These include *(a) cellular automata, (b) social force model and (c) molecular fluid dynamics*.

(a) *Cellular automata (CA)* has been used to simulate crowd dynamics in situations such as evacuation [2–5]. In this, CAs evaluate the feasibility of different evacuation scenarios. It has also been shown that CA can simulate certain effects such as line formation in the crowd [6]. However, CA does not aim to capture all the microscopic dynamics but only that which is necessary to derive a specific macro effect.

(b) *Social force model* is another popular method for crowd simulation [7, 8]. It has also been used to detect points with high social friction within the crowd [3, 9]. In particular, Helbing et al. noted [9]:

> The motion of pedestrians can be described as if they would be subject to 'social forces'. These 'forces' are not directly exerted by the pedestrians' personal environment, but are a measure for the internal motivations of the individuals to perform certain actions (movements).

In essence, the social force model is based on a simple model wherein the individuals move according to their goals and environmental constraints. It is assumed that each individual in the crowd has a desired direction and velocity while seeking to keep a social distance from other members of the crowd as well as avoiding hitting walls. To calculate the social force model, an estimate of the individual goals is required. Other methods have been proposed to estimate the individual desired directions and velocities in a crowd [3, 9]. A bag-of-words method is used to select features from within the social force fields in consecutive frames. These bag-of-words features are subjected to further learning of unusualness detection in crowds using latent Dirichlet allocation (LDA) [9].

(c) *Molecular fluid dynamics* has also been investigated for modelling pedestrian motions. Henderson was the first to propose a gas-kinetic model for pedestrian flows [10]. Using this basis of a Boltzmann-like gas-kinetic model, Helbing [11, 12] developed a special theory for pedestrians, distinguishing between different groups within the crowd with different types of motions and goals. Moore et al. [13] argue against the gas-kinetic-based modelling of crowd for high-density crowds and note that for a high-density crowd, the behaviour appears to be liquid-like with interaction forces dominating the motion of pedestrians.

We propose to use analogies from thermodynamics and statistical mechanics for describing the state of crowds. While thermodynamics is concerned with heat and temperature and their relationship to energy and work at molecular matter levels, our major interest here is to derive macroscopic properties of crowds from statistical mechanics, in terms of the microscopic constituents of crowds. It is this conceptual link between the microscopic constituents of matter and its macroscopic properties that we needed to borrow while adapting thermodynamics and statistical mechanics principles to derive macroscopic features of crowds from their microscopic constituents. These are individuals within the crowd.

Following the above, we set up holistic features in a way that would enable us to describe and differentiate between different kinds of crowds and also different states of a crowd. As will be discussed in the next section, three parameters are postulated. These are *structure*, *energy* and *translation*. In this case, any crowd system type can be projected onto a point within the structure-energy-translation dimensional space. The aim is to achieve a good separation between different types of crowd in this three-dimensional space. These parameters are then used in a contextual crowd behavioural model which models the normative behaviour of crowds within well-defined situations. If there are discrepancies between the expected and the perceived behaviours, the behaviour is deemed to be unusual.

2.2.2 Crowd Representation with Its Holistic Features

We assume that a force keeps the members of the crowd together. The strength of connections between the members of the crowd will be referred to as *structure*. Irrespective of the strength of connections, the crowd may be in an excited state, *high energy*' or a calm state, *low energy*. This feature of the crowd simply refers to *energy*. We also consider that the crowd moves in space, while we refer to as *translation*. Figure 5 shows a representation of the *structure-energy-translation* crowd space.

Table 1 also illustrates a set of hypothetical examples of various types of crowds, while Fig. 5 shows where these reside in the *structure-energy-translation* crowd space.

Regarding the values of the structure parameter, it is worthy to note that a high structure score may denote one of the two underlying reasons: *(i)* a high social interaction between the members of the crowd (the members of the crowd maintain a pattern within a crowd), or *(ii)* a high value for structure may also be the result of an enforced structure by the environment (barriers, passages and doorways are examples of elements which can impose environmental structure). The structure

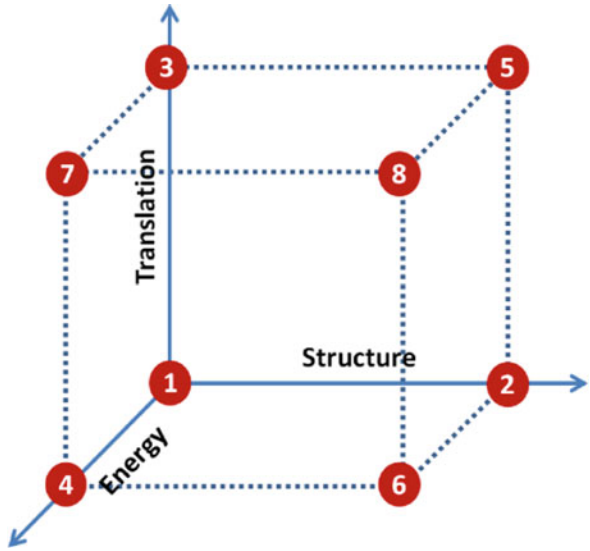


Fig. 5 Crowd space with hypothetical examples

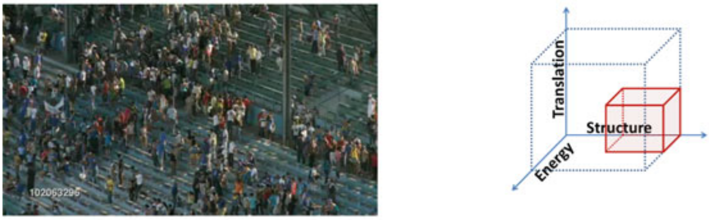
Table 1 Hypothetical examples of crowd states

1	A number of individuals walking slowly in different directions
2	Very unmotivated crowd at a less excited football game
3	Shopping centre at closing time
4	A group of panicking people locked in a room
5	People on an escalator
6	Crowd at a football match celebrating a winning goal
7	People escaping towards an exit
8	A bull run (when the bull arrives)

parameter only evaluates the level of structure in the crowd and does not differentiate between the pedestrian-imposed environment structures.

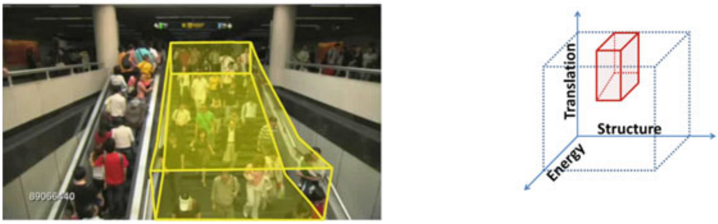
As shown in Fig. 5, a crowd may reside in any location in the *structure-energy-translation* crowd space. However, for any given situation, there would be an expectation of where the crowd should be, while a divergence from this expected, or desired, position may be a cause for alarm. Figure 6 shows sub-spaces of ‘usual’, or expected, crowd behaviour under various contextual situations and crowd types. By mapping the crowd into the *structure-energy-translation* crowd space and learning the limits of ‘usual’ behaviour, a crowd with unusual behaviour can be defined as

(a) *Spectators at a stadium while a football match is in progress. State 1 shows the low energy state when the crowd is motionless, while state 2 represents a high energy state. For example when the crowd is celebrating a winning goal.*

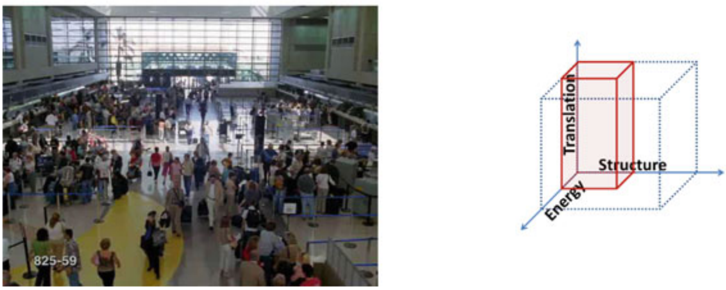


(b) *Spectators at a stadium (before and after a football match). Here the crowd has a non-zero translation.*

(c) *Walking crowd on an escalator with high Translation and Structure, but low Energy.*



(d) *Walking crowd on stairs with lower Structure, as each individual is moving with respective own speed, leading to higher Eneries, when compared to (c).*



(e) *Crowd at an airport main entrance hall. Low Structure is observed, with fluctuating values for Translation and low to medium Energy levels.*

Fig. 6 Various 'usual' behaviours in the *structure-energy-translation* crowd space. (**a**) Spectators at a stadium while a football match is in progress. State 1 shows the low energy state when the a crowd which does not fall within the limits of perceived ‘normality’ or strictly speaking ‘usualness’.

2.2.3 Approach

As mentioned before, we draw analogies from thermodynamics and statistical mechanics principles. The concept of crowd *energy* refers to its *internal energy*, while *translation* relates to crowd flow velocities. These velocities may be derived at various scales. Namely, at micro-, meso- or macroscales. As for the concept of crowd *structure*, it relates to the *entropy* of the states of a molecular system.

2.2.4 Translation Through Flow

As noted earlier, crowd flow can be derived at different scales. The most interesting of which is the one at mesoscale, as it concerns the flow of sub-groups within a crowd. Here, the term flow is used interchangeably with the term flow velocity. In fluid dynamics, flow velocity, v, is defined as

$$v = \frac{\dot{m}}{\rho . A} \tag{1}$$

where $\dot{m}$ denotes the fluid mass flow, ρ is its density and A is the flow cross-sectional area. With consideration of a sub-group within the crowd, its density ρ can be computed using the entire volume occupied by the sub-group. The number of individuals crossing cross-sectional planes of the crowd flow can be counted to find the mass flow $\dot{m}$.

In some circumstances a sub-group within the crowd can be represented by a ‘Gaussian blob’, of which its speed and direction can be denoted by the mean speed and direction of its constituents. This will be referred to as *translation*. Hence, *translation* is known as the measurement of how an entire sub-group (or whole group) travels in space, while flow measures the rate at which a mass of fluid crosses a plane.

Fig. 6 (continued) crowd is motionless, while state 2 represents a high energy state. For example, when the crowd is celebrating a winning goal. (**b**) Spectators at a stadium (before and after a football match). Here the crowd has a non-zero translation. (**c**) Walking crowd on an escalator with high translation and structure but low energy. (**d**) Walking crowd on stairs with lower structure, as each individual is moving with respective own speed, leading to higher energies, when compared to (**c**). (**e**) Crowd at an airport main entrance hall. Low structure is observed, with fluctuating values for translation and low to medium energy levels

2.2.5 Internal Kinetic Energy

The internal energy U of a crowd as a thermodynamic system can be used as a measure of how excited a crowd can be. It is defined as follows:

$$U = U_{\text{kinetic}} + U_{\text{potential}} \tag{2}$$

U_{kinetic} and $U_{\text{potential}}$ represent the *kinetic* and *potential* energies, respectively.

The kinetic energy U_{kinetic} is defined as follows:

$$U_{\text{kinetic}} = \frac{1}{2}\text{mv}^2 \tag{3}$$

where m and v represent the mass and internal flow velocity of a given sub-group within the crowd system.

As for the potential energy $U_{\text{potential}}$, its calculation is substantially complex to derive, particularly in context of crowd system thermodynamics. Its specifically relates to molecular systems which undergo thermodynamic phase transitions, where it is paramount to compute their potential energy.

However, some important pedestrian modelling approaches took inspiration from molecular systems theories with thermodynamic phase transitions [11, 12]. These postulate crowd potential energy as the ‘common sense’ of tasks pedestrians would take for reaching their expected destination. Nevertheless, such approach is not yet practical for us to implement in our experiments on crowd behaviour understanding. Therefore, we have not considered it in this study.

In the next section, we will particularly discuss the notion of entropy as an analogy to crowd structure. The computation of crowd *structure* while using analogous methods for calculating entropy is discussed with results presented.

2.2.6 Structure Through Entropy

Although initially defined within thermodynamics, the concept of entropy was generalised using Maxwell-Boltzmann classical statistical mechanics theory [14]. For example, entropy, S, is simply a measure of disorder in a molecular system:

$$S = -k_B \sum\nolimits_i p_i \ln p_i \tag{4}$$

where for a classic molecular system with a discrete set of microstates, p_i is the probability of occurrence for microstate i. k_B is the Boltzmann constant.

The same concept of entropy was also translated in 1948, under Shannon’s information theory in computer science and informatics [15]. Entropy, mostly denoted by H, is in this case a measure of uncertainties in random variables in data communication systems:

$$H = -\sum_i p_i \log_b p_i \tag{5}$$

Entropy is defined at a macroscopic level, where a given macroscopic state can have varying microscopic statistical realisations. The initial definition of entropy in statistical mechanics, $S = k_B \ln W$, connects entropy directly to the number of microstates, W, which corresponds to the macroscopic state of the given system.

Considering the states of matter, in classical terms which are solid, liquid or gas, the levels of entropy for these states can be intuitively understood. In a solid state, molecules oscillate in a vicinity of a fixed location, and the entropy is low. In a liquid state, molecules move freely but keep distances from one another, while the entropy is intermediate in values. However, in a gas state, molecules move randomly anywhere, while entropy increases to a higher level. Entropy, here, increases across these three matter states, with growing uncertainties on the location the constituting molecules of matter.

As noted above, *entropy* is really a measure of disorder, while in this case *structure* can be canonically described as a measure of *order*. For a normalised entropy in the range of [0, 1], structure and entropy are complementary and add up to unity. One of the challenges in evaluating the value of *structure* using the concept of *entropy* is that for each crowd example, only a sub-set of all possible microstates represents that macro-state is observed. Therefore, it is not possible to count the number of microstates or calculate their probabilities directly. An extra step is required to infer a model or a description for all the possible microstates using the observed microstates. Figure 7 shows a diagram of the required steps for calculating the entropy of crowd using the observed set of microstates.

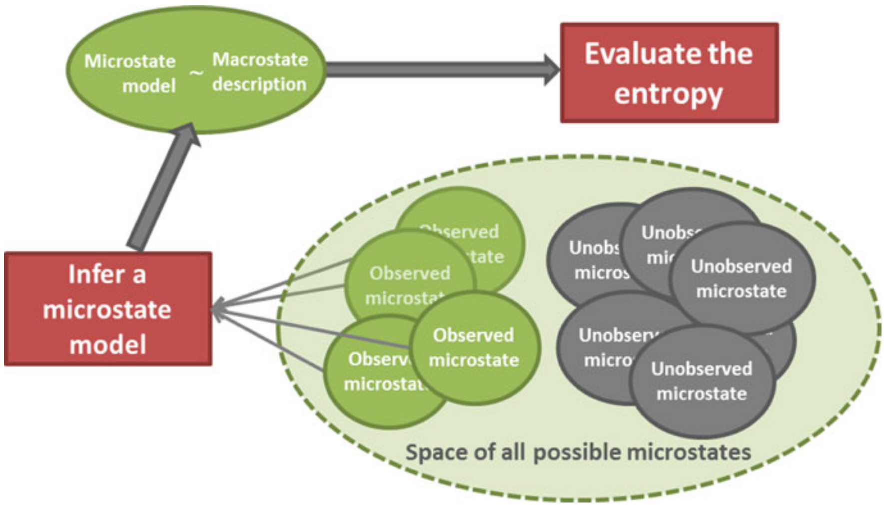


Fig. 7 Evaluating entropy through the observed microstates of a crowd

Table 2 Micro-space modelling parameters

Notation	Definition
N_f	Number of frames in the time window which is analysed
N_p	Average number of individuals in N_f frames
N_l	Number of spatial bins
l_i	ith spatial bin
Y	A discrete random variable defined on discrete sample space S_d
S_d	A numerical and discrete sample space of densities at a location
$(f_Y)_i$	Probability mass function for X at bin l_i

Fig. 8 Crowd density map

2.2.7 Calculating Entropy

Before discussing the model, it is useful to define our notations used in the calculation of *entropy*. These notations are listed in Table 2. Also, Fig. 8 illustrates the values for crowd density at the centre of each spatial bin.

2.2.8 Approach 1: Preserving the Density Pattern

The joint entropy of a population of N_p individuals scattered in N_l locations with probability mass function $(f_Y)_i$ is described as follows:

$$H\left(X_1, \cdots, X_{N_p}\right) = -\sum_{x_1 \in \mathcal{L}_X} \cdots \sum_{x_{N_p} \in \mathcal{L}_X} P\left(x_1, \cdots, x_{N_p}\right) \log\left[P\left(x_1, \ldots, x_{N_p}\right)\right] \quad (6)$$

where X_k is a triple $\left(x_k, \mathcal{L}_X, \mathcal{P}_{X_k}\right)$ and the outcome x is the value of a random variable, which takes on one of a set of possible values, $\mathcal{L}_X = \left\{l_1, l_2, \ldots, l_{N_l}\right\}$, having probabilities $\mathcal{P}_{X_k} = \left\{p_{k,1}, p_{k,2}, \ldots, p_{k,N_l}\right\}$, with $P(x_k = l_i) = p_{k,i}$. Here $\mathcal{P}_{X_k}$ and the joint probabilities, $P\left(x_1, \cdots, x_{N_p}\right)$, are unknown. The joint probabilities can be calculated using the probability mass functions $(f_Y)_i$. However, the computation cost is in the order of $O\left({N_l}^{N_p}\right)$. More efficient algorithms can reduce this computation cost. However, we argue against the validity of this approach, since it is prone to over-fitting the model to the sample set of observed microstates. Relaxing some of the conditions in this model may be favourable.

2.2.9 Approach 2: Preserving the Density Pattern with Independent Pedestrians

One of the conditions which can be relaxed in the first approach is the assumption of dependence between the positions of pedestrians. In the example below, it will be shown that although there is a reason to believe that these positions are in fact dependent, sufficient information is not available to understand their dependencies accurately and in an unbiased manner.

In support of the dependency argument, let us consider that people in a crowd system tend to keep distances from each other, known as personal space. Also depending on the relationships between the pedestrians, they may tend to further avoid other pedestrians or groupings. From a different point of view, consider the following example: A number of clusters of pedestrians are observed in different locations. There may be different causes for this effect. *Hypothesis A*: There might be some relationship between members of the crowd (the second person goes to the place where the first person randomly selected). In this, even if the initial selection for the person 1 was fully random with equal chances, due to the high correlation between the first and second persons, what is observed is an environment where a certain location seems very popular. However, equally probable is that the location itself is indeed popular and people cluster there for that reason (this will be called *Hypothesis B*). The point is that sufficient information is not given in favour of either *Hypothesis A or B* in the above example.

We propose then that when analysing crowd formation through a few correlated frames, the simpler model which can exhibit similar outcomes is more viable. In this model, the locations of pedestrians are considered to be independent. We hypothesise that a pattern is formed in the crowd if each individual is bounded by the same pattern. Also, when taking this approach, the calculation of *entropy* simplifies significantly.

Let $n_{i,j}$ be the number of times that individual j has been observed in bin l_i in N_f frames. The probability of selecting this bin, l_i, by individual j is

$$P\left(x_j = l_i\right) = \frac{n_{i,j}}{N_f} \tag{7}$$

Given that the location of individuals is considered as independent and that there is no differentiation between individuals, the probability of any individual selecting bin l_i is the same as any other. Thus, an estimate of the probability of selecting bin l_i, $P(x = l_i)$, can be given by

$$P\left(x=l_i\right)=\frac{\sum_{k=1}^{N_p} P\left(x_k=l_i\right)}{N_p}=\frac{\sum_{k=1}^{N_p} \frac{n_{i,k}}{N_f}}{N_p}=\frac{\sum_{k=1}^{N_p} n_{i,k}}{N_f N_p}=\frac{n_i}{N_f N_p} \tag{8}$$

where n_i is the sum of all density counts at bin l_i in N_f frames. Since the locations of individuals are independent of one another, the joint entropy of the crowd, $H\left(X_1, \cdots, X_{N_p}\right)$, simplifies to the following:

$$H\left(X_1, \cdots, X_{N_p}\right)=\sum_{k=1}^{N_p} H\left(X_k\right) \tag{9}$$

Note that the locations of all the individuals are based on the same location probabilities, $P(x = l_i)$.

Thus:

$$H\left(X_1\right)=H\left(X_2\right)=\cdots=H\left(X_{N_p}\right), \tag{10}$$

$$H\left(X_1, \ldots, X_{N_p}\right)=N_p H(X) \tag{11}$$

where X is a triple $(x, \mathcal{L}_X, \mathcal{P}_X)$, and the outcome x is the value of a random variable, which takes on one of a set of possible values, $\mathcal{L}_X = \left\{l_1, l_2, \ldots, l_{N_l}\right\}$, having probabilities $\mathcal{P}_X = \left\{p_1, p_2, \ldots, p_{N_l}\right\}$, with $P(x = l_i) = p_i$ as was defined in Eq. (5). The crowd entropy in Eq. (8) can be computed in linear time.

2.2.10 Pre-processing

Three pre-processing stages are required before the entropy can be computed. These are specified as follows:

2.2.11 Real-World Pedestrian Locations

The locations of pedestrians in an image have been subjected to projective transform. The real-world positions can be retrieved using the camera calibration matrix and head-height plane homography transforms.

2.2.12 Internal Position Estimation

The internal position of each pedestrian within the crowd, x_i, is also required. If the crowd is stationary, then the observed position, x_o, is equal to the internal position ($x_i = x_o$ iff $v_f = 0$). However if the crowd is moving with a flow velocity, v_f, the change in internal position in a time step dt can be calculated as

$$dx_i = dx_o - v_f \mathrm{d}t \tag{12}$$

where $v_f = \frac{\dot{m}}{\rho.A}$, $\dot{m}$ is the estimated mass flow, ρ is the mass density and A is the area. For a calibrated footage and given the density maps, A and ρ can be calculated. Given the tracks of pedestrians, the vertical and horizontal mass flows are estimated at two vertical and horizontal surface planes through the mid-point of the crowd's spatial space.

2.2.13 Internal Position Density Map

Once the internal positions of individuals are known, an internal density map can be created. Note that the size of the density map bins, w_{bin}, is a significant parameter in the calculation of entropy. In this, a too large bin will mask the very information that entropy is aiming to extract, while a too small bin will be prone to noise.

In addition to the above, entropy normalisation under the concept of specific entropy needs to be computed as follows:

2.2.14 Normalising Entropy

As well as the level of disorder in the crowd, the value of the crowd entropy depends on:

1. The number of individuals in the crowd
2. The extent of the crowd spatial area

2.2.15 Specific Entropy

Specific entropy is the entropy per unit of mass. Let each individual to have a unit of mass; the specific entropy, H_k, will be the entropy of one individual in this crowd:

$$H_k = H(X) \tag{13}$$

where X is a triple $(x, \mathcal{L}_{\mathrm{X}}, \mathcal{P}_X)$, as in Eq. 8.

2.2.16 Specific Entropy per Unit of Area

Entropy is maximised if $\mathcal{P}_X$ is uniform:

$H(X) \leq \log |\mathcal{L}_X|$ with equality iff $\forall i \in \{1, \cdots, N_l\}$ $p_i = \frac{1}{|\mathcal{L}_X|} = \frac{1}{N_l}$

It can be seen that the maximum value of entropy increases with the increase in the number of spatial bins, N_l. We borrow a concept from information theory called *redundancy*. Redundancy is a measure for the amount of wasted space when coding and transmitting data. The redundancy of X, $R(X)$, on alphabet $\mathcal{A}_X$ measures the fractional difference between $H(X)$ and its maximum possible value:

$$R(X) = 1 - \frac{H(X)}{\log \mid \mathcal{A}_X \mid} \tag{14}$$

Complementary to the concept of redundancy is *efficiency*, where the redundancy and efficiency of a code add up to one. Our notion of normalised specific entropy, h_k, is analogous to efficiency:

$$h_k = \frac{H_k}{\log N_l} \tag{15}$$

As noted, entropy is a measure of disorder, while structure can be described as a measure of order. For a normalised entropy in the range of $[0, 1]$, structure and entropy are complementary and add up to one. Let s_k be the normalised structure:

$$s_k = 1 - h_k \tag{16}$$

3 Experimental Results

Three crowd examples have been used in the experiments [16]. Experiment A shows a crowd of pedestrians climbing down a staircase. This motion of crowd is clearly unidirectional. This example depicts an indoor scene with artificial lighting, and the crowd is viewed from an oblique-frontal view. Similar crowds may be observed at a metro station or a stadium. Figure 9 shows one frame example of this crowd. This figure also shows three calibration planes. In this, the orange plane is the reference plane drawn manually. The blue plane and the yellow plane are the ground-level and the head-level planes, respectively, projected back to the image plane after calibration. The red circles show the position of the pedestrians' heads on the head-level plane. For this experiment, people's heads are labelled manually. Figure 10 shows the second crowd (Experiment B) which focuses on pedestrians on an escalator which is on the left-hand side of the same video footage. Here the pedestrians are standing still while the escalator carries them upwards. Finally, Fig. 11 (Experiment C) shows a larger crowd of people in an open indoor space with

Fig. 9 Crowd on stairs (Experiment A)

Fig. 10 Crowd on an escalator (Experiment B)

many pedestrians moving in different directions. This type of crowds may be found within airports or shopping malls and so forth. Following from the examples in Fig. 6, it is expected that (i) the crowd in Experiment B (Fig. 10) has the largest structure, since people are standing still; (ii) the crowd in Experiment A (Fig. 9) has a smaller structure than the crowd in Experiment B, but still larger than the crowd in Experiment C (Fig.11); and (iii) the smallest structure is envisaged for crowd in Experiment C.

Figure 12 shows the overall structure results from experiments A, B and C. The experiments were carried out for varying time window sizes (w_{tw}) and spatial bin widths (w_{bin}). Figure 12a shows the results, where a time window size of 5 s is used. In this, the values of structure are as expected:

$$s_k\left(X_{\text{exp}_\text{B}}\right) > s_k\left(X_{\text{exp}_\text{A}}\right) > s_k\left(X_{\text{exp}_\text{C}}\right)$$

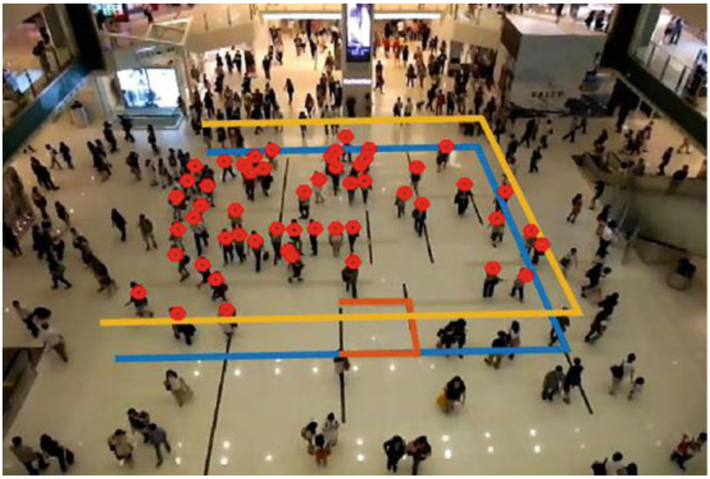


Fig. 11 Crowd in an open space (Experiment C)

Figure 12b shows the structure values for the same range of spatial bins, but the time window size has been reduced to 2 s. One can see that the order of structure values is still as expected. Nevertheless, while the separation between the various crowds is mostly achieved, the uncertainty on the value of structure increased considerably. Figure 12b, c also demonstrates the effects of spatial bin-size variations. The spatial bins in the range of 0.01 m $\leq w_{\text{bin}} \leq$ 0.6 m with a time window size of 2 s are investigated in these two graphs. One notes that the smallest bin size does not offer a good separation between crowds, while at the largest bin size of 0.6 m, all crowds show the same structure values. The best separation is achieved for bin sizes between 0.04 m and 0.1 m. Although as mentioned a time window of 5 s offers a much better separation, it must be noted that due to observing a non-stationary crowd with a stationary camera, it is possible that the crowd or the section of the crowd which is being analysed would move beyond the camera's field of view. As a consequence, the results for Experiment B when analysed with a 5-second time window may be considered as less reliable.

4 Ongoing Research and Future Perspectives

In our subsequent works, which we conducted recently, we have also looked into other possibilities which may provide an estimate for the structure of the crowd and compare the respective performances of these approaches with our proposed method over a larger set of crowd conditions. One of the methods which we have examined is that of Zhou et al. using the concept of 'crowd collectiveness' [17, 18]. With it we have been able to track individuals and groups in crowd-associated confined spaces

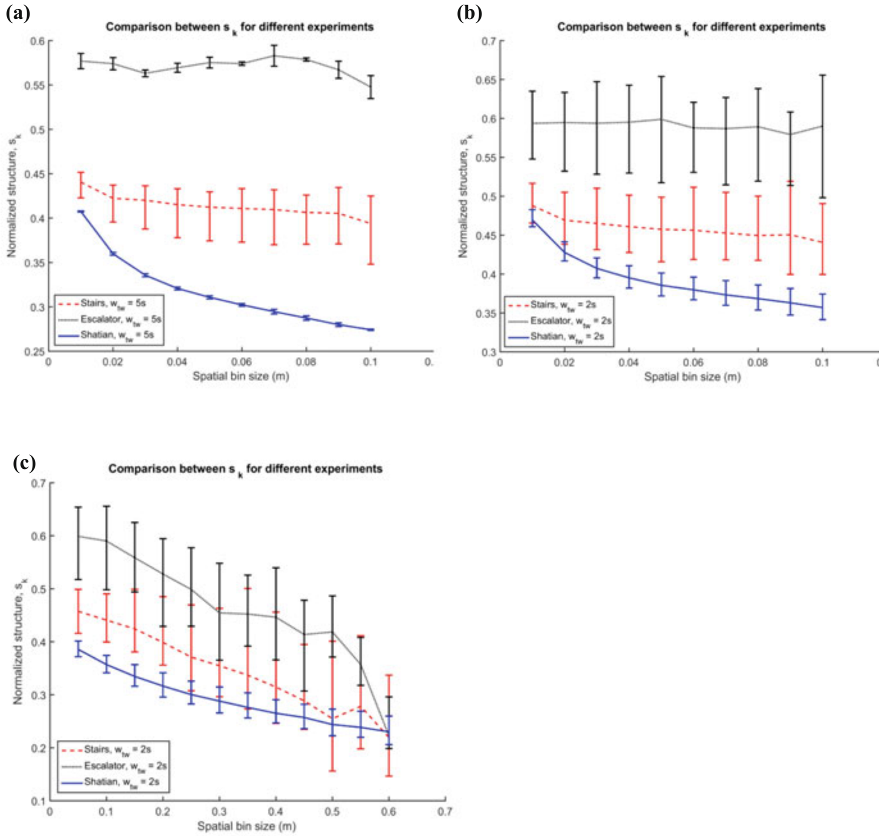


Fig. 12 Experiments with normalised specific entropy. (**a**) Experiments with a 5-second time windows ($w_{tw} = 5$s). (**b**) Experiments with a 2-second time windows ($w_{tw} = 2$s). (**c**) Experiments with larger spatial bins (0.5 m $\leq w_{\text{bin}} \leq 0.6$ m)

such as stadium arenas and in context of the event expected activities [19–21]. The unusualness of groups' behaviour is detected accordingly to provide an operational approach to security practitioners to respond in a scalable way. In this experiment, groups panic at a segment of the stadium arena and run towards the pitch. This is detected critically in time and produces an alert for security to focus on leading such type of distressed crowd to safety, as shown in Fig. 13.

We have further investigated on the actual mechanics of behaviour propagation in a crowd in recent years, particularly on behaviour which originates from a so-called seed which represents an individual or indeed a group of people behaving within the crowd. This is indeed of great importance to understand and capture trends in it so that we could develop a forecasting capability of behaviour. Although this research work is at its early stages while it is being conducted in the most recently launched S4AllCities research project [22, 23], it is showing us some

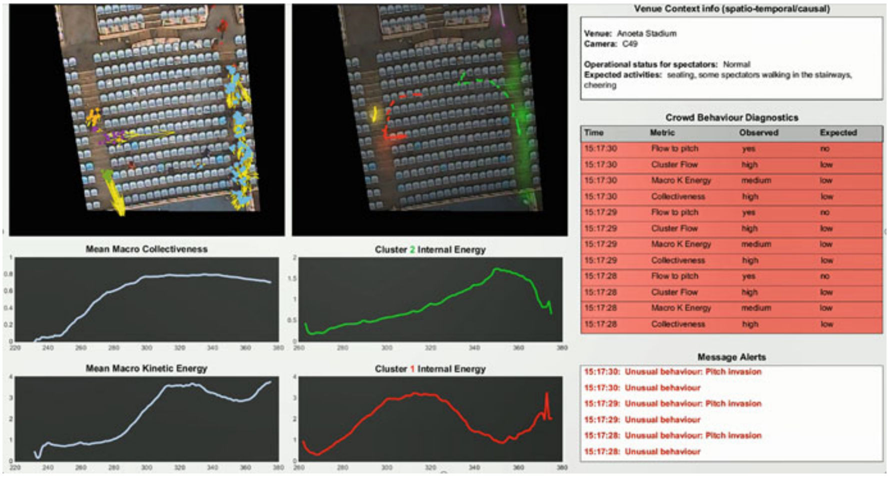


Fig. 13 Group panic behaviour detection in a stadium arena

encouraging findings. Namely, we are obtaining stable tracking of trajectories of individuals as well as groups, where we could derive their parametric functions. These will lead us onto developing data-driven models which will predict trends of such trajectories in future time frames. These of course will be derived with growing computed uncertainties downstream in time and space. We are therefore planning to computationally correcting these trends once new observation measurements are obtained in time in order to reduce and control such uncertainties, leading to a much improved learning process for understanding intentional behaviour in the near future.

Acknowledgements The authors are very grateful to all of the eVACUATE research partners, particularly Drs. Zlatko Zlatev, Gianluca Correndo and Lee Middleton, for their collaboration, while we initiated the research work at the University of Southampton. We are also thankful to Professor Nicola Bellomo for his important research discussions on the fundamental modelling of crowd when we initiated this research under the eVACUATE project in 2013. Our research work was partly supported by the European Commission, initially under H2020 Grant Number 313361 in the eVACUATE project (2013–2017) and more recently under H2020 Grant Number 883522 of the ongoing S4AllCities project (2020–2023).

References

1. Evacuate (2013–2017). A holistic, scenario independent, situation awareness and guidance system for sustaining the active evacuation route for large crowds. H2020 project ID number 313361. http://www.evacuate.eu/
2. Parisi, D. R. and Dorso, C. O. (2005). *Microscopic dynamics of pedestrian evacuation.* Physica A, Vol. 354, pp. 606–618.

3. Helbing, D. and Molnar, P. (1995). *Social force model for pedestrian dynamics.* Physical Review E, Vol. 51. 5, pp. 4282–4286
4. Burstedde, C., et al., et al. (2001). *Simulation of pedestrian dynamics using a two-dimensional cellular automaton.* Physica A, Vol. 295, pp. 507–525.
5. Kirchner, A., Nishinari, K. and Schadschneider, A. (2003). *Friction effects and clogging in a cellular automaton model for pedestrian dynamics.* Physical Review E, Vol. 67(5), pp. 056122–056122
6. Heneina, C. M. and White, T. (2007). *Macroscopic effects of microscopic forces between agents in crowd models.* Physica, Vol. 373, pp. 694–712.
7. Guo, R.Y. and Huang, H.J. (2008). *A mobile lattice gas model for simulating pedestrian evacuation.* Physica A, Vol. 387, pp. 580–586.
8. Weng, W. G., et al., et al. (2006). *Cellular automaton simulation of pedestrian counter flow with different walk velocities.* Physical Review E, Vol. 74(3), pp. 036102–036102
9. Helbing, D., Farkas, I. and Vicsek, T. (2000). *Simulating dynamical features of escape panic.* Nature, Vol. 407, pp. 487–490.
10. Mehran, R., Oyama, A. and Shah, M. (2009). *Abnormal Crowd Behavior Detection using Social Force Model. IEEE Conference on Computer Vision and Pattern recognition (*CVPR), 2009.pp. 935–942
11. Henderson, L. F. (1971). *The Statistics of Crowd Fluids.* Nature, Vol. 229, pp. 381 - 383.
12. Helbing, D. (1992). *A Fluid-Dynamic Model for the Movement of Pedestrians*. Complex Systems, Vol. 6, pp. 391–415.
13. Moore, B. E., et al., et al. (2011). *Visual Crowd Surveillance Through a Hydrodynamics Lens.* Communications of the ACM, Vol. 54, pp. 64–73.
14. Tolman, R. C. (1979). *The principles of statistical mechanics*. Courier Corporation.
15. C.E. Shannon. (1948). "A Mathematical Theory of Communication", *Bell System Technical J.*, vol. 27, pp. 379–423, July and Oct. 1948.
16. Middleton, L., Zlatev, Z., Sabeur, Z.A, Arbab-Zavar, B., Makantasis, K., Karatzalos, K., Christopoulou, I. and Bellomo, N. (2015). D3.4: Multi-scale behaviour recognition technical specification v2.0. Evacuate project deliverable. H2020 project ID number 313361. http://www.evacuate.eu/
17. Zhou, B., Tang, X., & Wang, X. (2012). Coherent filtering: Detecting coherent motions from crowd clutters. In *European Conference on Computer Vision*, pp. 857–871. Springer, Berlin, Heidelberg.
18. Zhou, B., Tang, X., & Wang, X. (2013). Measuring crowd collectiveness. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3049–3056.
19. Correndo, G., Arbab-Zavar, B., Zlatev, Z., Sabeur, Z.A. (2015). Context ontology modelling for improving situation awareness and crowd evacuation from confined spaces. In: Environmental Software Systems. Infrastructures, Services and Applications, pp. 407–416. Springer, New York.
20. Sabeur, Z.A., Doulamis, N., Middleton, L., Arbab-Zavar, B., Correndo, G., Amditis, A. (2015). Multi-modal computer vision for the detection of multi-scale crowd physical motions and behavior in confined spaces. In: Advances in Visual Computing, pp. 162–173. Springer, New York.
21. Arbab-Zavar, B. and Sabeur, Z.A. (2020). Multi-scale crowd feature detection using vision sensing and statistical mechanics principles. Machine Vision and Applications, 31(4), 1–16.
22. S4AllCities (2020–2023). Safe and Secure Smart Spaces for all Cities H2020 project ID number 883522. https://www.s4allcities.eu/project.
23. Sabeur Z., Angelopoulos C.M., Collick L., Chechina N., Cetinkaya D., Bruno A. (2021) Advanced Cyber and Physical Situation Awareness in Urban Smart Spaces. In: Ayaz H., Asgher U., Paletta L. (eds) Advances in Neuroergonomics and Cognitive Engineering. AHFE 2021. Lecture Notes in Networks and Systems, vol 259. pp. 428–441. Springer, Cham. https://doi.org/10.1007/978-3-030-80285-1_50

Applications of Crowd Dynamic Models: Feature Analysis and Process Optimization

Liang Li, Hong Liu, Yanbin Han, Guijuan Zhang, and Dianjie Lu

Abstract Recently, researchers studying crowd dynamics models have attempted to build a universal system for informing pedestrian traffic management. To achieve this goal, it is necessary to improve the research conducted on specific and detailed problems. In this chapter, we review our research work concerning two problems related to evacuation management: features analysis and process optimization. By summarizing previous studies and relevant literature, we will discuss the application of crowd dynamics models for solving the above-mentioned problems.

1 Introduction

Crowd dynamics models have developed rapidly [1, 3, 4, 35]. Owing to their advantages, such as easy reproducibility and verifiability, these models have been widely used in different applications. Applications of crowd dynamics models have diverse characteristics; they can be used to not only simulate crowd movements but also form a theoretical basis for other studies [50]. This chapter discusses the application of crowd dynamics models to evacuation management in detail.

Research on evacuation management has received significant attention and witnessed many achievements [5]; however, many of these results suffer from limitations in practical applications, owing to evacuation management's unique characteristics. For example, there is little scope for error in managing crowd evacuation, and few emergencies can be dealt with using the same approach. Thus, more comprehensive evacuation-management techniques are required. In particular, focusing on specific aspects has become a requirement for successfully conducting such research. Herein, we focus on feature analysis and process optimization.

The evacuation process has various features, including individual and group characteristics such as velocity, density, and self-organization phenomena. Given that

L. Li (✉) · H. Liu · Y. Han · G. Zhang · D. Lu
Shandong Normal University, Jinan, China
e-mail: zhangguijuan@sdnu.edu.cn; ludianjie@sdnu.edu.cn

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_4

evacuation in different situations is decided by diverse key features, specific-feature analysis is important for modeling crowd dynamics and evacuation management. However, feature analysis faces many challenges arising from the complexity of crowd motion. First, because most features are not independent, the effects of different characteristics on the target feature must be considered. Second, it is necessary to ensure that the features' mathematical expressions are correct. Finally, efficiency is also a requirement for analysis in evacuation management. Crowd dynamics models have gradually been reported to have applications in feature analysis. Owing to limited research on pedestrian behavior, the existing models usually focus on setting up the mathematical expression for one or a few features. Therefore, feature analysis can take advantage of mathematical expressions in the relevant crowd dynamics models, thus improving the analytical efficiency. Herein, we review our previous work to discuss the application of crowd dynamics models to feature analysis.

The disordered movement of pedestrians is the primary cause for direct and indirect dangers, particularly in emergencies such as fires or epidemics [45, 46]. Thus, the major objective of evacuation management is to optimize the crowd movement process to prevent disordered movement. As evacuation is a complex process, optimization can be achieved using various approaches; for example, optimization of the movement or emotional state can each improve the effectiveness of evacuation management. By contrast, different process optimizations may suffer similar problems. The main optimization problem is how to express crowd evacuation as an optimizable process. Moreover, combining optimization approaches such as intelligent algorithms with the understanding of crowd dynamics is another challenge. To further illustrate the aforementioned problems, we review some concrete work on process optimization in this chapter.

This review is mainly based on our previous papers [49–52, 87, 88] and is organized as follows: Sect. 2 presents the research on the analysis and alleviation of crowd congestion; Sect. 3 introduces different types of process optimizations used evacuation management; and Sect. 4 presents the conclusions.

2 Congestion Analysis and Alleviation for Managing Crowds During Evacuations

Congestion is a common phenomenon caused by an increase in the pedestrian population within a limited space. During emergencies, congestion can not only reduce the evacuation efficiency but also lead to dangerous situations; hence, many studies have considered congestion as an important feature of crowd dynamics [32, 33, 37, 47, 50]. In this chapter, we use our previous study on the analysis and alleviation of congestion as an example to illustrate research problems in feature analysis.

2.1 Congestion Analysis

Let us begin by briefly reviewing studies on congestion analysis. Many approaches have been developed to improve the effects of congestion analysis, and crowd density is a commonly studied feature [15, 48, 53]. In methods focusing on crowd density, congestion is typically used to distinguish between a crowded situation and a normal one; however, the crowd density is not the only relevant feature in establishing that one dense crowd is more congested than another [23]. A high level of crowd density does not always indicate that a situation is dangerous [29].

The fundamental diagram is widely used in the transportation theory; its applicability extends beyond the evaluation of pedestrian and vehicular flows. Nevertheless, it is commonly accepted that the validity of this diagram is restricted to a uniform steady flow, making it useless for analyzing crowds with complex interactions [85].

Reference [36] proposed the concept of crowd pressure to identify locations responsible for crowd turbulence and calculated crowd congestion by multiplying the crowd pressure by the local crowd density. Although this approach may be suitable for evaluating congestion in extremely crowded and dangerous situations, little is known about how crowd pressure behaves at low-to-medium crowd densities. As the crowd pressure reflects the variance in the macroscopic velocity, there are doubts as to whether this approach is appropriate for analyzing congestion in dense crowds with a low pedestrian velocity.

The congestion-level approach was proposed in [23] to analyze the congestion and intrinsic risks in crowds. This approach is inspired by the crowd pressure concept and defined by the velocity vector field obtained by analyzing the crowd motion trajectory to determine a congestion threshold. Compared with other approaches, the congestion level can better represent the crowd oscillation by varying the velocity direction and defining the danger region in a crowd.

Reference [29] designed a controlled experiment to analyze pedestrian dynamics, particularly its kinetic stress in the situations where swarms gather. The competitiveness of the pedestrians was considered to be the main cause for congestion, whereas the kinetic stress was used to characterize pedestrian dynamics. The results reported in [29] demonstrated that such a stress should be considered in congestion analysis.

In [90], it was shown that PDE-based models may be unable to define the congestion of dense crowds. A catastrophe model was developed to illustrate the congestion mechanism, demonstrating that accidents occur during the evacuation process. This model allows for analyses of crowd congestion without considering location or time.

Automatic detection of congestion was addressed in [85] by analyzing the behaviors of individual pedestrians. In particular, this study considered the direction weight calculated as the angle between two major directions (right and left) to reflect the lateral oscillation of people's upper bodies.

To study the press transfer in dense crowds, we presented an approach to analyzing the congestion of individual pedestrians [50], wherein the influence of different movement features on congestion is defined as different movement constraints. Although this approach can distinguish the congestion between different pedestrians, it still experiences many limitations.

The aforementioned studies show that research on congestion analysis has gradually developed with a focus on several key problems. The first problem is determining what the modeling target for congestion should be. In the density and fundamental approaches, this target is usually the entire region or crowd; in the crowd pressure and congestion-level approaches, the target is parts of the crowd; and in our approach, the target is individual pedestrians. Thus, the modeling target has gradually been refined from the global to local entities and then to the individual entity. Given that emergency situations are diversifying, the target of influence of congestion will need further refinement. For example, because of the requirements of epidemic prevention, the main feature influencing congestion may become the social-distancing requirements between pedestrians, as seen in Fig. 1.

The second problem is determining how many features in the crowd's movement process can influence congestion. According to the congestion analysis approach, it is obvious that refinement of the modeling target will increase the relevant features. At first, only the local density is considered; then, other features such as the flow velocity of the crowd, the velocity direction, and the oscillation must be added to the congestion analysis. In our approach, we have further considered the influence of press transfer; therefore, if we consider the congestion analysis as an equation-solving process, then the number of features continues to increase.

Here, a derivative problem should also be discussed: how to distinguish or weigh the influence of different features on congestion? In existing approaches (such as our method), it is common to confirm the feature that has the most influence; the remaining features can then be simply parameterized. As a result, there is a lack of comparison between different features. In our opinion, the weighting of other characteristics, influence on the target feature will become a major topic for future research.

The third problem is how to define the influence of other features on congestion. According to the existing approaches, there are two methods for representing this influence: direct and indirect. As shown in Fig. 2a, the first directly uses features to represent the congestion. For example, the final analysis of congestion in the congestion-level approach follows eq. (1), [23].

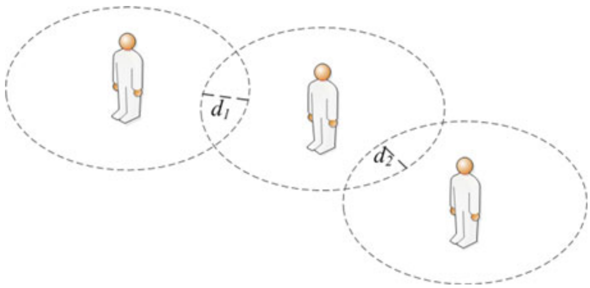


Fig. 1 To adhere to the requirements of epidemic prevention, the main factor influencing congestion may become the overlap between social-distancing zones

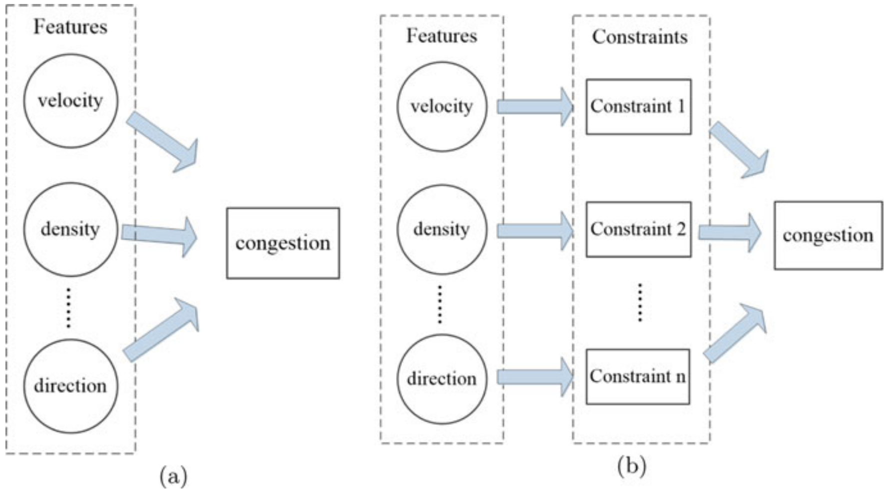


Fig. 2 Different methods in representing the influence of other features on congestion. (**a**): the direct method and (**b**): the indirect method

$$C_d = C_l \cdot \rho, \quad (1)$$

where C_d is the crowd danger (congestion), ρ is the local density, and C_l is a comprehensive variable.

In the indirect approach, different features are first summarized as results of crowd movement; then, the congestion is analyzed from the whole movement system. As shown in Fig. 2b, in our approach, different features are defined as corresponding constraints influencing the crowd motion, and the congestion is calculated according to these constraints.

Given that the development of congestion analysis is still in its early stages, there is no obvious evidence that proves one approach is superior to the other. The easily applicable direct approach is widely used in the existing congestion analysis; however, this approach may overestimate the influence of the main features, and its parameterization depends strongly dependent upon the quality of the experiments. Compared with the direct approach, the indirect approach is more complex because the influence of different features on crowd movement must be considered first, for which it is usually necessary to refer to theories in other fields such as analytical mechanics. As a result, the indirect approach may be more logical, which complies with the requirements of systematic construction in crowd dynamics.

The application of constraints for pedestrian movement is the main topic in our congestion analysis research. Given that crowd movement is an objective mechanical system, it is reasonable to apply constraints to this system, and this topic has already been studied by many researchers [9]. Here, we focus on the problem of how to effectively confirm the constraints for different features. There are two main requirements when confirming movement constraints; the first is that the constraints should be able to accurately represent the target features, for which an adequate understanding of crowd dynamics is required. The second is that the constraints should be described in an effective mathematical expression to ensure that the logic is reasonable.

Herein, we apply the existing crowd dynamics models to confirm the movement constraints [50]. As these models (which we assume to be built with a sufficiently high quality) are usually built according to deep research on some features or phenomena, their mathematical expressions are also convenient for confirming constraints. The concrete approach is to employ equations that define the target features, which we call local simplification. For example, according to the Hughes model, the macroscopic velocity of a pedestrian is calculated as follows:

$$v(\rho) = C \cdot \left(\frac{\rho_{trans} \cdot \rho_{crit}}{\rho_{max} - \rho_{crit}} \right)^{1/2} \cdot \frac{(\rho_{max} - \rho)^{1/2}}{\rho}, \quad (2)$$

and their microscopic velocity is defined as follows:

$$v_i = f_i(\rho) = \beta_i \cdot (A - B\rho), \quad (3)$$

where C, ρ_{trans}, ρ_{crit}, and ρ_{max} are constants. ρ_{trans}, ρ_{crit}, and ρ_{max} are set to distinguish the different congestion situations of the crowd; and A and B are constants. The detailed parameters are found in [38, 39]. The eq.2 and eq.3 describe the influence of the crowd on pedestrians' movements, although they still have their own velocity. Furthermore, these two velocities are defined by density. Therefore, it is possible to use the aforementioned equations as constraints on the pedestrian velocity. Although local simplification remains a crude method, we believe that it shows a reasonable direction for taking advantage of the existing crowd dynamics models.

The last problem is how to summarize the influences of different features. The problem is that it is necessary to formulate a reasonable theory to solve constraints; as the crowd movement exhibits many similar characteristics with particle systems, Lagrangian mechanics may be used to analyze the constraints of this system. The aforementioned opinion has already been considered by many researchers [25]. Herein, the constraints are also analyzed according to Lagrangian mechanics, and the relevant details can be found in [50].

Here, we also focus on a derivative problem: what is the requirement on crowd dynamics for the Lagrangian mechanics? The application of Lagrangian mechanics is based on the condition that some features of crowd dynamics are in a critical state during the movement process. In our approach, the equilibrium state (which is caused by the coordination movement of pedestrians) is considered to be the critical feature. The equilibrium state also provides many conveniences for the previous research work, such as macroscopic movement modeling. However, researchers have already found that the equilibrium state will disappear with the decrease of crowd density [5], suggesting that our approach cannot analyze the congestion of a crowd with low density. Therefore, confirming the critical features of a low-density crowd is a challenging and necessary problem for congestion analysis.

2.2 Congestion Alleviation

Given the negative influence of congestion, it is necessary to determine appropriate methods to alleviate it. Here, we focus upon congestion of a crowd that has already gathered, such as pedestrians at a bottleneck. In such situations, pedestrians are usually panicked and cannot be guided to change their target exit or to keep order. Therefore, our aim is to determine feasible objective methods for alleviating crowd congestion.

As congestion is caused by crowd motion, it can be alleviated by adjusting pedestrians' movement. Two main problems faced during congestion alleviation: the first is how to adjust pedestrians' movement. The adjustment approach should be acceptable by pedestrians as pedestrians are in a panic, and its adjustment on pedestrians should be stable. The second problem is how to confirm the relationship between the microcosmic pedestrian movement and the macrocosmic crowd congestion. As shown in Fig. 3, the adjustment of pedestrians' movement will change the crowd congestion via an intermediate method. In our previous work, we provided a feasible approach for alleviating crowd congestion [49]; specifically, we show that pedestrian movement can be adjusted by setting obstacles. The relationship is confirmed by examining an arch formation, which is a typical self-organization phenomenon.

Let us review studies on the setting of obstacles and arch formation. Many researchers have found that setting obstacles in a scene is an effective means of adjusting crowd movement. In panic situations, it is possible to increase the outflow by appropriately placing obstacles in front of the exit [40]; in a follow-up work, it was found that obstacles may increase pedestrian flow by 30% [36]. Reference [78]

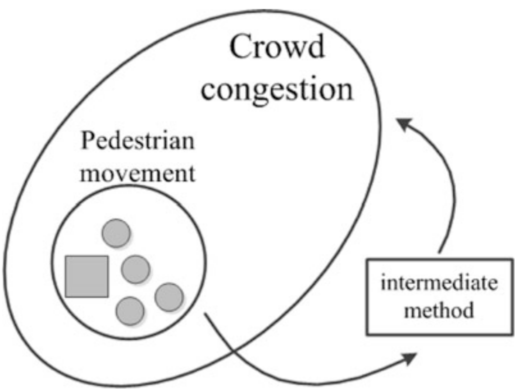


Fig. 3 An intermediate method is necessary for the connection of the pedestrians' movement and crowd congestion

reported that pedestrians' mean traveling time was reduced by 25% in an experiment when an obstacle was arranged in their path. However, obstacles with inappropriate configurations negatively affect evacuation efficiency; for example, [89] simulated an evacuation and found that placing obstacles symmetrically near the exit door may be harmful for evacuation. Therefore, the configuration of obstacles should be properly adjusted to obtain optimal results. Several studies have already been conducted to determine the proper configurations of obstacles. Reference [42] used a genetic algorithm to provide an obstacle layout design; however, they only considered obstacles structured like pillars. Other differently shaped obstacles, such as thin, flat panels, can also enhance the outflow efficiency. Reference [27] proposed that placing an obstacle in a panic situation might prevent congestion near the exit by absorbing pressure; consequently, the clogging effects are transferred to an early stage. Reference [16] introduced a new modeling technique that guarantees both the impermeability and opacity of the obstacles; their simulation results showed that the model can reduce the evacuation time from a room by adding multiple optimally placed and shaped obstacles to the walking area.

Our approach to addressing congestion is based on arch formation; therefore, we present a brief review of the research on this special self-organization phenomenon here. The arch-shaped structures (or arching phenomena) at bottlenecks are typically found in research fields such as traffic, architecture, granular flow through a hopper, and escape evacuation [68]. In some situations, pedestrian crowds exhibit collective phenomena similar to those observed in granular materials (for example, evacuation flows at bottlenecks) [17, 36, 54]. Herein, the arching effect in crowd evacuation is called arch formation. Such a formation is difficult to avoid when panicked pedestrians gather at an exit (Fig. 4). Arches may be formed and broken repeatedly, and this structure decreases the flow rate and increases congestion. Therefore, many studies have been conducted on avoiding arch formation. Reference [68] discussed the obstacle effect from the viewpoint of arch formation using the SFM; they discussed the possible physical mechanism behind the obstacle's effect on a pedestrian system and showed that it could take three forms: (i) influencing the space around the arch formation; (ii) shifting the center of the arch formation; and

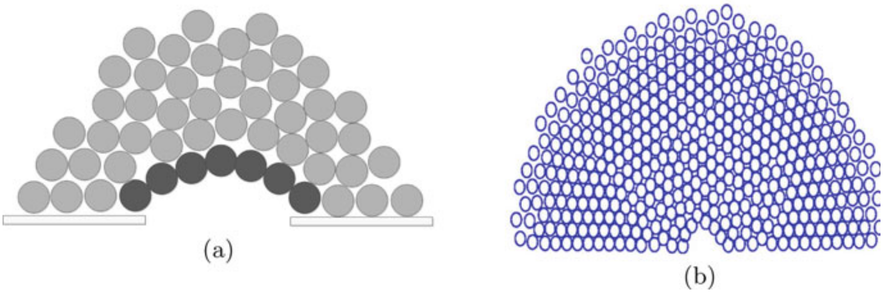


Fig. 4 Arch formation in a dense crowd. (**a**) An arch formation, which comprises pedestrians shown in dark gray. (**b**) A snapshot of a crowd in the simulation

(iii) distorting the arch formation. Reference [54] proposed a simple microscopic model for arch formation at bottlenecks; the dynamics of pedestrians in front of a bottleneck were described by a one-dimensional stochastic cellular automaton on a semicircular geometry. This model predicted the existence of a critical bottleneck size for a continuous evacuation flow; Reference [71] analyzed two macroscopic crowd dynamics models and studied the evacuation of pedestrians from a room with a narrow exit. In their simulations, the density profiles of the crowd exhibited a congestion situation with arch formation.

Herein, congestion alleviation is based on analysis of the arch formation. When obstacles are set in the region in which the arch is formed, they will change the movement of pedestrians who pass by. Then, the formation position and size of the arch will also be changed. According to previous research, crowd congestion will vary along with arch formation. Therefore, this phenomenon connects congestion with the movement of pedestrians.

Although the aforementioned approach can alleviate crowd congestion, it has many limitations. First, it can only be applied in a limited bottleneck situation wherein pedestrians gather at the exit; second, it depends hugely on the effects of congestion analysis. When adjusting arch formation, it is necessary to calculate the pressure acting on concrete pedestrians; thus, the aforementioned approach may be seen as a simple application of congestion analysis.

After reviewing the research on congestion alleviation, we discuss two hypotheses concerning self-organization phenomena. As phenomena such as arch formation can connect pedestrian and crowd motion, it can be considered on the mesoscopic scale. Research on mesoscopic models is currently an important avenue of crowd dynamics modeling; if such models can provide concrete definitions of some self-organization phenomena, they would be extremely helpful for feature analysis.

Herein, arch formation analysis is based on the theory for the similar phenomena observed in the research on granular motions. According to this theory, when a granular arch is formed, it will finally become a fixed structure, namely the reasonable arch axis [43]. The reasonable arch axis represents the balance of pressure. Although the arch formation in a crowd cannot always maintain its structure, it has many similar characteristics to the granular arch. In our opinion, arch formation can also be seen as an indication of the balance of crowd movement, which is similar to the equilibrium state of the crowd. Therefore, it is possible to consider arch formation, which is a self-organization phenomenon, as a critical feature. Thus, the question becomes can we always observe some self-organization phenomena that exist in a critical state during the crowd movement process? We believe that doing so will provide a feasible direction for confirming the critical features in research on a low-density crowd.

3 Process Optimization in Evacuation Management

Optimizing the pedestrian movement process is essential in evacuation management. Such research aims to avoid the danger caused by the gathering of panicked pedestrians. Approaches for process optimization are usually preventive methods such as path planning. Unlike reactive approaches such as congestion alleviation, the preventive approach can be achieved by many subjective methods that can take advantage of pedestrians' rational behavior; here, the critical problem is confirming guidance approaches. In this section, we optimize the evacuation process in two ways: the first is to optimize the target exit and path to improve the evacuation efficiency; the second is to control the propagation of the panic caused by the emergency. Although the aforementioned approaches have different optimization algorithms, both of their applications are based on concrete crowd dynamics.

3.1 Group-Based Approaches for Path Planning

Confirming the target exit for various pedestrians is an important problem for path planning. In addition to the distance, many other factors can influence the exit selection, such as the group phenomenon. Groups commonly form on the basis of pedestrians' relationships and cognition; although group formation will increase the difficulty of exit selection, it can also improve the effectiveness and performability of path planning. Herein, we try to study and apply the group phenomena to optimize path planning for crowd evacuation.

3.1.1 Group-Based Approach Without Navigation

The groups mentioned here are formed by pedestrians who gather and move to the same exit. In reality, when an emergency occurs, individuals usually attempt to stay closer to their friends and family, forming small, self-organized groups. Most will select the same exit and follow the leader who can reach it first; therefore, the grouping behavior is a common phenomenon in evacuation process.

Groups have a complex influence on crowd evacuation; on the one hand, the formation of a group can decrease the chaos level of a crowd and increase the evacuation efficiency; this is because, when pedestrians try to gather as a group, they will have a definite target. On the other hand, if the path and exit are not appropriately selected for a group, it is still possible to increase the evacuation time, leading to congestion. Thus, it is necessary to research the grouping strategy for evacuation management.

Group behavior is an important phenomenon in path planning; thus, many studies have focused on the influence of group behavior on the evacuation process [21, 41, 66, 86]. Crowds often comprise many small social groups based on friendship or kinship [58]. Under normal circumstances, both models and experience show that pedestrian crowds are self-organized [26]; when a social group meets severe environmental threats, they usually show strong emotional behavior and prefer to move together [86]. Individuals within the crowd can also form groups to cope with the emergency, even if they are not socially connected [66]. Small social groups based on kinship or friendships are ubiquitous in human crowds [58]; therefore, it is necessary to study the interaction between social groups and crowd evacuation. Reference [41] proposed a leader–follower model for crowd evacuation simulation. A crowd includes several groups, each having a leader and some followers. Here, leaders were responsible for determining the evacuation path for their followers. The objective of their simulation was to show the effects of different numbers of leaders upon the evacuation efficiency. Reference [72] proposed a cellular automaton model for crowd movement simulation by embedding the follow-the-leader technique as its fundamental driving mechanism. This study showed that it is possible to achieve path planning by controlling the grouping behavior.

Herein, we focus on a grouping strategy considering the relationship and the distance. Concretely, the individuals are grouped according to the selected exits, as well as the distances between exits and individuals. Pedestrians related to one another are divided into the same group. The group's movement direction is decided by the leader. Individuals with the smallest evacuation time, which is calculated as a fitness function, are selected as leaders. The fitness function considers the influence of distance and congestion; more details can be found in [52].

The application of a grouping strategy serves to simplify the evacuation process. When the strategy is applied, the process can be divided into two periods: group formation and movement of groups. Then, the requirement of optimal-path planning is to identify a path of minimum length from the starting node to the target, together with a collision-free path [56]. The advantage of this approach is that it can improve the performability of path planning and benefit the application of optimization algorithms. However, the definition of groups presents only an optimized state of the crowd for evacuation management. When pedestrians are forming the group, their movement is still influenced by the crowd dynamics. Therefore, the definition of the grouping strategy should be continuously modified according to study of the crowd dynamics.

After defining the group, it is necessary to discuss the main optimization problem in this work. Although we have proposed a fitness function to choose the leaders of every group, it is not sufficiently accurate to confirm most effective ones. Leaders with higher fitness may have lower efficiency in leading pedestrians than those with lower fitness. Therefore, we use this function to provide a threshold for selecting from many candidates who are sorted by their fitness value. The optimization problem is then to determine the leaders who can achieve the shortest evacuation time.

Researchers have proposed various methods for solving the path planning problem. In this section, we focus upon the application of swarm intelligence optimization algorithms, such as the ACO [91], PSO [69], and ABC algorithm [83], to evacuation management. Herein, we modified and applied the ABC algorithm to

Table 1 The process of the improved ABC algorithm

Step 1: Initialize the number and position of the individuals, the number of iterations, and the related parameters
Step 2: Evaluate the fitness value of each individual and sort the swarm in descending order according to the result
Step 3: Select the top 50% of individuals as the lead bees from the matrix obtained in Step 2. The individuals are grouped according to the selected exits and the distances between the exits and individuals
Step 4: Compute the selected probability of each lead bee
Step 5: Switch the roles of the remainders to scouts or onlookers in each group
Step 6: Update the individuals' positions according to their roles in each group
Step 7: Driven by an selected crowd dynamics model, the individuals move toward the exits
Step 8: Return to Step 2 if the iterative condition is met; otherwise, exit the iterations

optimize the selection of a leader for the evacuation process. The concrete algorithm is shown in Table 1. Unlike the original ABC algorithm, our improved approach only conducts an iterative calculation in every group to select the right leader. Thus, the algorithm's accuracy and efficiency can be improved.

After implementing the algorithm, the crowd dynamics models are applied via the iteration computation method to calculate the evacuation times for different leaders. To ensure the effect of the optimization, scholars in the fields of management and optimization usually apply existing crowd dynamics models to calculate the evacuation time. Therefore, the selection of an appropriate model is important for the optimization to work appropriately. This selection is mainly influenced by two factors: the simulation authenticity and the computational complexity.

The simulation authenticity guarantees the calculation of the evacuation time. In a study on path planning, the requirement for the authenticity is to represent the concrete movement behavior of pedestrians; for example, because of the influence of a leader at a microcosmic scale, we use the social force model to simulate the evacuation process. We have also modified this model to represent the grouping behavior in a crowd [52]. A vision factor is added to the social force model to represent pedestrians' ability to determine the group leader.

As crowd evacuation is a dynamic process, it must be ensured in real time or with a short time delay. Thus, the computational complexity of the simulation is a critical problem facing optimization. Herein, although the social force model can represent the group behavior, its computational complexity will rapidly increase with the population. From the aforementioned description, we should determine a selection that balances the requirements for simulation authenticity and computational complexity.

3.1.2 Optimization of Grouping Behavior with Navigation Knowledge

Although we have applied the grouping behavior to improving the path planning, there are many limitations to this approach. In this section, we focus on the following questions to improve this approach's effectiveness:

1. How can it be ensured that group members can always follow the leader? When pedestrians are in emergency situations, their ability to recognize the environment may be influenced by panic. Therefore, it is difficult to ask pedestrians to naturally remain in the group.
2. How can the groups' evacuation processes be optimized? The initial group is divided according to its best exit and path; however, these may change during the evacuation process. Thus, it is necessary to improve path planning for groups by accounting for the influence of environmental information.

In view of the aforementioned problems, we use information technology to optimize group-based path planning. Specifically, unlike in the Sect. 3.1.1, pedestrians are considered able to receive messages on their mobile phones; therefore, pedestrians can obtain the position of the group leader at any time, and this ability is not restricted by vision or emotional state. Thus, we propose a knowledge-based approach to improve the information utilization effect in path planning. According to this approach, a two-layer control mechanism is built to organize and provide the evacuation information for the group leader. This information is saved in the knowledge base, and information transmission is achieved by navigation agents, who correspond to the leaders. The belief space comprises different types of information for calculating personal and global best positions. The navigation agents conduct path planning for every leader to provide the next target position by considering the current position and the congestion of the obstacles and exits. Figure 5 shows the scheme of our two-layer control mechanism.

In our approach, the cultural algorithm (CA) and the agent-based model are used to collect, analyze, and transfer information for path planning. The CA was developed to model the evolving cultural components of an evolutionary computational system over time as it accumulates experience [18, 64]. The double-layered structure and evolutionary mechanism of this structure have been successfully applied to many research fields [31, 34, 51, 84]. Thus, it is possible to consider the CA as a framework for solving crowd evacuation problems. Agent-based modeling is an approach to modeling systems comprising individual, autonomous, interacting agents. Such a model is based on the representation of global behavior from the rules provided to individuals, which may enable them to view the macro-level consequences of micro-level interactions [57]. It is obvious that agent-based modeling is suitable for describing the crowd movement, and many scholars have studied its application effects [20, 65, 73].

Herein, the social force model is still used as the iteration calculation approach. As pedestrians can remain in groups using their mobile phones, the modification of the SFM has also been simplified. Unlike in the Sect. 3.1.1, we only change the direction of pedestrians' driving force to represent the group behavior. In this

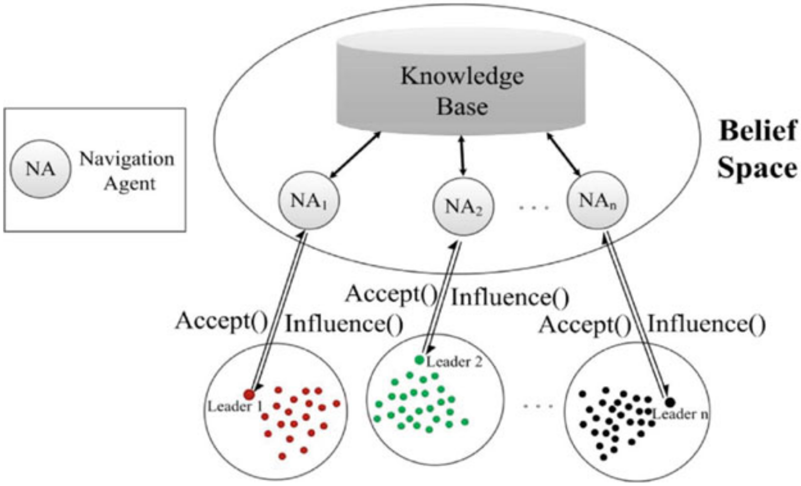


Fig. 5 Scheme of the two-layer control mechanism for group-based path planning

manner, the simulation authenticity of the original SFM is ensured while reducing the computational complexity.

To summarize, we presented a group-based approach to improving path planning during evacuations. According to this approach, the crowd evacuation process has been simplified to different periods. Then, intelligent algorithms can be easily applied to guide path planning. However, when calculating the optimization results, the complete crowd dynamics should still be considered. Therefore, the existing crowd dynamics models are used to iteratively calculate the evacuation time. The aforementioned work shows that the application of crowd dynamics models is a critical foundation for path planning and the application of intelligent algorithms. Thus, the next problem is selecting a model with the appropriate simulation authenticity and computational complexity.

3.2 Positive Emotional Contagion During Crowd Evacuation

When pedestrians are in emergency situations, their irrational behavior may cause dangerous situations, such as stampedes and crushes. Negative emotion has been considered a major cause for pedestrians' irrational behavior. Thus, it is important for evacuation management to reduce the influence of negative emotions. In this section, we introduce approaches that aim to optimize positive emotional contagion to decrease the influence of negative emotions. The concrete method of these approaches is to dispatch safety officers to calm panicked pedestrians down. Our approaches are expected to assist in emergency preplanning and to provide guidance for emergency management.

3.2.1 Strategies for Utilizing Positive Emotional Contagion

This work studies how to dispatch safety officers to avoid the danger caused by negative emotion. Here, the term safety officer refers to a person who is responsible for the "safety" of the people who work in or visit an area. The positive emotions of safety officers can help to calm the crowd and assist in an orderly evacuation, thereby effectively avoiding a stampede. To achieve the aforementioned target, we focus on the following question: how can one describe the process of positive emotional contagion during crowd evacuation? When modeling pedestrians' emotional contagion, existing studies have not fully considered the effects of positive emotions [11]. Furthermore, because of the influence of safety officers, other features, such as the trust relationship, should be considered. How can the effectiveness of safety officers be optimized? The number of such officers is usually limited, and their calming effect is limited by their physical location and the number of pedestrians passing by.

We first briefly review studies of emotional contagion during crowd evacuation. Many factors can influence crowd evacuation, with the leading emotion in the crowd exerting the greatest influence upon the effectiveness of rescue work during emergencies [75]. Thus, increasing efforts have been made to study emotional models in recent years [2, 30, 61, 67]; for example, the OCC (Ortony, Clore, Collins) emotional model is widely used in AI applications because of its structural, rule-based form [62].

Research on emotional contagion is an important topic in emotional modeling, and some previous works have investigated the dynamics of emotional contagion within a social network [6, 7, 70, 79]. The ASCRIBE model was the first to describe the emotional contagion within a crowd [8, 10, 59, 60, 76]; this model assumes that the emotional contagion process among individuals is similar to the heat dissipation phenomenon studied in thermodynamics. As an alternative to the ASCRIBE model, many researchers have studied the process of emotional contagion using epidemiology-based methods inspired by the spread of disease, as described in the epidemiology literature [11, 19, 22]. In these approaches, emotional contagion is assumed to be similar to the diffusion of infectious diseases. However, most methods used in the field of crowd evacuation focus on the negative emotional contagion observed in emergency situations.

Studying the problem of how the positive emotions of safety officers help to calm a crowd is crucial in practical applications. The influence of positive emotions has been extensively studied [22, 28]; during crowd evacuations, such emotions can guide individuals to establish rational behaviors [12]. However, unlike the in-depth research on negative emotional contagion, only a few studies have explored the positive case. The predefined rules used in these methods oversimplify real-world situations, making the analysis unquantifiable. In addition, the authors have not studied how to maximize the influence of positive emotions.

Results show that positive emotional contagion is mainly influenced by the trust relationships among pedestrians [87]. Here, the trust relationships represent the probability of a pedestrian being infected by another one. For example, in emergency situations, people tend to be more willing to believe information from safety officers than from ordinary people. We assumed that pedestrians can only receive information by observing their environment; therefore, the emotional contagion between two pedestrians is influenced by their visibility. Then, a trust-based emotional contagion network (Trust-ECN) is developed to describe the emotional contagion process. The Trust-ECN $G(V, E)$ is a directed graph in which each node $v \in V$ represents an individual in the crowd and each directed edge $e(i, j) \in E$ from node i to node j denotes that j trusts i and that i is visible to j, meaning that j can be directly infected by the positive emotion of i.

Calculating the speed of the spread is an important step for optimizing positive emotion. We computed the spreading speeds of emotional contagion by considering the relevant factors. In particular, the ability to express or assimilate positive emotions depends upon an individual's personality.

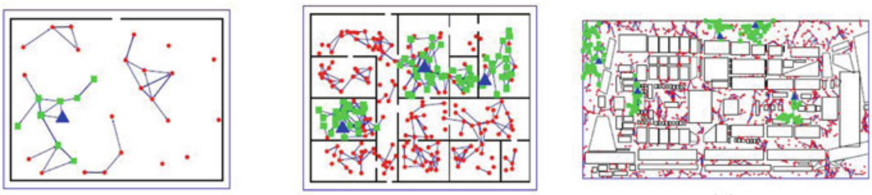
Following the lines of [55], we apply the Big Five model to describe the influence of pedestrians' personalities. Extroversion is used to describe their ability to express positive emotion, and neuroticism is used to represent their ability to assimilate it. To facilitate analysis, we assume that both extroversion and neuroticism can be numerically formalized by values drawn from a specified probability distribution.

We also assume that positive emotional contagion decays smoothly with increasing distance between i and j because common sense suggests that our visual and auditory systems become more likely to produce errors with the increase in the transmission distance. After considering the aforementioned factors, the spreading speed of emotional contagion from i to j is expressed as follows

$$v_(ij) = \varepsilon_i \cdot R_{ij} \cdot a_j. \tag{4}$$

Here, R_{ij} represents emotional attenuation, ε_i denotes the ability of i to express positive emotions, and a_j denotes the ability of j to assimilate positive emotions.

The effect of the positive emotional contagion network is verified using the number of pedestrians infected in a given time window T, see Fig. 6. The emotional state of individual i is represented by the infection probability $P(t_i < T|S, V)$,



(a) (b) (c)

Fig. 6 Emotional contagion networks and the results of positive emotional contagion from [87]. The blue markers (triangles) are safety officers, whereas the red/green markers (circles/squares) are individuals with/without negative emotions. (**a**) The room network; (**b**) the office network; (**c**) the square network (colored in the online version) where S is the set of m safety officers selected at time $t = 0$, and V is the set of emotional contagion speeds. The infection probability of each safety officer is set to 1. As for the other pedestrians, the calculation of their infection probability is based upon the time difference, which is a random variable associated with each edge in $G(V, E)$.

Given a crowd with negative emotions, a time window T, and a variable m, our target is to select safety officers optimal positions for to maximize the "calm-down" effect. The optimization problem is presented as follows:

$$S^{\circ} = max_{s=m}(\mu(S, T, V)) = EI(S, T, V) = \sum_{i=1}^{C} P(t_i < T|S, V), \tag{5}$$

where S° is the number of safety officers with optimal positions, $I(S, T, V)$ is the number of individuals who can be infected up to time T, $\mu(S, T, V)$ is the expectation value of $I(S, T, V)$, which is defined by the average number of nodes infected prior to time T, t_i is the infection time of i, C is the size of the crowd, and $P(t_i < T|S, V)$ is the infection probability of individual i prior to T. The positions of the safety officers can be obtained intuitively because the individuals' initial positions are given. Herein, this problem is optimized using an artificial bee colony optimized emotional contagion (ABCEC) algorithm.

In this section, we have proposed a feasible framework for combining the modeling strategies and optimization algorithms for positive emotional contagion. The dispatching of safety officers can truly improve evacuation management; however, research on positive emotional contagion is just beginning. Therefore, many limitations should still be considered. A critical reason for this is that many of the features of the crowd dynamics have been ignored to simplify the evacuation process. Given that the emotional-infection probability is influenced by the distance between pedestrians, it is necessary to consider the crowd movement in the iterative calculation. To improve the optimization efficiency, we define a random variable to represent the influence of distance on the infection probability. Thus, the iterative calculation only uses the initial distribution of pedestrians. However, the aforementioned approach also leads to a requirement that the emotional contagion network should be defined so as to consider more features of crowd dynamics. This is a direction for future study.

3.2.2 Optimization of the Positive Emotional Contagion

After confirming our approach for positive emotional contagion, we focus upon improving this method's effectiveness. Although our previous work combined safety officer guidance and positive influence maximization, it is limited by the simplification of crowd dynamics. Therefore, we optimize the positive emotional contagion by adding more features of crowd dynamics.

Here, we focus on extending the application of positive emotional contagion to affect more random individuals in emergency situations. To achieve this, we apply IoT devices to capture more details of crowd dynamics and to develop an entropy-based anisotropic emotional contagion model to account for the influence of crowd chaos, which is a critical factor determining crowd behaviors.

Recently, IoT devices have been widely applied to capturing the complex features of crowd movement. Research on IoT-based crowd evacuation has mostly focused on discovering evacuation routes [24, 44, 74, 77], navigating evacuation [14, 44, 82], and simulating crowd evacuation [13, 63, 80, 81]. In evacuation route discovery, the predeployed IoT devices are used to explore the areas in which emergencies occur and to discover the optimal evacuation paths from the site of the accident to the emergency exits. In the evacuation navigation approaches, the escape indication devices are used to guide the crowd toward the exits. Moreover, the data sensed by IoT devices are also used to enhance the visual realism of the crowd simulation. The aforementioned literature shows that IoT devices have sufficient accuracy to capture real-time information to improve positive emotional contagion. Herein, we use the data captured by IoT devices to quantify the chaos of crowd behavior, which is a fundamental factor affecting crowd evacuation.

To further improve the positive emotional contagion, it is necessary to take more details of crowd dynamics into account. Herein, we focus on the influence of crowd chaos [88]; here, "chaos" refers to a nonuniform pedestrian movement. Pedestrians in a crowd may have different directions, and their emotions are more likely influenced by other pedestrians moving in the same direction. The positive emotion propagates quickly in an ordered crowd; however, it plays a less important role in crowd chaos regulation, because most individuals are already ordered. Positive emotion has a greater calming effect in a chaotic crowd, but it propagates slowly. Those chaotic individuals whose directions of motion are inconsistent with the crowd cause the greatest damage to the safety of crowd evacuation. Thus, understanding chaos can be used to define and develop new strategies that build upon the effect of safety officers' positive emotional contagion during crowd evacuation.

There are many differences between crowd chaos and congestion. On the one hand, congestion is caused by the phenomenon whereby space cannot satisfy pedestrians' requirements, and it is mainly embodied as an objective interaction between pedestrians; crowd chaos, on the other hand, it is mainly embodied as the uncertain and nonuniform movement caused by negative emotional contagion. In our opinion, crowd chaos is similar to entropy in a thermodynamic system. Thus, we introduce the definitions of crowd entropy to quantify crowd chaos; specifically, during the emotional contagion process, entropy has a nonuniform impact on the propagation rate: the larger the entropy is, the more chaotic the crowd, and the slower the propagation rate will be.

Herein, we define the crowd entropy from both macroscopic and microscopic perspectives [88]. The macroscopic entropy is used to measure the influence of the whole crowd's chaos upon the local pedestrians at the same orientation, whereas the microscopic entropy is used to measure the degree of microchaos among individuals

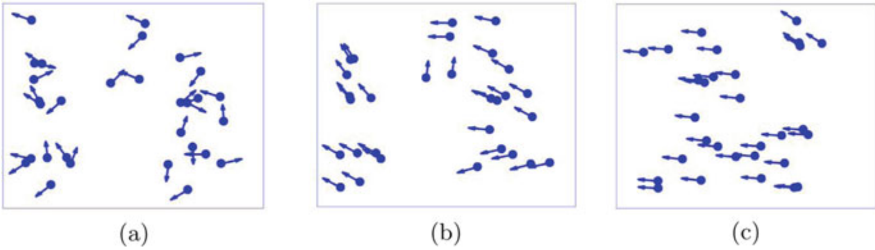


Fig. 7 The effect of crowd entropy from [88]. There are 30 particles moving at a constant speed of 0.03, and the direction of each particle's motion is the average direction of neighboring particles with some random perturbation: (**a**) $E = 1.3$; (**b**) $E = 0.34$; and (**c**) $E = 0.01$

of the same orientation, and it is defined as the velocity correlation between two pedestrians. After confirming the entropy, the emotional propagation rate is influenced by the following rule: the smaller the entropy, the more orderly the crowd, and the faster the propagation rate will be. The effect of the crowd's entropy is shown in Fig. 7; initially, the particles move randomly and their entropy attains a high value. The particle motion becomes increasingly orderly and the entropy decreases accordingly, showing that the proposed definition of crowd entropy effectively quantifies the chaos of the crowd's motion.

The crowd's entropy represents the influence of complex dynamic features upon positive emotional contagion. When defining the aforementioned variable, we consider the macroscopic orientation difference and the microscopic transfer chains; then, the effect of the crowd entropy is represented by the propagation rate of positive emotion. In our opinion, the definition of crowd entropy accords with the current developmental tendency (mesoscopic) of crowd dynamics modeling. However, this definition is unable to fully take advantage of crowd dynamics. This is attributed to the entropy being computed using a snapshot of the crowd at an instant in time, with each individual's motion direction assumed to be fixed. When considering more complex crowd dynamics (e.g., dynamic evolution of chaos), the number of microstates will become unaccountably infinite.

4 Summary

In this chapter, we reviewed several works on feature analysis and process optimization to discuss the application of crowd dynamics models to evacuation management. We first presented an introduction to research on crowd congestion. Thus, this work has improved the effectiveness of evacuation management through congestion analysis and alleviation. Furthermore, we have discussed several critical problems and the role of crowd dynamics models in features analysis. The main conclusions and hypotheses are presented as follows:

1. It is feasible to achieve feature analysis by defining and analyzing the constraints for crowd movement; the existing models offer definitions of the constraints.
2. The main problem in feature analysis is to confirm the critical characteristics; this problem is also important for research on crowd dynamics.
3. The self-organization phenomenon, which is usually mesoscopic, can connect the microscopic pedestrian movement and its macroscopic features.

We have reviewed works on process optimization in evacuation management. The grouping behavior is used to optimize path planning to improve the evacuation efficiency. To reduce panic during an evacuation, we also study the optimization of positive emotional contagion. Although process optimization works usually have different targets, there are still many similar characteristics between them. First, the optimization is conducted on a process that is simplified from the complete crowd movement, and the simplification is based on the analysis of a key feature; second, with this study's development, the optimization has to consider more details of crowd dynamics. Finally, given that process optimization focuses more on the application of other technologies and algorithms, the understanding and applications of crowd dynamics are usually based on existing research. In our opinion, the aforementioned characteristics show some feasible directions for the application for crowd dynamics models.

Although we have applied many interdisciplinary approaches (theories and algorithms) to research on evacuation management, their effectiveness is mainly decided by the current understanding of crowd dynamics. For example, the effects of congestion analysis, group behavior, and emotional contagion are all determined by the definition of the interaction between the crowd and individual pedestrians. Therefore, understanding the multiscale interaction remains one of the main challenges that modeling of the crowd dynamics must face in the future.

References

1. G. Albi, N. Bellomo, L. Fermo, S. Y. Ha and J. Soler, Vehicular traffic, crowds, and swarms. From kinetic theory and multiscale methods to applications and research perspectives. *Mathematical Models and Methods in Applied Sciences* **29** (2019) 1901–2005.
2. M. Belhaj, F. Kebair and L. B. Said, A computational model of emotions for the simulation of human emotional dynamics in emergency Situations. *International Journal of Computer Theory and Engineering* **6** (2014) 227–233.
3. N. Bellomo and A. Bellouquid, On multiscale models of pedestrian crowds from mesoscopic to macroscopic. *Communications in Mathematical Sciences* **13** (2015) 1649–1664.
4. N. Bellomo and L. Gibelli, Toward a mathematical theory of behavioral social dynamics for pedestrian crowds. *Mathematical Models and Methods in Applied Sciences* **25** (2015) 2417–2437.
5. N. Bellomo and L. Gibelli, Behavioral Human Crowds. *Crowd Dynamics Vol.1* (2018).
6. N. Bellomo, L. Gibelli and N. Outada, On the interplay between behavioral dynamics and social interactions in human crowds. *Kinetic and Related Models*, **12** (2019) 397–409.
7. A. L. Bertozzi, J. Rosado, M. B. Short, and L. Wang, Contagion shocks in one dimension. *Journal Statistical Physics*, **158** (2015) 647–664.

8. J. Bordas, Omnemotion: the propagation of emotions. *Proc. of the Virtual Reality International Conference* (2014).
9. Boris Andreianov, Carlotta Donadello, Ulrich Razafison and Massimiliano Daniele Rosini, One-Dimensional Conservation Laws with Nonlocal Point Constraints on the Flux. *Crowd Dynamics Vol.1* (2018).
10. T. Bosse, R. Duell, Z. A. Memon, J. Treur and C. N. Wal, A Multi-agent model for mutual absorption of emotions. *Proc. Of European Council on Modeling and Simulation* (2009) 212–218.
11. M. Cao, G. Zhang, M. Wang, D. Lu and H. Liu, A method of emotion contagion for crowd evacuation. *Physica A: Statistical Mechanics and its Applications* **483** (2017) 250–258.
12. J. R. Carl, D. P. Soskin, C. Kerns and D. H. Barlow, Positive emotion regulation in emotional disorders: a theoretical review. *Clinical Psychology Review* **33** (2013) 343–360.
13. P. Charalambous and Y. Chrysanthou, The PAG crowd: A graph based approach for efficient data-driven crowd simulation. *Computer Graphics Forum* **33** (2015) 95–108.
14. L. Chen, J. Cheng and Y. Tseng, Distributed emergency guiding with evacuation time optimization based on wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems* **27** (2016) 419–427.
15. J. C. Chu, A. Y. Chen and Y. F. Lin, Variable guidance for pedestrian evacuation considering congestion, hazard, and compliance behavior. *Transportation Research Part C: Emerging Technologies* **85** (2017) 664–683.
16. E. Cristiani and D. Peri, Handling obstacles in pedestrian simulations: models and optimization. *Applied Mathematical Modelling* **45** (2017) 285–302.
17. E. Cristiani, F. S. Priuli and A. Tosin, Multiscale Modeling of Granular Flows with Application to Crowd Dynamics. *Multiscale Modeling and Simulation* **9** (2011) 155–182.
18. J. G. Digalakis and K. G. Margaritis, A multipopulation cultural algorithm for the electrical generator scheduling problem. *Mathematics and Computers in Simulation* **60** (2002) 293–301.
19. P. Dodds and D. J. Watts, A generalized model of social and biological contagion. *Journal of Theoretical Biology* **232** (2005) 587–604.
20. R. J. Dawson, R. Peppe and M. Wang, An agent-based model for risk-based flood incident management. *Nature Hazards* **59** (2011) 167–189.
21. J. Drury, C. Cocking and S. Reicher, Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. *British Journal of Social Psychology* **48** (2009) 487–506.
22. H. Faroqi and M. Mesgari, Agent-based crowd simulation considering emotion contagion for emergency evacuation problems. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2015) 193–196.
23. C. Feliciani and K. Nishinari, Measurement of congestion and intrinsic risk in pedestrian crowds. *Transportation Research Part C: Emerging Technologies* **91** (2018) 124–155.
24. E. Ferranti and N. Trigoni, Robot-assisted discovery of evacuation routes in emergency scenarios. *Proceedings of IEEE International Conference on Robotics and Automation* (2008) 2824–2830.
25. M. Filipovic, A. Djuric and L. Kevac, The significance of adopted Lagrange's principle of virtual work used for modeling aerial robots, *Appl. Math. Modell.* **39** (2015) 1804–1822.
26. L. Fisher, The Perfect Swarm: The Science of Complexity in Everyday Life. *Basic Books* (2009).
27. G. A. Frank and C. O. Dorso, Room evacuation in the presence of an obstacle. *Physica A: Statistical Mechanics and its Applications* **390** (2011) 2135–2145.
28. B. L. Fredrickson, The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist* **56** (2001) 218–226.
29. A. Garcimartín, J. M. Pastor, C. Martín-Gómez, D. Parisi and I. Zuriguel, Pedestrian collective motion in competitive room evacuation. *Scientific Reports* **7** (2017) 10792.
30. J. Gratch and S. Marsella, A domain-independent framework for modeling emotion. *Cognitive Systems Research* **5** (2004) 269–306.

31. Y. N. Guo, D. Liu , J. Cheng , Multi-population cooperative cultural algorithms. *International Conference on Intelligent Computing (ICIC): Bio-Inspired Computing and Applications* (2011) 199–206.
32. Y. Han and H. Liu, Modified social force model based on information transmission toward crowd evacuation simulation, *Physica A* **469** (2017) 499–509.
33. Y. Han, H. Liu and P. Moore, Extended route choice model based on available evacuation route set and its application in crowd evacuation simulation, *Simul. Model. Pract. Theory.* **75** (2017) 1–16.
34. D. X. He, G. H. Xu, Z. H. Guan , M. Chi and D.F. Zheng, Hybrid coordination of multi-agent networks with hierarchical leaders. *Communications in Nonlinear Science and Numerical Simulation* **27** (2015) 110–119.
35. D. Helbing, I. Farkas and T. Vicsek, Simulating dynamical features of escape panic. *Nature* **407** (2000) 487–490.
36. D. Helbing, A. Johansson and H. Z. Al-Abideen, The dynamics of crowd disasters: an empirical study. *Physical Review E* **75** (2007) 046109.
37. D. Helbing and P. Molnár, Social force model for crowd dynamics, *Physical Review E* **51** (1995) 4282–4286.
38. R. L. Hughes, The flow of large crowds of pedestrians. *Mathematics and Computers in Simulation* **53** (2000) 367–370.
39. R. L. Hughes, A continuum theory for the flow of pedestrians. *Transportation Research Part B-Methodological* **36** (2002) 507–535.
40. A. Johansson and D. Helbing,. Pedestrian flow optimization with a genetic algorithm based on Boolean grids. *Pedestrian and Evacuation Dynamics 2005. Springer Berlin Heidelberg* (2007).
41. Q. G. Ji, C. Gao, Simulating crowd evacuation with a leader-follower model. *International Journal of Computer Sciences and Engineering Systems* **1** (2007) 249–252.
42. L. Jiang, J. Li, C. Shen, S. Yang and Z. Han, Obstacle optimization for panic flow-reducing the tangential momentum increases the escape speed. *PloS one.* **9** (2014) e115463.
43. H. L. Jia, C. H. Wang and J. H. Li, Discussion on some issues in theory of soil arch. *Journal of Southwest Jiao tong University* **38** (2003) 398–402.
44. M. Khan, O. Khalid, Y. Huang, R. Ranjan, F. Zhang, J. Cao, B. Veeravalli, S. Khan, K. Li and A. Zomaya, MacroServ: A route recommendation service for large-scale evacuations. *IEEE Transactions on Services Computing* **10** (2017) 589–602.
45. D. Kim and A. Quaini, A kinetic theory approach to model pedestrian dynamics in bounded domains with obstacles. *Kinetic and Related Models* **12** (2019) 1273–1296.
46. D. Kim and A. Quaini, Coupling kinetic theory approaches for pedestrian dynamics and disease contagion in a confined environment. *Mathematical Models and Methods in Applied Sciences* **30** (2020) 1893–1915.
47. A. Kirchner, K. Nishinari and A. Schadschneider, Friction effects and clogging in a cellular automaton model for pedestrian dynamics. *Physical Review E* **67** (2003) 056122.
48. B. Krausz and C. Bauckhage, Analyzing pedestrian behavior in crowds for automatic detection of congestions. *IEEE International Conference on Computer Vision Workshops (ICCV 2011 Workshops, Barcelona, Spain, 2011)*, 144–149.
49. L. Li, H. Liu and Y. Han, Arch formation-based congestion alleviation for crowd evacuation. *Transportation Research Part C: Emerging Technologies* **100** (2019) 88–106.
50. L. Li, H. Liu and Y. Han, An approach to congestion analysis in crowd dynamics models. *Mathematical Models and Methods in Applied Sciences* **30** (2020) 867–890.
51. H. Liu, B. Liu, H. Zhang, L. Li, X. Qin and G. Zhang, Crowd evacuation simulation approach based on navigation knowledge and two-layer control mechanism. *Information Sciences* **436–437** (2018) 247–267.
52. H. Liu, B. Xu, D. Lu and G. Zhang, A path planning approach for crowd evacuation in buildings based on improved artificial bee colony algorithm. *Applied Soft Computing* **68** (2018) 360–376.
53. Y. Ma, E. W. M. Lee, R. K. K. Yuen, Dual effects of pedestrian density on emergency evacuation. *Physics Letters A* **381** (2017) 435–439.

54. T. Masuda, K. Nishinari and A. Schadschneider, Critical bottleneck size for jamless particle flows in two dimensions. *Physical Review Letters* **112** (2014) 138701.
55. R. R. Mccrae and P. T. Costa, Personality in adulthood: a five- factor theory perspective. *International and Cultural Psychology* (2003) 303–322.
56. Q. Meng and X. B. Qu, Uncertainty propagation in quantitative risk assessment modeling for fire in road tunnels. *IEEE Trans. Cybern. Part C* **42** (2012) 1454–1464.
57. A. G. Milla, F. Borges, R. Suppi and E. Luque, Crowd evacuations SaaS: an ABM approach. *Procedia Computer Science* **51** (2015) 473–482.
58. M. Moussaïd, N. Perozo, S. Garnier, D. Helbing and G. Theraulaz, The walking behavior of pedestrian social groups and its impact on crowd dynamics. *PLoS One* **5** (2010) e.0010047.
59. A. Neto, C. Pelachaud, and S. Musse, Emotion contagion model for crowds. *SBC Journal on Interactive Systems* **6** (2015) 37–44.
60. A. Neto, C. Pelachaud, and S. Musse, Giving emotional contagion ability to virtual agents in crowds. *Proceedings of the International Conference on Intelligent Virtual Agents* (2017) 63–72.
61. M. Ntika, I. Sakellariou, P. Kefalas, and I. Stamatopoulou, Experiments with emotion contagion in emergency evacuation simulation. *Proceedings of International Conference* (2014).
62. A. Ortony, G. Clore and A. Collins, The Cognitive Structure of Emotions. *Cambridge University Press, Cambridge* (1988).
63. A. Rimboux, R. Dupre, E. Daci, T. Lagkas, P. G. Sarigiannidis, P. Remagnino and V. Argyriou, Smart IoT cameras for crowd analysis based on augmentation for automatic pedestrian detection, simulation and annotation. *Proceedings of the 15th International Conference on Distributed Computing in Sensor Systems* (2019) 304–311.
64. R. G. Reynolds, An introduction to cultural algorithms. *Proceedings Third Annu. Conf. Evol. Program* (1994) 131–139.
65. J. Shi, A. Ren and C. Chen, Agent-based evacuation model of large public buildings under fire conditions. *Automation in Construction* **18** (2009) 338–347.
66. J. D. Sime, Affiliative behaviour during escape to building exits. *Journal of Environmental Psychology* **3** (1983) 21–41.
67. J. Sinclair and C. S. M. Lui, Integrating personality and emotion for human crowd simulation. *Proceedings of the Fifteenth International Conference on Electronic Business* (2015) 300–307.
68. K. Suzuno, A. Tomoeda, M. Iwamoto and D. Ueyama, Dynamic structure in pedestrian evacuation: Image processing approach. *Traffic and Granular Flow'13. Springer, Cham* (2015) 195–201.
69. T. ThoaMac, C. Copot, D.T. Tran and R. D. Keyser, A hierarchical global path planningapproach for mobile robots based on multi-objective particle swarmoptimization. *Applied Soft Computing* **59** (2017) 68–76.
70. J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen, and A. C. M. Fong, Quantitative Study of Individual Emotional States in Social Networks. *IEEE Transactions on Affective Computing* **3** (2012) 132–144.
71. M. Twarogowska, P. Goatin and R. Duvigneau, Macroscopic modeling and simulations of room evacuation. *Applied Mathematical Modelling* **38** (2014) 5781–5795.
72. C. Vihas, I.G. Georgoudas and G.Ch. Sirakoulis, Cellular automata incorporatingfollow-the-leader principles to model crowd dynamics. *Journal of Cellular Automata* **8** (2013) 333–346.
73. N. Wagner and V. Agrawal, An agent-based simulation system for concert venue crowd evacuation modeling in the presence of a fire disaster. *Expert Systems with Applications* **41** (2014) 2807–2815.
74. Q. Wang, B. Guo, L. Wang, T. Xin, H. Du, H. Chen, and Z. Yu, CrowdWatch: Dynamic sidewalk obstacle detection using mobile crowd sensing. *IEEE Internet of Things Journal* **4** (2017) 2159–2171.
75. J. Wang, S. Lo, J. Sun, Q. Wang and H. Mu, Qualitative simulation of the panic spread in large-scale evacuation. *Simulation* **88** (2012) 1465–1474.

76. L. Wang, M. Short and A. Bertozzi, Efficient numerical methods for multiscale crowd dynamics with emotional contagion. *Mathematical Models and Methods in Applied Sciences* **27** (2016) 205–230.
77. X. Xu, L. Zhang, S. Sotiriadis, E. Asimakopoulou, M. Li and N. Bessis, Clotho: A large-scale internet of things-based crowd evacuation planning system for disaster management. *IEEE Internet of Things Journal* **5** (2018) 3559–3568.
78. D. Yanagisawa, R. Nishi, A. Tomoeda, K. Ohtsuka and K. Nishinari, Study on efficiency of evacuation with an obstacle on hexagonal cell space. *SICE Journal of Control Measurement and System Integration* **3** (2010) 395–401.
79. Y. Yang, J. Jia, B. Wu and J. Tang, Social role-aware emotion contagion in image social networks. *Proc. of the 30th AAAI Conference on Artificial Intelligence* (2016) 65–71.
80. Z. Yao, G. Zhang, D. Lu and H. Liu, Data-driven crowd evacuation: A reinforcement learning method. *Neurocomputing* **366** (2019) 314–327.
81. Z. Yao, G. Zhang, D. Lu and H. Liu, Learning crowd behavior from real data: A residual network method for crowd simulation. *Neurocomputing* **404** (2020) 173–185.
82. K. Yu, H. Hsu, N. Chung, C. Lien, S. T. Chen, M. Y. Lei,and N. Tsai, Construct an intelligent evacuation guidance system with open system architecture. *Proceedings Of International Conference on Advanced Control* (2018).
83. T. Yue and H. Y. Chung, Using ABC and RRT algorithms to improve mobile robotpath planning with danger degree. *2016 Fifth International Conference onFuture Generation Communication Technologies, IEEE, Luton, UK* (2016) 21–26.
84. P. M. Zadeh and Z. Kobti, A multi-population cultural algorithm for community detection in social networks. *Procedia Computer Science* **52** (2015) 342–349.
85. J. Zhang, W. Klingsch, A. Schadschneider and A. Seyfried, Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. *Journal of Statistical Mechanics:Theory and Experiment* (2012) P02002.
86. P. Zhang, H. Liu and Y.H. Ding, Dynamic bee colony algorithm based on multi-species co-evolution. *Applied Intelligence* **40** (2014) 427–440.
87. G. Zhang, D. Lu and H. Liu, Strategies to utilize the positive emotional contagion optimally in crowd evacuation. *IEEE Transactions on Affective Computing* **11** (2018) 708–721.
88. G. Zhang, D. Lu and H. Liu, IoT-based Positive Emotional Contagion for Crowd Evacuation. *IEEE Internet of Things Journal* **8** (2020) 1057–1070.
89. Y. Zhao, M. Li, X. Lu, L. Tian, Z. Yu, K. Huang, Y. Wang and T. Li, Optimal layout design of obstacles for panic evacuation using differential evolution, *Physica A: Statistical Mechanics and its Applications* **465** (2017) 175–194.
90. X. P. Zheng and J. H. Sun, Analyzing pedestrian behavior in crowds for automatic detection of congestions. *Building and Environment* **45** (2010) 1755–1761.
91. Q. Zhu, J. Hu, W. Cai and L. Henschen, A new robot navigation algorithm for dynamic unknown environments based on dynamic path re-computation and an improved scout ant algorithm. *Applied Soft Computing* **11** (2011) 4667–4676.

Optimized Leaders Strategies for Crowd Evacuation in Unknown Environments with Multiple Exits

Giacomo Albi, Federica Ferrarese, and Chiara Segala

Abstract In this chapter, we discuss the mathematical modeling of egressing pedestrians in an unknown environment with multiple exits. We investigate different control problems to enhance the evacuation time of a crowd of agents, by few informed individuals, named leaders. Leaders are not recognizable as such and consist of two groups: a set of unaware leaders moving selfishly toward a fixed target, whereas the rest is coordinated to improve the evacuation time introducing different performance measures. Follower-leader dynamics is initially described microscopically by an agent-based model, subsequently a mean-field type model is introduced to approximate the large crowd of followers. The mesoscopic scale is efficiently solved by a class of numerical schemes based on direct simulation Monte-Carlo methods. Optimization of leader strategies is performed by a modified compass search method in the spirit of metaheuristic approaches. Finally, several virtual experiments are studied for various control settings and environments.

All authors acknowledge the support of the Italian Ministry of Instruction, University and Research (MIUR) with funds coming from PRIN Project 2017 (No. 2017KKJP4X entitled "Innovative numerical methods for evolutionary partial differential equations and applications") and from RIBA 2019 (prot. RBVR199YFL entitled "Geometric Evolution of Multi Agent Systems").

G. Albi (✉)
Dipartimento di Informatica, Università di Verona, Verona, Italy
e-mail: giacomo.albi@univr.it

F. Ferrarese · C. Segala
Dipartimento di Matematica, Università di Trento, Trento, Italy
e-mail: federica.ferrarese@unitn.it; chiara.segala-1@unitn.it

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_5

1 Introduction

Control methodologies for crowd motion are of paramount importance in real-life applications for the design of safety measures and risk mitigation. The creation of virtual models of a large ensemble of pedestrians is a first step for reliable predictions, otherwise not easily reproducible with real-life experiments.

Pedestrians have been properly modeled by means of different agent-based dynamics such as lattice models [21, 39], social force models [43, 50], or cellular automata models [1, 52]. A different level of description is obtained using mesoscopic models [2, 4, 36] where the quantities of study are densities of agents; at a larger scale macroscopic models [20, 22, 32] describe the evolution of moments such as mass and momentum. Multiscale models have been also considered, to account for situations where different scales coexist, we refer in particular to [26, 27]. Such a hierarchy of models is able to capture coherent global behaviors emerging from local interactions among pedestrians. These phenomena are strongly influenced by the social rules, the *rationality* of the crowd, and the knowledge of the surrounding environment. In the case of egressing pedestrians in an unknown environment with limited visibility we expect people to follow basically an instinctive behavior [12, 20, 21, 39], whereas a perfectly rational pedestrian will compute an optimal trajectory towards a specific target (the exit), forecasting exactly the behavior of other pedestrians [1, 46].

In this manuscript, we focus on the evacuation problem in an unknown environment with multiple exits. We aim at influencing their behavior towards the desired target with minimal intervention. Starting from the seminal work [4] we consider a bottom-up approach where few informed agents are acting minimizing verbal directives to individuals and preserving as much as possible their natural behavior. This approach is expected to be efficient in situations where direct communication is impossible, for example, in the case of very large groups, emergencies, violent crowds reluctant to follow directions; or in panic situations where rational behavior is overtaken by instinctive decisions. Furthermore, we consider few additional agents, who are informed about the position of some exit and acting as unaware leaders. Hence, their dynamics will influence the global behavior of the crowd, introducing inertia that may constitute an additional difficulty in the optimization problem, for example, increasing congestions next to the exits or increasing the level of uncertainty.

The control problem associated with the evacuation of a crowd falls in the larger research field aimed at investigating the control of *self-organizing agents*. From the mathematical view point, this type of problem is challenging due to the presence of non-local interaction terms and their high dimensionality. Control of alignment-type dynamics, such as the Cucker-Smale model [29], have risen a lot of interest in the mathematical community, where several strategies have been explored to enforce the emergence of consensus, see, for example, [5, 11, 14, 44]. At the same time, to cope with the high dimensionality of such optimal control problems, reduced approaches have been explored [15, 18, 37, 38], promoting sparsity of the control acting only on few agents. In biological models, it has been shown that a small percentage of individuals can influence the whole group towards a desired target, see [23]. Similarly, leaders in crowd can act as control signals to enforce alignment towards a desired direction as recognizable leaders [3, 9, 16, 34, 48], or moving undercover [4, 23, 33, 40, 41], or even in a repulsive way [17]. These strategies heavily rely on the power of the *social influence* (or herding effect), namely the natural tendency of people to follow other mates in situations of emergency or doubt.

Alternative control methodologies consist in optimal design of the surrounding environment such as obstacles [6, 24, 25], or evacuation signage [53, 54], or exit locations [52].

The manuscript is organized as follows in Sect. 2 we introduce the mathematical framework for the microscopic dynamics of leader-follower type and we formulate different scenarios for the optimal control problem to be solved. In particular, we will distinguish between minimum time of evacuation, total mass evacuated, and optimal mass splitting among the multiple exits. In this work the word mass denotes the total amount of pedestrians. Section 3 is devoted to the description of the mesoscopic scale, first we introduce the mean-field type model, second we sketch an efficient Monte-Carlo algorithm for its simulation. In Sect. 4 we focus on the numerical realization of the optimized strategies. We start introducing the algorithmic procedure used for the solution of the large-scale optimization problem, and we compare microscopic and mean-field dynamics in several scenarios and with different target functionals. Finally in Sect. 5 we outline possible extensions and further perspectives.

2 Control of Pedestrian Dynamics Through Leaders

In this section, we focus first on the mathematical description of pedestrian dynamics in complex environments. We consider an ensemble of agents, *followers*, in an unknown environment trying to reach exit locations, at the same time the crowd population includes few informed agents, *leaders*, acting as controllers but not distinguishable from followers. In particular, we account for a mixed approach where leaders are either aware of their role, then responding to an optimal force as the result of an offline optimization procedure, *optimized leaders*, or unaware of their role and moving with a greedy strategy towards a target exit position, *selfish leaders*. The main mechanisms ruling the behaviors among the followers are isotropic interactions with other agents based on metrical short-range repulsion, induced by social distancing and collisional avoidance, and topological long-range alignment dynamics. Leaders instead consider only short-range repulsion. Additionally, for followers, we account self-driving forces describing the exploration phase, preferential direction, and desired speed. The overall dynamics will be influenced by the surrounding environment when the exits are visible or close to obstacles.

In the following sections, we describe first the microscopic dynamics of the follower-leader system and later different control tasks for different applications.

2.1 Microscopic Model with Leaders and Multiple Exits

Following the approach proposed in [4, 6] we model leaders by a first-order model and followers by a second-order one, where both positions and velocities are state variables. We denote by d the dimension of the space in which the motion takes place (typically $d = 2$), by N^{F} the number of followers and by $N^{\mathrm{L}} \ll N^{\mathrm{F}}$ the number of leaders. We also denote by $\Omega \equiv \mathbb{R}^d$ the walking area, and we identify the different exits by $x_e^\tau \in \Omega$ with To define each target's visibility area, we consider the set Σ_e, with $x_e^\tau \in \Sigma_e \subset \Omega$, and we assume that the target is completely visible from any point belonging to Σ_e and completely invisible from any point belonging to $\Omega \backslash \Sigma_e$, namely we also assume that visibility areas are disjoint sets, i.e., $\Sigma_{e_i} \cap \Sigma_{e_j} = \varnothing$ for all $e_i, e_j \in \{1, .., N_e\}$.

For every $i = 1, \ldots, N^{\mathrm{F}}$, let $(x_i(t), v_i(t)) \in \mathbb{R}^{2d}$ denote position and velocity of the agents belonging to the population of followers at time $t \geq 0$ and, for every $k = 1, \ldots, N^{\mathrm{L}}$, let $(y_k(t), w_k(t)) \in \mathbb{R}^{2d}$ denote position and velocity of the agents among the population of leaders at time $t \geq 0$. Let us also define $\mathbf{x} := (x_1, \ldots, x_{N^{\mathrm{F}}})$ and $\mathbf{y} := (y_1, \ldots, y_{N^{\mathrm{L}}})$.

The microscopic dynamics described by the two populations is given by the following set of ODEs for $i = 1, \ldots, N^{\mathrm{F}}$ and $k = 1, \ldots, N^{\mathrm{L}}$,

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = S(x_i, v_i) + \sum_{j=1}^{N^{\mathrm{F}}} m_j^F H^{\mathrm{F}}(x_i, v_i, x_j, v_j; \mathbf{x}, \mathbf{y}) + \sum_{\ell=1}^{N^{\mathrm{L}}} m_\ell^L H^{\mathrm{L}}(x_i, v_i, y_\ell, w_\ell; \mathbf{x}, \mathbf{y}), \\ \dot{y}_k = w_k = \sum_{j=1}^{N^{\mathrm{F}}} m_j^F K^{\mathrm{F}}(y_k, x_j) + \sum_{\ell=1}^{N^{\mathrm{L}}} m_\ell^L K^{\mathrm{L}}(y_k, y_\ell) + \xi_k u_k^{\mathrm{opt}} + (1 - \xi_k) u_k^{\mathrm{self}}, \end{cases} \tag{1}$$

with initial data for followers $(x_i(0), v_i(0)) = (x_i^0, v_i^0)$ and leaders $(y_k(0), w_k(0)) = (y_k^0, w_k^0)$. The quantities m_i^F, m_k^L weight the interaction of followers and leaders, in what follows we will assume that $m_1^F = \ldots = m_{N^{\mathrm{F}}}^F = m_1^L = \cdots = m_{N^{\mathrm{L}}}^L$ and the following mass constraint holds

$$m_i^F = \frac{\rho^F}{N^{\mathrm{F}}}, \qquad m_k^L = \frac{\rho^L}{N^{\mathrm{L}}}, \qquad \rho^F + \rho^L = 1, \tag{2}$$

for ρ^F, ρ^L positive quantities.

1. S is a self-propulsion term, given by the relaxation toward a random direction or the relaxation toward a unit vector pointing to the target (the choice depends on the position), plus a term which translates the tendency to reach a given characteristic speed $s \geq 0$ (modulus of the velocity), i.e.,

$$S(x, v) := C_s(s^2 - |v|^2)v + \sum_{e=1}^{N_e} \psi_e(x) C_\tau \left(\frac{x_e^\tau - x}{|x_e^\tau - x|} - v \right), \tag{3}$$

where $\psi_e : \mathbb{R}^d \to [0, 1]$ is the characteristic function of Σ_e, and C_τ, C_s are positive constants.

2. The interactions follower-follower and follower-leader account a repulsion and an alignment component, as follows

$$\begin{aligned} H^{\mathrm{F}}(x, v, x', v'; \mathbf{x}, \mathbf{y}) &:= -C_r^{\mathrm{F}} R_{\gamma,r}(x, x')(x' - x) \\ &+ (1 - \psi(x))C_{al}^{\mathrm{F}} A(x, x'; \mathbf{x}, \mathbf{y})(v' - v), \\ H^{\mathrm{L}}(x, v, y, w; \mathbf{x}, \mathbf{y}) &:= -C_r^{\mathrm{L}} R_{\gamma,r}(x, y)(y - x) \\ &+ (1 - \psi(x))C_{al}^{\mathrm{L}} A(x, y; \mathbf{x}, \mathbf{y})(w - v), \end{aligned} \tag{4}$$

for given positive constants C_r^{F}, C_{al}^{F}, C_{al}^{L}, C_{at}, r, γ, and where the characteristic function of the unknown environment $\Omega \setminus \cup_e \Sigma_e$, such that

$$\psi(x) := \sum_{e=1}^{N_e} \psi_e(x).$$

The first term on the right hand side of (4) represents the metrical repulsion force, where the intensity is modulated by the function $R_{\gamma,r}$ defined as

$$R_{\gamma,r}(x, y) = \begin{cases} \frac{e^{-|y-x|^\gamma}}{|y-x|} & \text{if} y \in B_r(x)\setminus\{x\}, \\ 0 & \text{otherwise}, \end{cases} \tag{5}$$

where $B_r(x)$ is the ball of radius $r > 0$ centered at $x \in \Omega$. The second term accounts for the (topological) alignment force, which vanishes inside the visibility regions, and where

$$A(x, y; \mathbf{x}, \mathbf{y}) := \chi_{\mathcal{B}_{\mathcal{N}}(x;\mathbf{x},\mathbf{y})}(y), \tag{6}$$

and by $\mathcal{B}_{\mathcal{N}}(x; \mathbf{x}, \mathbf{y})$ the *minimal* ball centered at x encompassing at least $\mathcal{N}$ agents.

3. The interactions leader-follower and leader-leader reduce to a mere (metrical) repulsion, i.e., $K^{\mathrm{F}} = K^{\mathrm{L}} = -C_r^{\mathrm{L}} R_{\zeta,r}$, where $C_r^{\mathrm{L}} > 0$ and $\zeta > 0$ are in general different from C_r^{F} and γ, respectively.
4. $u_k^{\mathrm{opt}}, u_k^{\mathrm{self}} : \mathbb{R}^+ \to \mathbb{R}^{dN^{\mathrm{L}}}$ characterize the strategies of the leaders and are chosen in a set of admissible control functions. The parameter $\xi_k \in \{0, 1\}$ identifies for $\xi_k = 1$ leaders aware of their role, whose movements are the result of an optimization process, and alternatively for $\xi_k = 0$ leaders moving "selfishly" towards a specific exit. A specific description of leaders' strategy will be discussed in Sect. 4. Hence we account for situations where a small part of

the mass is informed about exit positions, but policymakers have no control over them.

Remark 1

- Differently from the model proposed in [4, 6] the dynamics do not include random effects. However, we consider this uncertainty by assuming that the initial velocity directions of followers are distributed according to a prescribed density $v_i^0 \sim p_v(\mathbb{R}^d)$, for example, a uniform distribution over the unitary sphere $\mathbb{S}_{d-1}$.
- The choice $C_{al}^{\mathrm{F}} = C_{al}^{\mathrm{L}}$ leads to $H^{\mathrm{F}} \equiv H^{\mathrm{L}}$ and, therefore, the leaders are not recognized by the followers as special. This feature opens a wide range of new applications, including the control of crowds not prone to follow authority's directives.
- The pedestrian microscopic model (1) allows agent movements in space without any constriction. However, in real applications, dynamics are constrained by walls or other kinds of obstacles. There are several ways of dealing with this feature in agent-based mode and we refer to [24, Sect. 2] for a review of obstacles handling techniques such as repulsive obstacle, rational turnaround, velocity cut-off. The choice for obstacle handling will be discussed in Sect. 4.

2.2 Control Framework for Pedestrian Dynamics

In order to define the strategies of optimized leaders, we formulate an optimal control problem to exploit the tendency of people to follow group mates in situations of emergency or doubt. The choice of a proper functional to be minimized constitutes a modeling difficulty, and it is typically a trade-off between a realistic task and a viable realization of its minimization. In general we will set up the following constrained optimal control problem

$$\begin{aligned} &\min_{\mathbf{u}^{\mathrm{opt}}(\cdot)\in U_{adm}} \mathcal{J}(\mathbf{u}^{\mathrm{opt}}), \\ &\quad \text{s.t.} \qquad (1), \end{aligned} \tag{7}$$

where $\mathbf{u}^{\mathrm{opt}} = (u_k^{opt}(\cdot))$ is the control vector associated with the optimized leaders, given a set of admissible controls U_{adm}. In what follows we will specify different functionals for different type of applications. For later convenience we introduced the empirical distributions defined as follows

$$f^{N^{\mathrm{F}}}(\cdot, x, v) = \sum_{i=1}^{N^{\mathrm{F}}} m_i^F \delta(x - x_i(\cdot))\delta(v - v_i(\cdot)), \tag{8}$$

$$g^{N^{\mathrm{L}}}(\cdot, x, v) = \sum_{j=1}^{N^{\mathrm{L}}} m_j^L \delta(x - y_j(\cdot))\delta(v - w_j(\cdot)). \tag{9}$$

- *Evacuation time.* In a situation where egressing pedestrians are in an unknown environment the most natural functional is the evacuation time, which we may define as follows

$$\mathcal{J}(\mathbf{x}, \mathbf{y}, \mathbf{u}^{\mathrm{opt}}) = \left\{ t > 0 \mid (x_i(t), y_j(t)) \notin \Omega \; \forall i = 1, \dots N^F, \forall j = 1, \dots, N^{\mathrm{L}} \right\}, \tag{10}$$

where we explicit the dependency on the states vector of follower positions $\mathbf{x} \in \mathbb{R}^{dN^{\mathrm{F}}}$. This cost functional is extremely irregular, therefore the search of minima is particularly difficult, additionally the evacuation of the total mass in some situations cannot be completely reached.
- *Total mass with multiple exits.* Instead of minimizing the total evacuation, we fix a final time $T > 0$ and we aim to minimize the total mass inside the computational domain $\Omega \setminus \cup_e \Sigma_e$, which coincides with maximizing the mass inside the visibility areas. The functional reads

$$\mathcal{J}(\mathbf{x}, \mathbf{y}, \mathbf{u}^{\mathrm{opt}}) = \int_{\mathbb{R}^d} \int_{\Omega \setminus \cup_e \Sigma_e} (f^{N^{\mathrm{F}}}(T, x, v) + g^{N^{\mathrm{L}}}(T, x, v)) dx dv. \tag{11}$$

- *Optimal mass splitting over multiple exits.* In complex environments, it may happen that total mass does not distribute in an optimal way between the target exits. This may lead to problems of heavy congestions and overcrowding around the exits that, in real-life situations, can cause injuries due to over-compression and suffocation. Hence we ask to distribute the total evacuated mass at final time T among the exits according to a given desired distribution. To this end we set

$$\mathcal{J}(\mathbf{x}, \mathbf{y}, \mathbf{u}^{\mathrm{opt}}) = \sum_{e=1}^{N_e} \left| \mathcal{M}_e^F(T) - \mathcal{M}_e^{\mathrm{des}} \right|^2, \tag{12}$$

where $\mathcal{M}_e^{\mathrm{des}}$ is the desired mass to be reached in the visibility area Σ_e and $\mathcal{M}_e^F(T)$ is the total mass of followers and leaders who reached exit x_e^τ up to final time T.

3 Mean-Field Approximation of Follower-Leader System

Mean-field scale limit for large number of interacting individuals has been investigated in several directions for single and multiple population dynamics, see, for example, [19, 31], and it is a fundamental step to tame the curse of dimensionality arising for coupled systems of ODEs.

In the current setting, we want to give a statistical description of the followers-leaders dynamics considering a continuous density for followers and maintaining leaders microscopic. Hence, we introduce the non-negative distribution function of followers $f = f(t, x, v)$ with $x \in \mathbb{R}^d$, $v \in \mathbb{R}^d$ at time $t \geq 0$, the meso-micro system corresponding to (1) reads as follows

$$\begin{aligned} \partial_t f + v \cdot \nabla_x f &= -\nabla_v \cdot \left(f \left(S(x, v) + \mathcal{H}^F[f, g^{N^{\mathrm{L}}}] + \mathcal{H}^L[f, g^{N^{\mathrm{L}}}] \right) \right), \\ \dot{y}_k = w_k &= \int_{\mathbb{R}^{2d}} K^{\mathrm{F}}(y_k, x) f(t, x, v)\, dx\, dv + \sum_{\ell=1}^{N^{\mathrm{L}}} m_\ell^L K^{\mathrm{L}}(y_k, y_\ell) \\ &\quad + \xi_k u_k^{\mathrm{opt}} + (1 - \xi_k) u_k^{\mathrm{self}}, \end{aligned} \tag{13}$$

where the followers dynamics is described by a kinetic equation of Vlasov-type, and where we use the corresponding empirical distribution for leaders $g^{N^{\mathrm{L}}}$. Furthermore we assume that the follower and leader densities are such that their number densities are

$$\varrho^F = \int_{\mathbb{R}^{2d}} f(t, x, v)\, dx\, dv, \qquad \varrho^L = \int_{\mathbb{R}^{2d}} g^{N^{\mathrm{L}}}(t, x, v)\, dx\, dv.$$

We observe that the terms $S(\cdot)$, $K^{\mathrm{F}}(\cdot)$ and $K^{\mathrm{L}}(\cdot)$ are defined, respectively, as in the microscopic setting, whereas the non-local operators $\mathcal{H}^F$, $\mathcal{H}^L$ correspond to the following integrals

$$\begin{aligned} \mathcal{H}^F[f, g^{N^{\mathrm{L}}}](t, x, v) &= -C_r^{\mathrm{F}} \int_{\mathbb{R}^d} \int_{B_r(x)} R_{\gamma,r}(x, x')(x' - x) f(t, x', v')\, dx' dv' \\ &\quad + C_{al}^{\mathrm{F}}(1 - \psi(x)) \int_{\mathbb{R}^d} \int_{\mathcal{B}_{r_*}(t,x)} (v' - v) f(t, x', v')\, dx' dv', \end{aligned} \tag{14}$$

$$\begin{aligned} \mathcal{H}^L[f, g^{N^{\mathrm{L}}}](t, x, v) &= -C_r^{\mathrm{L}} \int_{\mathbb{R}^d} \int_{B_r(x)} R_{\gamma,r}(x, x')(x' - x) g^{N^{\mathrm{L}}}(t, x', v')\, dx' dv' \\ &\quad + C_{al}^{\mathrm{L}}(1 - \psi(x)) \int_{\mathbb{R}^d} \int_{\mathcal{B}_{r_*}(t,x)} (v' - v) g^{N^{\mathrm{L}}}(t, x', v')\, dx' dv', \end{aligned} \tag{15}$$

where the first term corresponds to the metrical repulsion as in (5), and the second part accounts the topological ball $\mathcal{B}_{r_*}(t, x) \equiv \mathcal{B}_{r^*}(t, x; f, g^{N^{\mathrm{L}}})$ whose radius is defined for a fixed $t \geq 0$ by the following variational problem

$$r^*(t, x) = \arg\min_{\alpha > 0} \left\{ \int_{\mathbb{R}^d} \int_{B_\alpha(x)} \left(f(t, x, v) + g^{N^{\mathrm{L}}}(t, x, v) \right) dx\, dv \geq \varrho_{\mathrm{top}} \right\}, \tag{16}$$

where $\varrho_{\mathrm{top}} > 0$ is the target topological mass.

Remark 2

- Rigorous derivation of the mean-field limit (13) from (1) is a challenging task due to the strong irregularities induced by the behavior of topological-type interactions. We refer to [42] for possible regularization in the case of Cucker-Smale type dynamics, and to [13, 30] for alignment driven by jump-type processes.
- Alternative derivation of mesoscopic models in presence of diffusion has been obtained in [4], where the authors derived a Fokker-Planck equation of the original microscopic system via quasi-invariant scaling of binary Boltzmann interactions. This technique, analogous to the so-called grazing collision limit in plasma physics, has been thoroughly studied in [51] and allows to pass from a Boltzmann description to the mean-field limit, see, for example, [49].
- For optimal control of large interacting agent systems, the derivation of a mean-field approximation involves the convergence of minimizers from microscopic to mesoscopic scale. This problem has been addressed from different directions, and we refer to [15, 38].

Remark 3 In order to obtain a closed hydrodynamic system for (13) a standard assumption is to assume the velocity distribution to be mono-kinetic, i.e. $f(t, x, v) = \rho(t, x)\delta(v - V(t, x))$, and the fluctuations to be negligible. Hence, computing the moments of (13) leads to the following macroscopic system for the density ρ and the bulk velocity V,

$$\begin{cases} \partial_t \rho + \nabla_x \cdot (\rho V) = 0, \\ \partial_t(\rho V) + \nabla_x \cdot (\rho V \otimes V) = \mathcal{G}_m \left[\rho, \rho^{\mathrm{L}}, V, V^{\mathrm{L}}\right] \rho, \\ \dot{y}_k = w_k = \displaystyle\int_{\mathbb{R}^d} K^{\mathrm{F}}(y_k, x)\rho(t, x)\, dx + \sum_{\ell=1}^{N^{\mathrm{L}}} K^{\mathrm{L}}(y_k, y_\ell) \\ \qquad\qquad + \xi_k u_k^{\mathrm{opt}} + (1 - \xi_k) u_k^{\mathrm{self}}, \end{cases} \tag{17}$$

where $\rho^{\mathrm{L}}(x, t)$, $V^{\mathrm{L}}(x, t)$ represent the leaders macroscopic density and bulk velocity, respectively, and $\mathcal{G}_m$ the macroscopic interaction operator associated with the followers, we refer to [8, 19] for further details.

3.1 MFMC Algorithms

For the numerical solution of the mean-field followers dynamics in (13) we employ mean-field Monte-Carlo methods (MFMCs) generalizing the approaches proposed in [7, 49]. These methods fall in the class of fast algorithms developed for interacting particle systems such as direct simulation Monte-Carlo methods (DSMCs), and they are strictly related to more recent class of algorithms named Random Batch Methods (RBMs) [45].

In order to approximate the evolution of the followers density, first we sample N_s^F particles from the initial distribution $f^0(x, v)$ in the phase space, i.e., $\{(x_i^0, v_i^0)\}_{i=1}^{N_s^F}$. Furthermore we consider a subsample of M particles, $j_1, \ldots, j_M$ uniformly without repetition such that $1 \le M \le N_s^F$. In order to approximate the non-local terms $\mathcal{H}^F, \mathcal{H}^L$ we evaluate the interactions with a subsample of size M at every time step. Hence we define the discretization step as

$$v_i^{n+1} = v_i^n + \Delta t S(x_i^n, v_i^n) - \Delta t \left[\hat{R}_i^{F,n}(\hat{X}_i^n - x_i^n) + \hat{R}_i^{L,n}(\hat{Y}_i^n - x_i^n) \right] \tag{18}$$

$$+ \Delta t (1 - \psi(x_i^n)) \left[\hat{A}_i^{F,n}(\hat{V}_i^n - v_i^n) + \hat{A}_i^{L,n}(\hat{W}_i^n - v_i^n) \right], \tag{19}$$

where we defined the following auxiliary variables for the repulsion term (5),

$$\begin{aligned} \hat{R}_i^{F,n} &= \frac{C_r^{\mathrm{F}} \varrho^F}{M} \sum_{k=1}^{M} R_{\gamma,r}(x_i^n, x_{j_k}^n), \quad \hat{X}_i^n = \frac{C_r^{\mathrm{F}} \varrho^F}{M} \sum_{k=1}^{M} \frac{R_{\gamma,r}(x_i^n, x_{j_k}^n)}{\hat{R}_i^{F,n}} x_{j_k}^n, \\ \hat{R}_i^{L,n} &= \frac{C_r^{\mathrm{L}} \varrho^L}{N^{\mathrm{L}}} \sum_{\ell=1}^{N^{\mathrm{L}}} R_{\gamma,r}(x_i^n, y_\ell^n), \quad \hat{Y}_i^n = \frac{C_r^{\mathrm{L}} \varrho^L}{N^{\mathrm{L}}} \sum_{k=1}^{N^{\mathrm{L}}} \frac{R_{\gamma,r}(x_i^n, y_\ell^n)}{\hat{R}_i^{L,n}} y_\ell^n. \end{aligned} \tag{20}$$

For the topological alignment we have

$$\begin{aligned} \hat{A}_i^{F,n} &= \frac{C_{al}^{\mathrm{F}} \varrho^F}{M} \sum_{k=1}^{M} \chi_{\mathcal{B}_{r_M^*}(x_i;\mathbf{x},\mathbf{y})}(x_{j_k}), \quad \hat{V}_i^n = \frac{C_{al}^{\mathrm{F}} \varrho^F}{M} \sum_{k=1}^{M} \frac{\chi_{\mathcal{B}_{r_M^*}(x_i;\mathbf{x},\mathbf{y})}(x_{j_k})}{\hat{A}_i^{F,n}} v_{j_k}^n, \\ \hat{A}_i^{L,n} &= \frac{C_{al}^{\mathrm{L}} \varrho^L}{N^{\mathrm{L}}} \sum_{\ell=1}^{N^{\mathrm{L}}} \chi_{\mathcal{B}_{r_M^*}(x_i;\mathbf{x},\mathbf{y})}(x_{j_k}), \quad \hat{W}_i^n = \frac{C_{al}^{\mathrm{L}} \varrho^L}{N^{\mathrm{L}}} \sum_{k=1}^{N^{\mathrm{L}}} \frac{\chi_{\mathcal{B}_{r_M^*}(x_i;\mathbf{x},\mathbf{y})}(x_{j_k})}{\hat{A}_i^{L,n}} w_\ell^n, \end{aligned} \tag{21}$$

where the topological ball $\mathcal{B}_{r_M^*}(x)$ is the topological ball defined over the subsample of M agents, with radius such that

$$r_M^*(t, x_i) = \arg\min_{\alpha>0} \left\{ \frac{\varrho^F}{M} \sum_{k=1}^{M} \chi_{B_\alpha(x_i)}(x_{j_k}) + \frac{\varrho^L}{N^{\mathrm{L}}} \sum_{\ell=1}^{N^{\mathrm{L}}} \chi_{B_\alpha(x_i)}(y_\ell) \ge \varrho_{\mathrm{top}} \right\}. \tag{22}$$

From the above considerations we obtain the following Algorithm in the time interval $[0, T]$.

Algorithm (MFMC Follower-Leader)

1. Given N_s^F samples v_i^0, with $i = 1, \ldots, N_s^F$ computed from the initial distribution $f(x, v)$ and $M \leq N_s^F$;
2. `for` $n = 0$ `to` n_{tot}
   a. `for` $i = 1$ `to` N_s^F
      i. sample M particles $j_1, \ldots, j_M$ uniformly without repetition among all particles;
      ii. compute the quantities $\hat{R}_i^{L,n}$, $\hat{R}_i^{F,n}$, $\hat{X}_i^n$ and $\hat{Y}_i^n$ from (20);
      iii. compute the quantities $\hat{A}_i^{L,n}$, $\hat{A}_i^{F,n}$, $\hat{V}_i^n$ and $\hat{W}_i^n$ from (21);
      iv. compute the velocity change v_i^{n+1} according to (18);
      v. compute the position change

$$x_i^{n+1} = x_i^n + \Delta t v_i^{n+1}.$$

   `end for`

`end for` □

Remark 4

- By using this Monte-Carlo algorithm we can reduce the computational cost due to the computation of the interaction term from the original $O\left({N_s^F}^2\right)$ to $O\left(M N_s^F\right)$. For $M = N_s^F$ we obtain the explicit Euler scheme for the original N_s^F particle system.

4 Numerical Optimization of Leaders Strategies

In this section we focus on the numerical realization of the general optimal control problem of type

$$\min_{\mathbf{u}^{\text{opt}}(\cdot) \in U_{adm}} \mathcal{J}(\mathbf{u}^{\text{opt}}), \tag{23}$$

constrained to the evolution of microscopic (1) or mean-field system (13). We observe that the minimization task for evacuation time or total mass can be extremely difficult, due to the strong irregularity and the presence of many local minima.

In order to optimize (23) we propose instead an alternative suboptimal, but computationally efficient strategy, named modified Compass Search (CS). This method falls in the class of metaheuristic algorithms, it ensures the convergences towards local minima, without requiring any regularity of the cost functional [10].

We use the CS method in order to optimize the trajectory of the *aware* leaders. The idea is to start from an initial guess $u_k^{\text{opt},(0)}$ which produces an admissible trajectory toward a target exit, for example, as follows

$$u^{\text{opt}},\ (0)_k(t) = \beta \frac{\Xi_k(t) - y_k(t)}{\|\Xi_k(t) - y_k(t)\|} + (1-\beta)(m_F(t) - y_k(t)), \tag{24}$$

where $\Xi_k(t)$ is the target position at time t, depending on the environment and such that $\Xi_k(t) = x_e^\tau$ for $t > t_*$. The parameter $\beta \in [0, 1]$ measures the tendency of leaders to move toward the target $\Xi_k(t)$ or staying close to followers center of mass $m_F(t)$.

We will refer to (24) as "go-to-target" strategy. Then CS method iteratively modifies the current best control strategy found so far computing small random piecewise constant variation of points on the trajectories. Then, if the cost functional decreases, the variation is kept, otherwise it is discarded. We consider piecewise constant trajectories, introducing suitable switching times for the leaders controls.

We summarize this procedure in the following algorithm.

Algorithm (Modified Compass Search)

1. Select a discrete set of sample times $S_M = \{t_1, t_2, \ldots, t_M\}$, the parameters $j = 0$, $j_{\texttt{max}}$ and J_E.
2. Select an initial strategy u^* piecewise constant over the set S_M, e.g., constant direction and velocity speed towards a fixed target $\Xi_k(t)$, $k = 1, \ldots, N^L$, see Equation (24). Compute the functional $\mathcal{J}(\mathbf{x}, \mathbf{u}^*)$.
3. Perform a perturbation of the trajectories over a fixed set of points $P^*(t)$ on current optimized leader trajectories with small random variations over the time-set S_M,

$$P^{(j)}(t_m) = P^*(t_m) + B_m, \quad m = 1, \ldots, M, \tag{$\mathcal{P}$}$$

where $B_m \sim \text{Unif}([-1, 1]^d)$ is a random perturbation and set for $m = 1, \ldots, M$,

$$u^{\text{opt},(j)}(t) = \frac{P^*(t_{m+1}) - P^*(t_m)}{\|P^*(t_{m+1}) - P^*(t_m)\|}, \quad t \in [t_m, t_{m+1}].$$

Finally compute $\mathcal{J}(\mathbf{x}, \mathbf{u}^{(j)})$.

4. `while` $j < j_{\texttt{max}}$ AND $\mathcal{J}(\mathbf{x}, \mathbf{u}^*) < \mathcal{J}_E$
   a. Update $j \leftarrow j + 1$.
   b. Perform the perturbation ($\mathcal{P}$) and compute $\mathcal{J}(\mathbf{x}, \mathbf{u}^{(j)})$.
   c. `If` $\mathcal{J}(\mathbf{x}, \mathbf{u}^{(j)}) \le \mathcal{J}(\mathbf{x}, \mathbf{u}^*)$
   set $u^* \leftarrow u^{(j)}$ and $\mathcal{J}(\mathbf{x}, \mathbf{u}^*) \leftarrow \mathcal{J}(\mathbf{x}, \mathbf{u}^{(j)})$.

`repeat` □

Remark 5

- Compass search does not guarantee the convergence to a global minimizer, on the other hand it offers a good compromise in terms of computational efficiency.
- Alternative metaheuristic schemes can be employed to enhance leader trajectories and improve the convergence towards the global minimizer, among several possibilities we refer to genetic algorithms, and particle swarm based optimizations.
- The synthesis of control strategies via compass search for the microscopic and the mean-field dynamics can produce different results, due to the strong non-linearities of the interactions, and the non-convexity of the functional considered, such as the evacuation time. However, in any case, the solutions retrieved by this approach satisfy local optimality criteria by construction.

4.1 Numerical Experiments

We present three different numerical experiments at microscopic and mesoscopic levels, corresponding to the minimization of cost functionals presented in Sect. 2.2.

Numerical Discretization The dynamics at microscopic level is discretized by a forward Euler scheme with a time step $\Delta t = 0.1$, whereas the evolution of the mean-field dynamics is approximated by MFMCs algorithms. We choose a sample of $\mathcal{O}(10^3)$ particles for the approximation of the density and we reconstruct their evolution in the phase space by kernel density estimator with a multivariate standard normal density function with bandwidth $h = 0.4$. Table 1 reports the parameters of the model for the various scenarios unchanged for every test. The number of leaders instead changes and it will be specified later.

Obstacles Handling In order to deal with obstacles we use a *cut-off velocity* approach, namely we compute the velocity field first neglecting the presence of the obstacles, then nullifying the component of the velocity vector which points inside the obstacle. This method is used in, e.g., [4, 26, 28] and requires additional conditions to avoid situations where pedestrians stop walking completely because both components of the velocity vector vanish, e.g., in presence of corners, or when obstacles are very close to each other. We refer to [24] for more sophisticated approaches of obstacles handling.

Table 1 Model parameters for the different scenarios

N^F	$\mathcal{N}$	C_r^F	C_r^L	C_a^L	C_a^F	C_τ	C_s	s^2	$r \equiv \zeta$	γ
150	20	2	1.5	3	3	1	0.5	0.4	1	1

4.1.1 Test 1: Minimum Time Evacuation with Multiple Exits

In this first test, leaders aim to minimize the time of evacuation (10), hence trying to enforce crowd towards the exit avoiding congestion and ease the outflux of the pedestrian. We assume that leaders informed about exits position follow 'go-to-target' strategy defined as in (24), where the target is defined by the different exits and will be specified for each leader. In what follows we account for two different settings comparing microscopic and mesoscopic dynamics.

Setting a) Three Exits

We consider the case of a room with no obstacles and three exits located at $x_1^\tau = (35, 10)$, $x_2^\tau = (16, 20)$, $x_3^\tau = (10, 10)$ with visibility areas $\Sigma_e = \left\{x \in \mathbb{R}^2 : |x - x_e^\tau| < 5\right\}$. We consider two different types of leaders; we call selfish leaders y^{self} the agents who do not care about followers and follow the direction that connects their positions to the exits. While the optimized leaders y^{opt} are aware of their role and they move with the aim to reach the exits and to maintain contact with the crowd, only the trajectories of this type of leaders will be optimized. The admissible leaders trajectories are defined as in Eq. (24), we choose $\beta = 1$ for selfish leaders, $\beta = 0.6$ for optimized leaders and the target position as $\Xi_k(t) = x_e^\tau$ $\forall t$ and for every leader k. At initial time leaders and followers are uniformly distributed in the domain $[17, 29] \times [6.5, 13.5]$ where followers velocities are sampled from a normal distribution with average -0.5 and variance 0.1, hence biased towards the wrong direction. We report in Fig. 1 the initial configuration for both microscopic and mesoscopic dynamics.

Microscopic Case We consider $N^L = 9$ leaders, three optimized and six unaware leaders. Each leader is associated with an exit: unaware leaders move towards the nearest exit, whereas each optimized leader is assigned to a different exit.

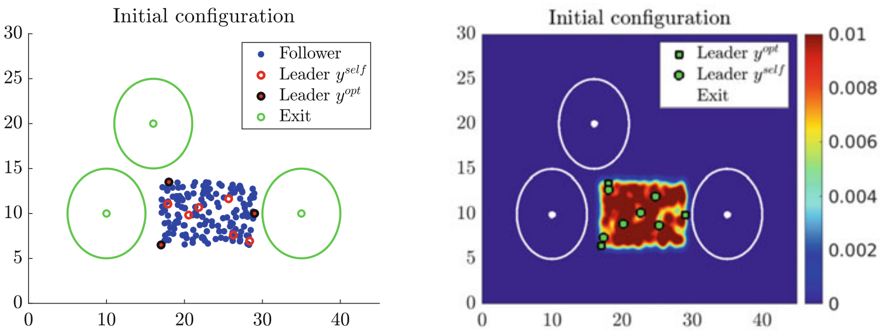


Fig. 1 *Test 1a*. Minimum time evacuation with multiple exits, initial configuration

Figure 2 shows the evolution of the agents with the go-to-target strategy on the left and with the optimal strategy obtained by the compass search algorithm on the right. As it can be seen in Fig. 2 with the go-to-target strategy the whole crowd reaches the exit, after 850 time steps. We distinguish optimized leaders y^{opt} with a dashed black line. Optimized movements for leaders are retrieved by means of Algorithm 4, with initial guess go-to-target strategy, we report in Fig. 3 the decrease of the performance function (10) as a function of the iterations of compass search. Eventually optimized leaders influence the crowds for a larger amount of time and the total mass is evacuated after 748 time steps, as shown in Table 2. Figure 4 compares the evacuated mass and the occupancy of the exits visibility zone as a function of time for the uncontrolled case, the go-to-target strategy and the optimal compass search strategy. Dashed lines indicate times of total mass evacuation.

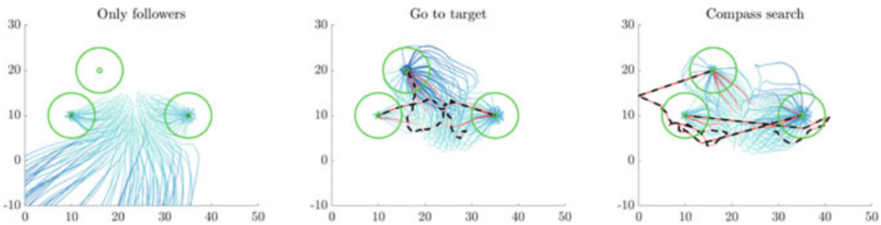


Fig. 2 *Test 1a*. Microscopic case: minimum time evacuation with multiple exits. On the left the uncontrolled case, in the center the go-to-target and on the right the optimal compass search strategy

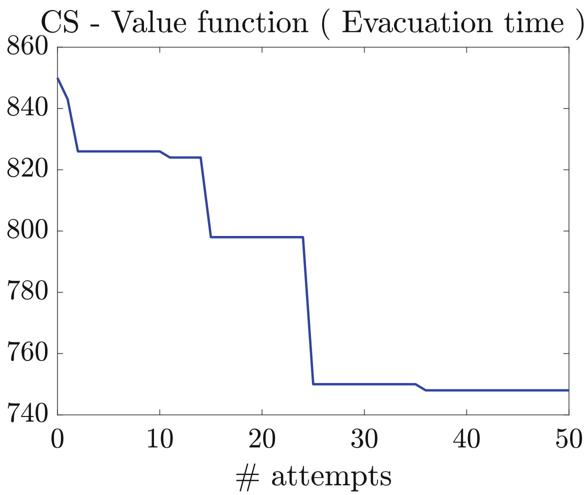


Fig. 3 *Test 1a*. Microscopic case: decrease of the value function (10) as a function of compass search iteration

Table 2 *Test 1a*. Performance of leader strategies over microscopic dynamics

	Uncontrolled	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 1000	850	748
Evacuated mass (percentage)	46%	100%	100%

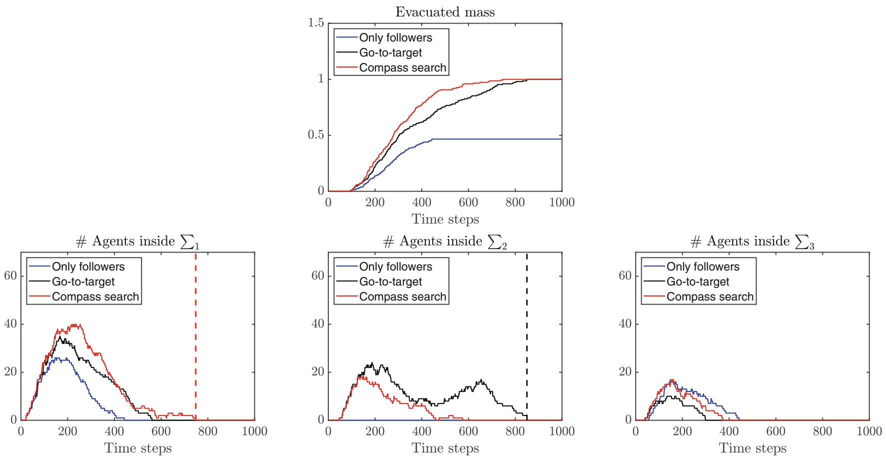


Fig. 4 *Test 1a.* Microscopic case: minimum time evacuation with multiple exits. Evacuated mass (first row), occupancy of the visibility area Σ_1 (second row, left), Σ_2 (second row, center), and Σ_3 (second row, right) as a function of time for uncontrolled, go-to-target and optimal compass search strategies. The dot line denotes the time step in which the whole mass is evacuated, the line is black for the go-to-target and red for the optimal compass search strategy

Table 3 *Test 1a*. Performance of leader strategies over mesoscopic dynamics

	Uncontrolled	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 1000	> 1000	897
Evacuated mass (percentage)	84%	99%	100%

Mesoscopic Case We consider now a continuous density of followers, in the same setting of the previous microscopic case: we account for $N^L = 9$ microscopic leaders moving in a room with no obstacles and three exits. Hence we compare uncontrolled dynamics, go-to-target strategies, and optimized strategies with compass search. In Table 3 we show that without any control followers are unable to reach the total evacuation reaching 84% of total mass evacuated. Go-to-target strategy improves total mass evacuated, however, a small part of the mass spreads around the domain and is not able to reach the target exit. Eventually, with optimized strategies, we reach the evacuation of the total mass in 897 simulation steps. The better performance of the optimized strategy can be observed directly from Fig. 5, where functional (10) is evaluated at subsequent iterations of Algorithm 4. In Fig. 6 we show three snapshots of the followers density comparing leaders with different strategies and the uncontrolled case. In the upper row, we report the evolution

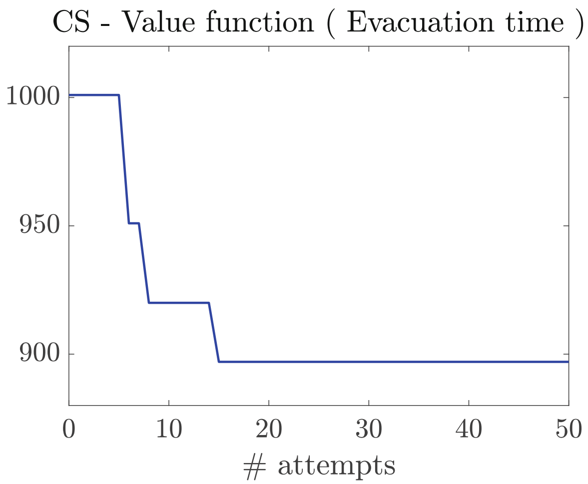


Fig. 5 *Test 1a.* Mesoscopic case: minimum time evacuation with multiple exits. Decrease of the value function (10) as a function of attempts

without any control. The middle row shows leaders driven by a go-to-target strategy promoting evacuation of followers density. At time $t = 50$ leaders are moving to influence the followers towards the three exits. At time $t = 100$, the followers mass splits and starts to reach the exits. At time $t = 1000$, complete evacuation is almost reached. The bottom row depicts improved strategies of leaders, where total mass is evacuated at time step 912.

Finally in Fig. 7 we summarize the results showing the evacuated mass as the cumulative distribution of agents who reached the exit, and the occupancy of the visibility areas in terms of total mass percentage for the various exits. Dashed red line indicates time of complete evacuation.

Setting b) Two Exits in a Closed Environment

Assume now to have a room with walls that contains two exits, $x_1^\tau = (50, 0)$ and $x_2^\tau = (30, 50)$. Followers are uniformly distributed in $[0, 10] \times [0, 10]$. Assume that initially two unaware leaders y^{self} move towards exit x_1^τ with selfish strategy, i.e., $\beta = 0$ in (24). Hence the goal is to minimize the total evacuation time as reported in (10) introducing two additional leaders y^{opt} moving towards exit x_2^τ, for these two leaders we choose the parameter $\beta = 0.6$ in (24). The target position is $\Xi_k(t) = x_e^\tau \; \forall t$ and for every leader k. Figure 8 shows the initial configuration in the microscopic and mesoscopic case, and with an initial position of $N^L = 4$ unaware and aware leaders.

Microscopic Case In Fig. 9 we report the crowd's evolution in various scenarios: left plot shows the trajectories where only unaware leaders are present, in this case,

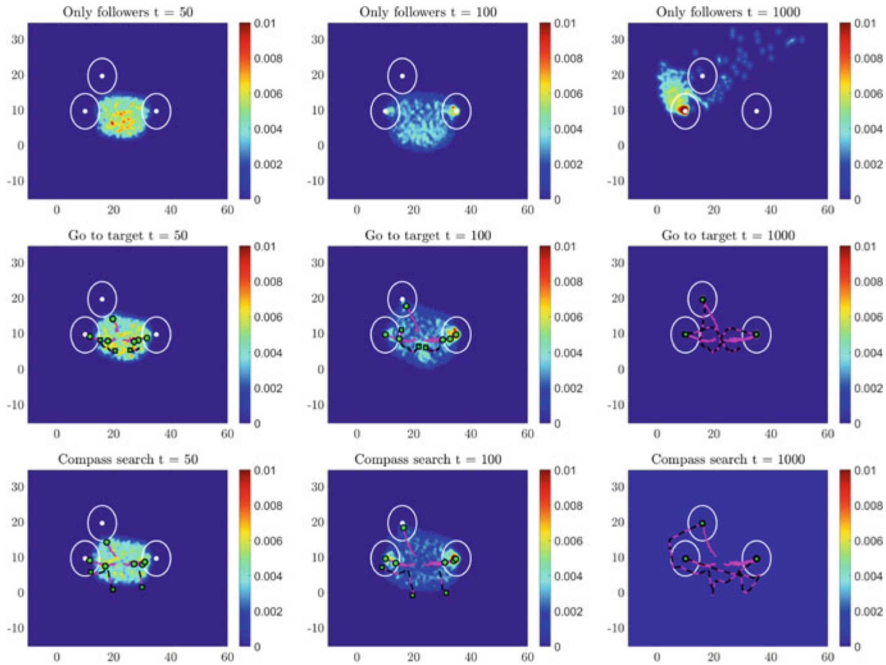


Fig. 6 *Test 1a.* Three snapshots taken at time $t = 50$, $t = 100$, $t = 1000$ of the mesoscopic densities for the minimum time evacuation with multiple exits. In the upper row the uncontrolled case, in the central row the three aware leaders, follows a go-to-target strategy, whereas in the bottom row their trajectories are optimized according to CS algorithm

the whole crowd reaches the exit x_1^τ; central and right plots show the influence of two aware leaders moving to x_2^τ, respectively, with fixed and optimized strategies. Unaware leaders influence the whole crowd to move towards the exit, however, generating overcrowding at x_1^τ and leaving some agents getting lost. Introducing two aware leaders with fixed strategies the whole mass is evacuated in 1966 time steps, with optimized strategies evacuation time is further reduced to 1199 time steps. In these last cases, the mass is split between the two exits and hence overcrowding phenomena are reduced. In Table 4 the total evacuation time and the corresponding evacuated mass for the three scenarios are reported, where we indicate that optimized strategy is obtained after 50 iterations of compass search. Finally, in Fig. 10 we report the occupancy of the visibility areas and the cumulative distribution of the mass evacuate as a function of time for the various scenarios.

Mesoscopic Case We consider now the mean-field approximation of the microscopic setting. We report in Fig. 11 three snapshots of followers density and trajectories of leaders, for each scenario. In this case, unaware leaders moving selfishly towards exit x_1^τ are able to influence followers and evacuate 81% at final time, whereas the rest of the mass is congested around the exit. Introducing

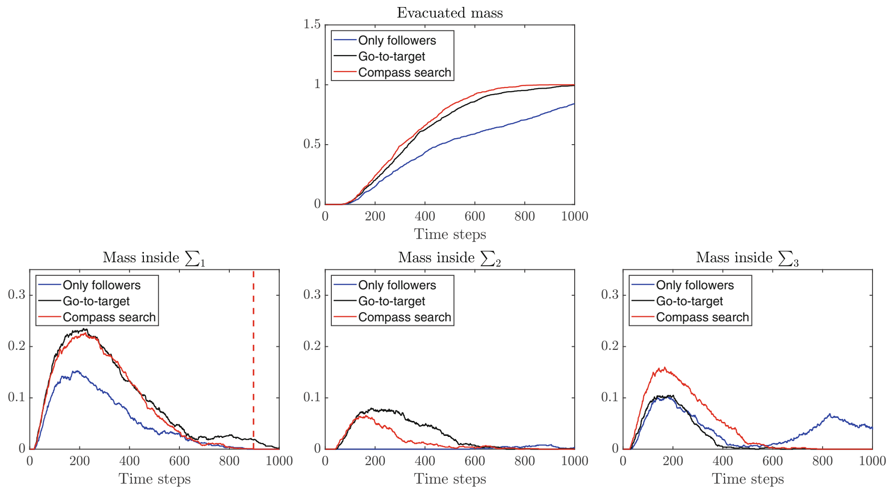


Fig. 7 *Test 1a*. Mesoscopic case: minimum time evacuation with multiple exits. Evacuated mass (first row), occupancy of the visibility area Σ_1 (second row, left), Σ_2 (second row, center) and Σ_3 (second row, right) as a function of time for go-to-target and optimal compass search strategies. The dot line denotes the time step in which the whole mass is evacuated with the optimal compass search strategy

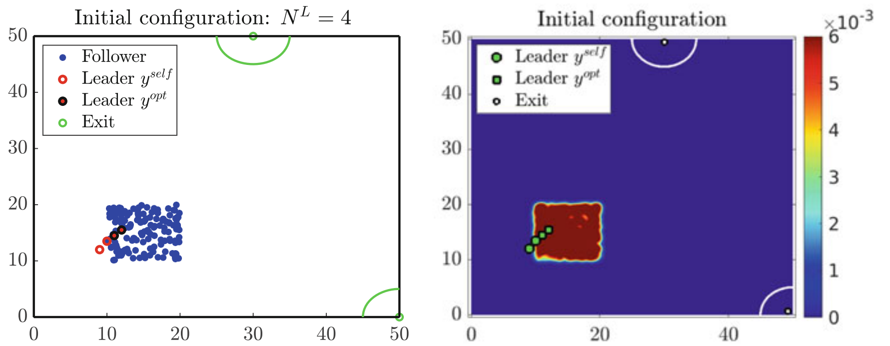


Fig. 8 *Test 1b*. Minimum time evacuation with multiple exits and obstacles, initial configuration for microscopic and mesoscopic case

two aware leaders with a fixed strategy toward x_2^τ is not sufficient to reach total evacuation at final time which is and at final time 95% of the mass is evacuated. The bottom row depicts the case with optimized leaders strategies, in this case, the total mass is evacuated at time step 1750. We summarize the performances of the results in Table 5, and in Fig. 12 we report the occupancy of the visibility areas and the cumulative distribution of mass evacuated as a function of time.

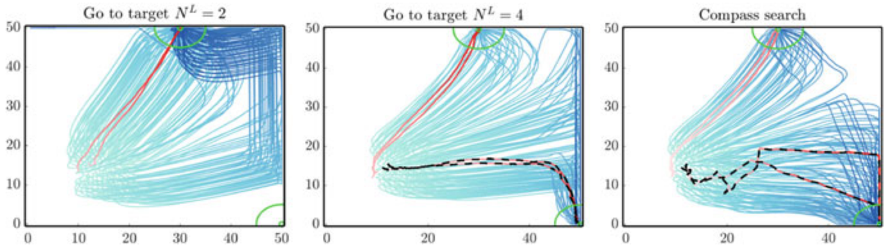


Fig. 9 *Test 1b*. Microscopic case: minimum time evacuation with multiple exits and obstacles. Go-to-target $N^L = 2$ (left), go-to-target $N^L = 4$ (center), optimal compass search (right)

Table 4 *Test 1b*. Performance of leader strategies over microscopic dynamics

	Go-to-target $N^L = 2$	Go-to-target $N^L = 4$	CS (50 it)
Evacuation time (time steps)	>2000	1966	1199
Evacuated mass (percentage)	99%	100%	100%

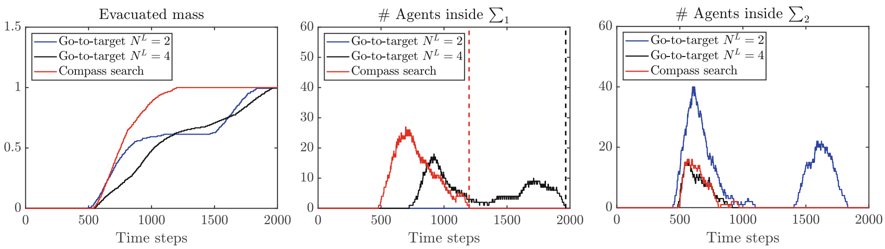


Fig. 10 *Test 1b*. Microscopic case: minimum time evacuation with multiple exits and obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies. The black and red dot lines denote the time step in which the whole mass is evacuated with the go-to-target ($N^L = 4$) and optimal compass search strategy, respectively

4.1.2 Test 2: Mass Evacuation in Presence of Obstacles

We consider two rooms, one inside the other, where the internal room is limited by three walls while the external one is bounded by four walls. We assume that walls are nonvisible obstacles, i.e., people can perceive them only by physical contact. This corresponds to an evacuation in case of null visibility (but for the exit points which are still visible from within Σ_1 and Σ_2). Consider the case of two exits, $x_1^\tau = (2, 78)$ and $x_2^\tau = (45, 2)$ positioned in the external room. Figure 13 provides a description of the initial configuration. Note that in order to evacuate, people must first leave the inner room, in which they are initially confined, and then search for exits. Evacuation in presence of obstacles is not always feasible. Instead of minimizing the total evacuation time as in Sect. 4.1.1, we aim to minimize the total mass inside the domain as reported in (11) and hence to maximize the total evacuated mass.

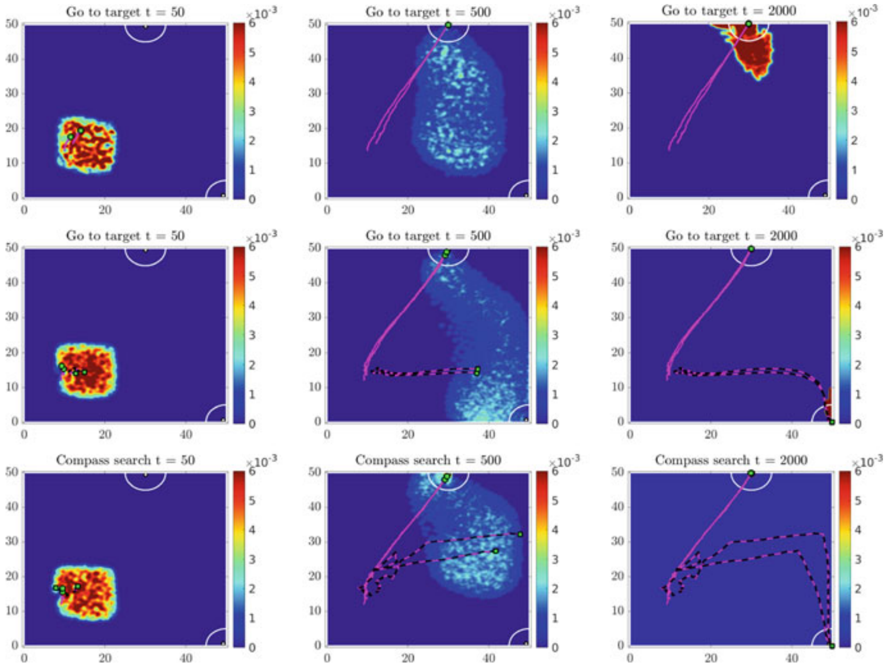


Fig. 11 *Test 1b.* Mesoscopic case: minimum time evacuation with multiple exits and obstacles. Three snapshots taken at time $t = 50$, $t = 500$, $t = 2000$ with the go-to-target strategy in the case $N^L = 2$ (upper row), $N^L = 4$ (central row) and with the optimized compass search strategy (lower row)

Table 5 *Test 1b.* Performance of leader strategies over mesoscopic dynamics

	Go-to-target $N^L = 2$	Go-to-target $N^L = 4$	CS (50 it)
Evacuation time (time steps)	>2000	>2000	1750
Evacuated mass	81%	95%	100%

Each leader will move toward one of the exits following a go-to-target, similar to (24), and such that it is admissible for the configuration of the obstacles. We choose $\beta = 1$ for every leader. The target position is $\Xi_k(t) = x_e^\tau$ for $t > t_*$, while for $t < t_*$ we consider one intermediate point in order to let the leaders to evacuate the inner room.

Microscopic Case We consider $N^L = 6$ leaders, with two aware leaders. Initially, followers have zero velocity. Three leaders, only one aware, will move towards exit x_1^τ, and the remaining towards exit x_2^τ. We report in Fig. 14 the evolution with the go-to-target strategy on the left, and with optimized strategies for the two aware leaders on the right. With go-to-target strategy leaders first leave the room and then move towards the exits. Since leaders move rapidly towards the exits, their influence over followers vanishes after a certain time. Indeed, part of the followers hits the

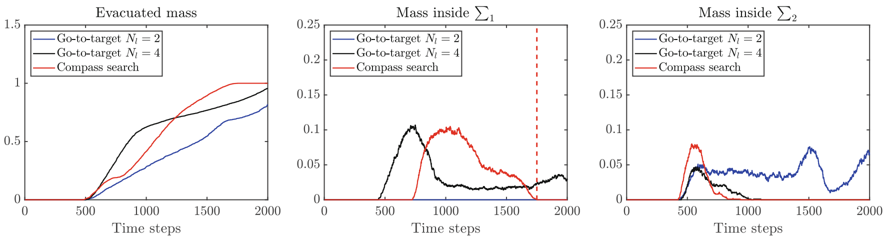


Fig. 12 *Test 1b.* Mesoscopic case: minimum time evacuation with multiple exits and obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies. The red dot line denotes the time step in which the whole mass is evacuated with the optimal compass search strategy

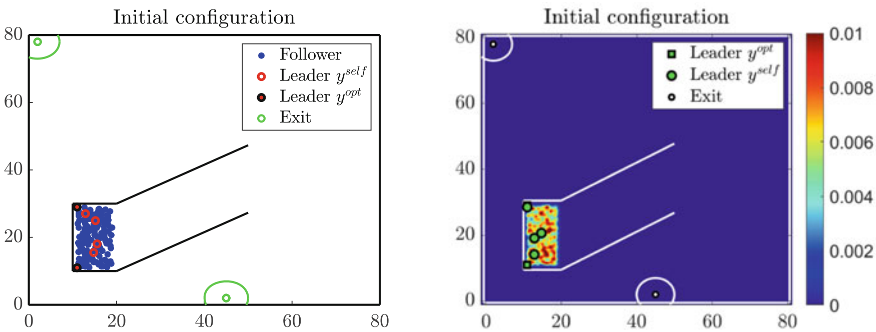


Fig. 13 *Test 2.* Maximization of mass evacuated in presence of obstacles, initial configuration

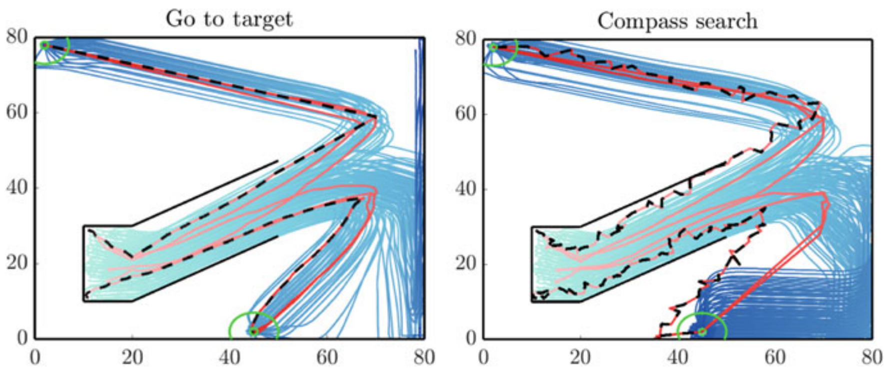


Fig. 14 *Test 2.* Microscopic case: mass maximization in presence of obstacles. On the left, go-to-target. On the right, optimal compass search

right boundary wall and does not reach the exits. Instead, with optimized strategies, leaders are slowed down, as consequence followers are influenced by leaders for a larger amount of time. Table 6 reports the comparison between two strategies

Table 6 *Test 2*. Performance of various strategies for obstacle case with two exits in the microscopic case

	Go-to-target	CS (3 it)
Evacuation time (time steps)	>3000	2948
Evacuated mass (percentage)	42%	100%

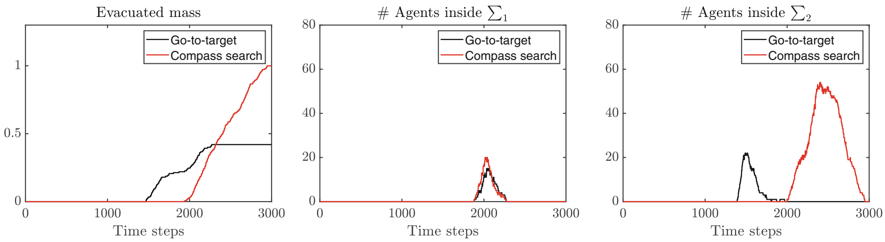


Fig. 15 *Test 2*. Microscopic case: mass maximization in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies

in terms of evacuated mass, where with only three iterations of the optimization method total evacuation is accomplished. In Fig. 15 we compare the cumulative distribution of evacuated mass and the occupancy of the exits visibility areas as a function of time for go-to-target strategy and optimized strategy. We remark that with minimal change of the fixed strategy we reach evacuation of the total mass.

Mesoscopic Case Consider now the case of continuous mass of followers, and the equivalent setting as in the microscopic case. Initial configuration is reported in Fig. 13. We report the evolution of the two scenarios in Fig. 16, where in the upper row we depict three different time frames of the dynamics obtained with go-to-target strategy. Once leaders have moved outside the inner room, at time $t = 1400$, followers mass splits into two parts. However, only leaders moving towards the lower exit x_2^τ are able of steering the followers towards the target, the rest of the followers moving upwards get lost and at final time $t = 3000$ is located close to the left wall. Hence, partial evacuation of followers is achieved, as shown in Table 7 we retrieve 78.8% of total mass evacuated. Only one exit is used, this may cause problems of heavy congestion around the exits. Bottom row of Fig. 16 shows the situation with optimized leaders strategy. Differently from the previous case at time $t = 2380$ the whole mass has been evacuated, part of the followers mass reaches the lower exits and the remaining mass reaches x_1^τ after a while. In Table 7 we reported the performances of the two approaches. In Fig. 17 we compare the evacuated mass and the occupancy of the exits visibility zone as a function of time for go-to-target strategy and optimized strategy after 5 iterations of compass search method.

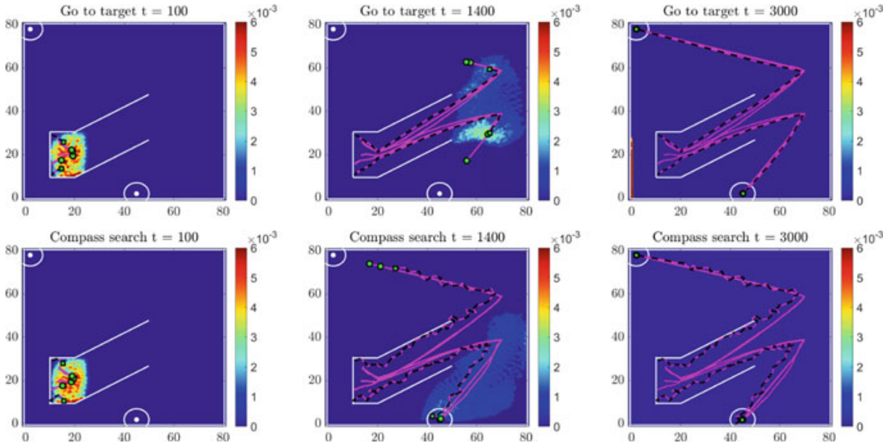


Fig. 16 *Test 2.* Mesoscopic case: mass maximization in presence of obstacles. Upper row: three snapshots taken at time $t = 100$, $t = 1400$, $t = 3000$ with the go-to-target strategy. Lower row: three snapshots taken at time $t = 100$, $t = 1400$, $t = 3000$ with the optimized compass search strategy

Table 7 *Test 2.* Performances of total mass evacuation problems in the mesoscopic case

	Go-to-target	CS (5 it)
Evacuation time (time steps)	>3000	2380
Evacuated mass (percentage)	78,8%	100%

4.1.3 Test 3: Optimal Mass Splitting over Multiple Exits

Problems of heavy congestion and overcrowding around the exits arise naturally in evacuation and, in real-life situations, they can cause injuries due to over-compression and suffocation. Instead of maximizing the total evacuated mass or the minimum time, we ask to distribute the total evacuated mass at final time T between all the exits as reported in (12). The choice of mass redistribution among the different exits can be done according to the specific application and environment. In what follows we consider two different examples, both with two exits, and we will require that mass splits uniformly between the two targets.

Setting 1) Two Exits in a Close Environment

As first example we consider the same setting of Test 2, where complete evacuation was achieved, but all followers were directed toward a single exit. In this case we aim to optimize leaders strategies in order to equidistribute the total mass of follower among the two exits.

Microscopic Case In Fig. 18 we depict the scenario for the fixed strategy and the optimized one. We observe that again with go-to-target strategy the complete

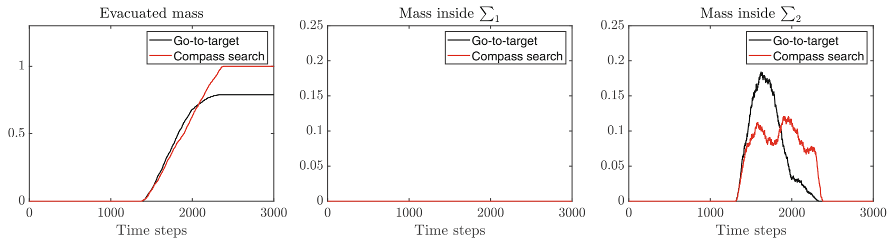


Fig. 17 *Test 2.* Mesoscopic case: mass maximization in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies

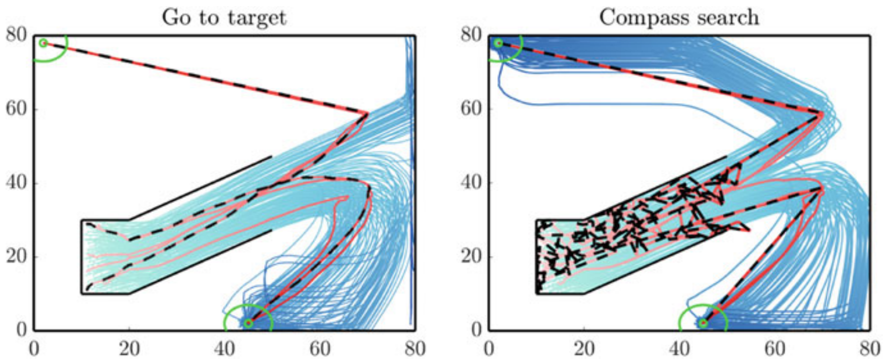


Fig. 18 *Test 3a.* Microscopic case: mass splitting in presence of obstacles. On the left, go-to-target. On the right, optimal compass search

Table 8 *Test 3a.* Performances of mass splitting in the microscopic case

	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 3000	2704
Mass evacuated from E_1	0%	45%
Mass evacuated from x_2^τ	72%	55%
Total mass evacuated	72%	100%

evacuation is not achieved. Moreover, since the vast majority of followers reach the lower exit x_2^τ, heavy congestion is formed in the visibility area Σ_2. On the other hand with an optimized strategy two aware leaders slow down their motion spending more time inside the inner room. In this way, followers are split between the two exits, and the entire mass is evacuated at final time. In Table 8 we report the performances of the two strategies, where for optimized strategy we have 45% of mass in x_1^τ and 55% in x_2^τ. In Fig. 19 we report the evacuated mass and the occupancy of the exits visibility zone as a function of time for go-to-target strategy and optimal compass search strategy. Note that, with the compass search strategy, the whole mass is split between the two exits reducing the overcrowding in the visibility region.

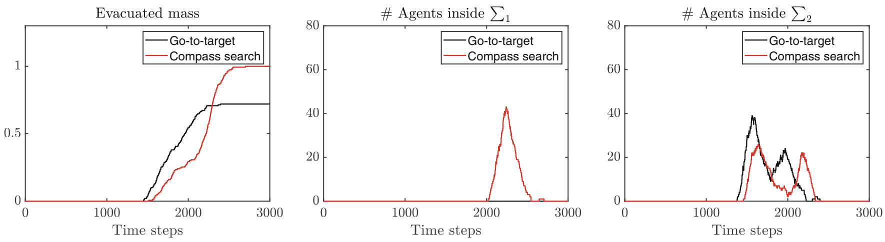


Fig. 19 *Test 3a.* Microscopic case: mass splitting in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies

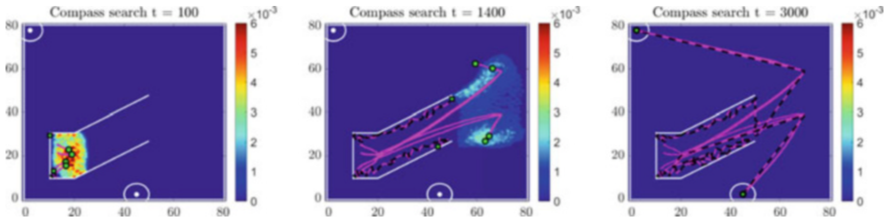


Fig. 20 *Test 3a.* Mesoscopic case: mass splitting in presence of obstacles. Three snapshots taken at time $t = 100$, $t = 1400$, $t = 3000$ with the optimal compass search strategy. For the go-to-target case we refer to the first row of Fig. 16

Table 9 *Test 3a.* Performances of mass splitting in the mesoscopic case

	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 3000	> 3000
Mass evacuated from x_1^τ	0%	49%
Mass evacuated from x_2^τ	78, 8%	50%
Total mass evacuated	78,8%	99%

Mesoscopic Case We report now the case of a continuum density of followers. For the go-to-target strategy, we consider the same dynamics of the previous test, in this case the mass of followers does not split between the two exits, as shown in Fig. 16, and the 78, 8% reaches exit x_2^τ. In Fig. 20, three snapshots were taken at three different times with the compass search strategy. At time $t = 100$, leaders move to evacuate the followers mass out of the inner room. At time $t = 1400$, the followers mass splits in two masses, one moving towards the upper and the other towards the lower exits. At time $t = 3000$, almost all the followers mass is evacuated. The mass is split between the two exits as shown in Table 9. In Fig. 17 we compare the evacuated mass and the occupancy of the exits visibility zone as a function of time for go-to-target strategy and optimal compass search strategy. With the compass search technique the occupation of the visibility areas is reduced since the splitting of the total mass between the two exits is optimized. Hence, the risk of injuries due to overcrowding in real-life situations should be reduced, as depicted in (Fig. 21).

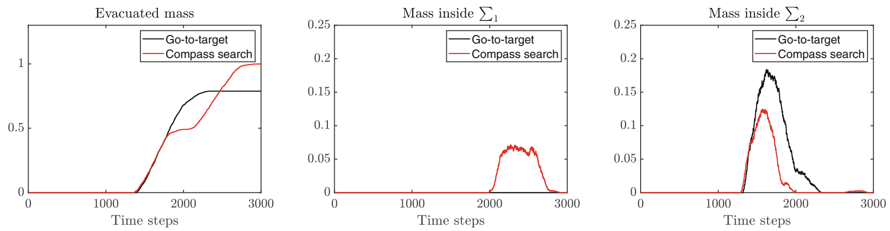


Fig. 21 *Test 3a.* Mesoscopic case: mass splitting in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies

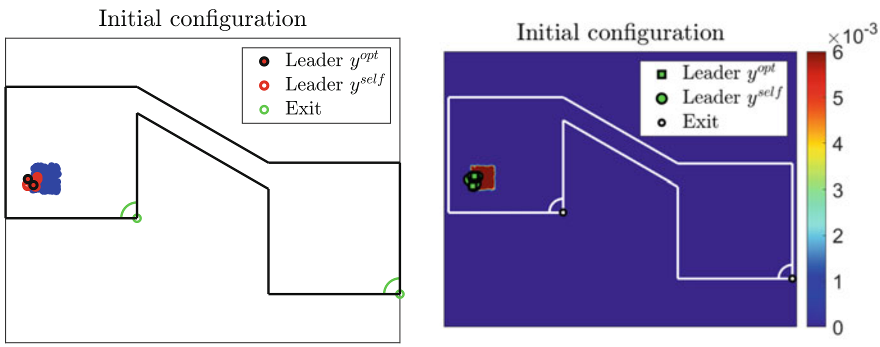


Fig. 22 *Test 3b.* Mass splitting in presence of staircases, initial configuration

Setting b) Two Exits with Staircases

Consider two rooms and two exits, limited by walls, positioned at different floors, and connected by a staircase. Each room has an exit located in the bottom right corner. We assume followers and leaders to be uniformly distributed in a square inside the first room. Similar to the previous case, we assume that the model includes eight unaware and two aware leaders in total $N^L = 10$. The admissible leaders trajectories are defined as in Eq. (24), we choose $\beta = 1$ for every leaders. The target position is $\Xi_k(t) = x_1^\tau \; \forall t$ for the leaders moving towards the exit in the first room. While for the others is $\Xi_k(t) = x_2^\tau$ for $t > t_*$ and for $t < t_*$ we select two intermediate points in such a way that first leaders reach the staircases and then the second room. Indeed, to evacuate, agents must either reach the exit in the first room, called exit x_1^τ, or move towards the staircase, reach the second room and then search for the other exit, called exit x_2^τ. The initial configuration is shown in Fig. 22.

Microscopic Case Consider $N^L = 10$ leaders. Assume that two leaders are aware of their role while the remaining are selfish leaders. Exits for every leader are chosen at time $t = 0$ in such a way that five unaware leaders move toward exit x_1^τ and the remaining toward exit x_2^τ. Among them, one of the two aware leaders moves towards one exit and the other towards the other exit.

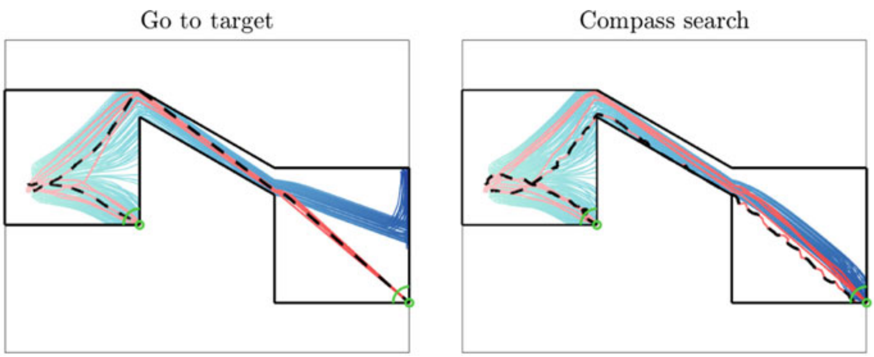


Fig. 23 *Test 3b*. Microscopic case: mass splitting in presence of staircases. On the left, go-to-target. On the right, compass search

Table 10 *Test 3b*. Performances of mass splitting in the microscopic case

	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 3000	2627
Mass evacuated from x_1^τ	57%	62%
Mass evacuated from x_2^τ	0%	38%
Total mass evacuated	57%	100%

In the case of go-to-target strategy, leaders drive some followers to exit x_1^τ and some others to the staircase. The ones that reach the staircase move from the upper to the lower room and then are driven by leaders to exit x_2^τ. As shown in Fig. 23 on the left, some followers are able to reach exit x_1^τ and some others to reach the second room. However, since the vast majority of leaders are unaware and move selfishly towards the exits, followers do not evacuate completely. Hence the only exit useful for evacuation is the one placed in the first room, exit x_1^τ, whose visibility area is overcrowded. On the right of Fig. 23 leaders movement follows an optimized strategy allowing followers to split between the two exits. In this case, complete evacuation is achieved. Table 10 reports the performances of the two strategies. With the go-to-target strategy, all the evacuated followers reach the visibility area Σ_1 and hence are evacuated from exit x_1^τ. With an optimized strategy instead, a larger amount of followers is evacuated and the overcrowding of the visibility areas is reduced.

In Fig. 24 we compare the evacuated mass and the occupancy of the exits visibility zone as a function of time for go-to-target strategy and optimal compass search strategy. Note that, with the compass search strategy, the whole mass is split between the two exits while with the go-to-target strategy the evacuated mass reaches only exit x_1^τ.

Mesoscopic Case Consider the case of a continuous mass of followers. Similar to the microscopic case we observe in Fig. 25 the evolution of the dynamics with fixed strategy and with the optimized one. The upper row shows that with the go-to-target

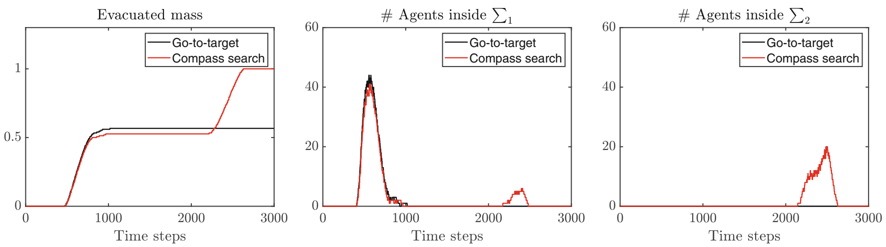


Fig. 24 *Test 3b.* Microscopic case: mass splitting in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and compass search strategies

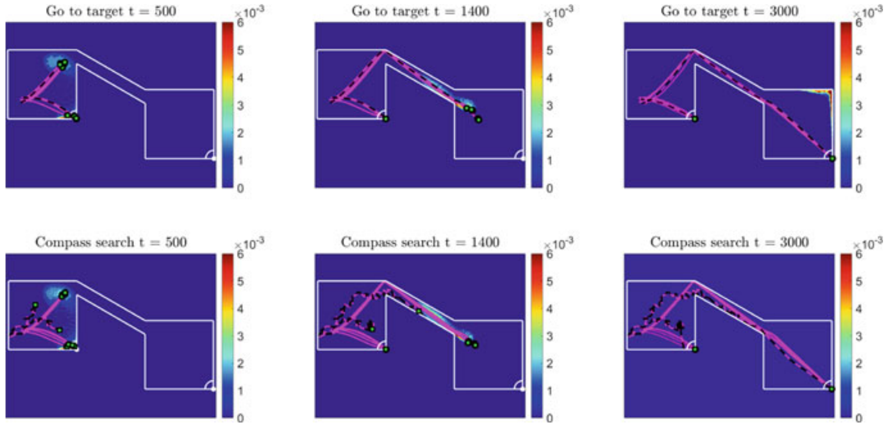


Fig. 25 *Test 3b.* Mesoscopic case: mass splitting in presence of staircases. Upper row: three snapshots taken at time $t = 500$, $t = 1400$, $t = 3000$ with the go-to-target strategy. Lower row: three snapshots taken at time $t = 500$, $t = 1400$, $t = 3000$ with the optimized compass search strategy

strategy the total evacuated mass is not split between the two exits since just the $1, 2\%$ of mass reaches exit x_2^τ. However, as shown in Table 11 a larger percentage of followers reaches exit x_1^τ and the remaining part spreads in the second room without evacuate.

The lower row of Fig. 25 shows the dynamics obtained with the optimized compass search strategy. At the time $t = 500$ a larger follower mass is moving towards the staircase. At time $t = 3000$ almost all the mass is evacuated and split between the two exits. In Table 11 we compare the two strategies showing that with the compass search technique it is possible to improve the mass splitting.

In Fig. 26 we compare the evacuated mass and the occupancy of the exits visibility zone as a function of time for go-to-target strategy and optimal compass search strategy. Note that, with the compass search strategy a larger percentage of mass reaches exit x_2^τ than with the go-to-target strategy.

Table 11 *Test 3b.* Performances of mass splitting in the mesoscopic case

	Go-to-target	CS (50 it)
Evacuation time (time steps)	> 3000	> 3000
Mass evacuated from x_1^τ	46, 6%	51%
Mass evacuated from x_2^τ	1, 2%	48%
Total mass evacuated	47, 8%	99%

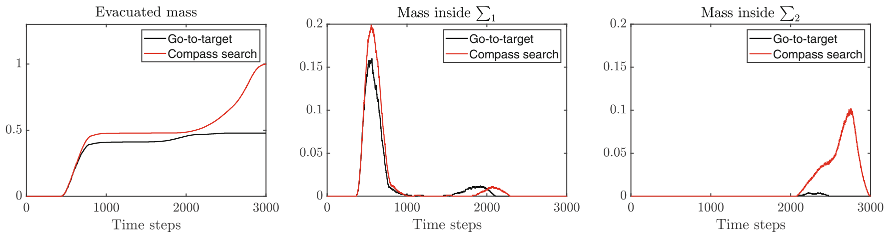


Fig. 26 *Test 3b.* Mesoscopic case: mass splitting in presence of obstacles. Evacuated mass (left), occupancy of the visibility area Σ_1 (center) and Σ_2 (right) as a function of time for go-to-target and optimal compass search strategies

4.2 Discussion and Comparison

In the previous tests we have considered different scenarios to create more complex situations in relation to the functionals chosen, [55]. In general, given a certain setting, it is difficult to choose the optimal number of leaders that guarantee evacuation, and a high number of leaders does not necessarily imply better evacuation efficiency, see, for example, [47]. Another challenging aspect is to give a uniform measure of the performance of the different strategies in such different contexts. A viable option is to quantify the congestion around the exits to exclude dangerous situations. Following the idea in [35] we consider the congestion value

$$cong_{\Sigma_i}(t) = \rho_{\Sigma_i}(t) var_{\Sigma_i}(v(t)),$$

where $\rho_{\Sigma_i}(t)$ is the number of agents (mass) in the microscopic (mesoscopic) case inside Σ_i at time t and

$$var_{\Sigma_i}(v(t)) = \frac{1}{\rho_{\Sigma_i}(t)} \sum_{j \in \Sigma_i} \left(|v_j(t)| - s \right)^2 .$$

We consider also m_{Σ_i} the maximum number of pedestrians over time inside the visibility area Σ_i and l_{Σ_i} the percentage of time in which the visibility area Σ_i is not empty, finally we denote by M_{Σ_i} the percentage of mass inside Σ_i in the mesoscopic case.

Table 12 Comparison of the congestion in the visibility areas for the microscopic case. In red the maximum value of $cong_{\Sigma_i}$ among the visibility areas Σ_i

	$cong_{\Sigma_1}$	$cong_{\Sigma_2}$	$cong_{\Sigma_3}$	m_{Σ_1}	m_{Σ_2}	m_{Σ_3}	l_{Σ_1}	l_{Σ_2}	l_{Σ_3}
Test 1a	0.039	0.011	0.012	40	19	17	0.73	0.51	0.33
Test 1b	0.013	0.009	–	27	16	–	0.36	0.22	–
Test 2	0.009	0.056	–	20	54	–	0.13	0.31	–
Test 3a	0.035	0.027	–	43	26	–	0.19	0.29	–
Test 3b	0.024	0.006	–	41	20	–	0.28	0.16	–

Table 13 Comparison of the congestion in the visibility areas for the mesoscopic case. In red the maximum value of $cong_{\Sigma_i}$ among the visibility areas Σ_i

	$cong_{\Sigma_1}$	$cong_{\Sigma_2}$	$cong_{\Sigma_3}$	M_{Σ_1}	M_{Σ_2}	M_{Σ_3}	l_{Σ_1}	l_{Σ_2}	l_{Σ_3}
Test 1a	0.025	0.005	0.016	0.22	0.6	0.16	0.88	0.79	0.75
Test 1b	0.010	0.005	–	0.1	0.08	–	0.51	0.26	–
Test 2	0	0.009	–	0	0.12	–	0	0.36	–
Test 3a	0.005	0.011	–	0.07	0.12	–	0.3	0.32	–
Test 3b	0.013	0.004	–	0.2	0.1	–	0.41	0.3	–

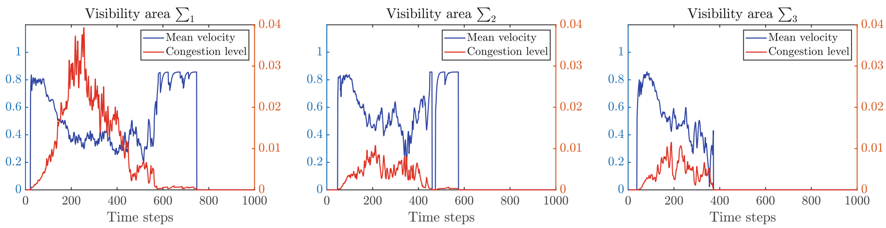


Fig. 27 *Test 1a*. Microscopic case: number of agents and mean velocity of the visibility areas

In this way we can compare the congestion of the various exits for different settings, showing that the more desirable situations are when $cong_{\Sigma_i}$ and m_{Σ_i} (M_{Σ_i}) are small and l_{Σ_i} is high. We reported in Tables 12 and 13, respectively, the values for the microscopic and the mesoscopic setting.

Finally, Figs. 27, 28 show the mean velocity and the congestion level for the case of evacuation with three exits (*Test 1a*) in the microscopic and mesoscopic case, respectively. These plots underline that if the congestion level is higher then the mean velocity is lower.

5 Conclusions

This work has been devoted to the study of optimized strategies for the control of egressing pedestrians in an unknown environment. In particular, we studied situations with complex environments where multiple exits and obstacles are

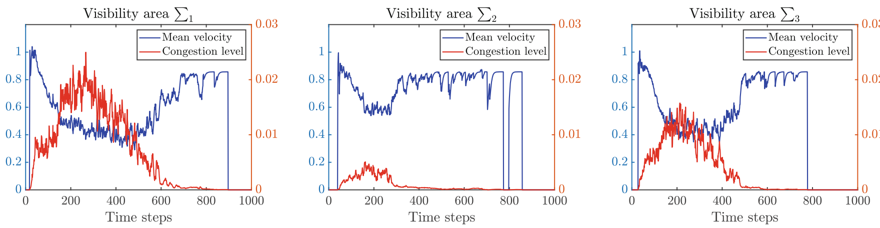


Fig. 28 *Test 1a*. Mesoscopic case: mass of agents and mean velocity of the visibility areas

present. Few informed agents act as controllers over the crowd, without being recognized as such. Indeed it has been shown that minimal intervention can change completely the behavior of a large crowd, and at the same time avoiding adversarial behaviors. On the other hand, we observed that if part of the informed agents moves without coordinated action, this may cause critical situations, such as congestion around the exit. Hence it is important to have a clear understanding of different strategies to enhance the safe evacuation of the crowd. To this end, we explored various optimization tasks such as minimum time evacuation, maximization of mass evacuated, and optimal mass distribution among exits.

We investigated these dynamics at the various scales: from the microscopic scale of agent-based systems to the statistical description of the system given by mesoscopic scale. Numerically we proposed an efficient scheme for the simulation of the mean-field dynamics, whereas we use a metaheuristic approach for the synthesis of optimized leaders strategies. The proposed numerical experiments suggest that the optimization of leaders movements is enough to de-escalate critical situations.

Different questions arise at the level of control through leaders with multiple exits and obstacles. In such a rich environment several research directions can be explored, such as optimal positioning and amount of leaders within the crowd, or different type of cooperative strategies among different groups of leaders to optimally distribute the followers crowd.

References

1. Abdelghany, A., Abdelghany, K., Mahmassani, H., Alhalabi, W.: Modeling framework for optimal evacuation of large-scale crowded pedestrian facilities. European J. Oper. Res. **237**(3), 1105–1118 (2014). https://doi.org/10.1016/j.ejor.2014.02.054
2. Agnelli, J.P., Colasuonno, F., Knopoff, D.: A kinetic theory approach to the dynamic of crowd evacuation from bounded domains. Math. Models Methods Appl. Sci. **25**(1), 109–129 (2015). https://doi.org/10.1142/S0218202515500049
3. Albi, G., Bellomo, N., Fermo, L., Ha, S.Y., Kim, J., Pareschi, L., Poyato, D., Soler, J.: Vehicular traffic, crowds, and swarms: from kinetic theory and multiscale methods to applications and research perspectives. Math. Models Methods Appl. Sci. **29**(10), 1901–2005 (2019). https://doi.org/10.1142/S0218202519500374

4. Albi, G., Bongini, M., Cristiani, E., Kalise, D.: Invisible control of self-organizing agents leaving unknown environments. SIAM J. Appl. Math. **76**(4), 1683–1710 (2016). https://doi.org/10.1137/15M1017016
5. Albi, G., Choi, Y.P., Fornasier, M., Kalise, D.: Mean field control hierarchy. Applied Mathematics & Optimization **76**(1), 93–135 (2017)
6. Albi, G., Cristiani, E., Pareschi, L., Peri, D.: Mathematical models and methods for crowd dynamics control. In: Crowd Dynamics, Volume 2, pp. 159–197. Springer (2020)
7. Albi, G., Pareschi, L.: Binary interaction algorithms for the simulation of flocking and swarming dynamics. Multiscale Modeling & Simulation **11**(1), 1–29 (2013). https://doi.org/10.1137/120868748
8. Albi, G., Pareschi, L.: Modeling of self-organized systems interacting with a few individuals: from microscopic to macroscopic dynamics. Appl. Math. Lett. **26**, 397–401 (2013)
9. Albi, G., Pareschi, L., Zanella, M.: Boltzmann-type control of opinion consensus through leaders. Phil. Trans. R. Soc. A **372**, 20140138/1–18 (2014). https://doi.org/10.1098/rsta.2014.0138
10. Audet, C., Dang, K.C., Orban, D.: Optimization of algorithms with OPAL. Math. Prog. Comp. **6**(3), 233–254 (2014). https://doi.org/10.1007/s12532-014-0067-x
11. Bailo, R., Bongini, M., Carrillo, J.A., Kalise, D.: Optimal consensus control of the Cucker-Smale model. IFAC-PapersOnLine **51**(13), 1–6 (2018)
12. Bailo, R., Carrillo, J.A., Degond, P.: Pedestrian models based on rational behaviour. In: Crowd Dynamics, Volume 1, pp. 259–292. Springer (2018)
13. Blanchet, A., Degond, P.: Kinetic models for topological nearest-neighbor interactions. J. Stat. Phys. **169**(5), 929–950 (2017). https://doi.org/10.1007/s10955-017-1882-z
14. Bongini, M., Fornasier, M., Kalise, D.: (Un)conditional consensus emergence under perturbed and decentralized feedback controls. Discrete Contin. Dyn. Syst. **35**(9), 4071–4094 (2015). https://doi.org/10.3934/dcds.2015.35.4071
15. Bongini, M., Fornasier, M., Rossi, F., Solombrino, F.: Mean-field Pontryagin maximum principle. Journal of Optimization Theory and Applications **175**(1), 1–38 (2017)
16. Borzì, A., Wongkaew, S.: Modeling and control through leadership of a refined flocking system. Math. Models Methods Appl. Sci. **25**(2), 255–282 (2015). https://doi.org/10.1142/S0218202515500098
17. Burger, M., Pinnau, R., Totzeck, C., Tse, O., Roth, A.: Instantaneous control of interacting particle systems in the mean-field limit. Journal of Computational Physics **405**, 109181 (2020)
18. Caponigro, M., Fornasier, M., Piccoli, B., Trélat, E.: Sparse stabilization and optimal control of the Cucker-Smale model. Math. Control Relat. Fields **3**(4), 447–466 (2013). https://doi.org/10.3934/mcrf.2013.3.447
19. Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. In: Mathematical modeling of collective behavior in socio-economic and life sciences, pp. 297–336. Springer (2010)
20. Carrillo, J.A., Martin, S., Wolfram, M.T.: An improved version of the Hughes model for pedestrian flow. Math. Models Methods Appl. Sci. **26**(4), 671–697 (2016). https://doi.org/10.1142/S0218202516500147
21. Cirillo, E.N.M., Muntean, A.: Dynamics of pedestrians in regions with no visibility - A lattice model without exclusion. Physica A **392**, 3578–3588 (2013)
22. Colombo, R.M., Lécureux-Mercier, M.: Nonlocal crowd dynamics models for several populations. Acta Mathematica Scientia **32**(1), 177–196 (2012)
23. Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A.: Effective leadership and decision-making in animal groups on the move. Nature **433**, 513–516 (2005)
24. Cristiani, E., Peri, D.: Handling obstacles in pedestrian simulations: Models and optimization. Appl. Math. Model. **45**, 285–302 (2017)
25. Cristiani, E., Peri, D.: Robust design optimization for egressing pedestrians in unknown environments. Appl. Math. Model. **72**, 553–568 (2019)
26. Cristiani, E., Piccoli, B., Tosin, A.: Multiscale modeling of granular flows with application to crowd dynamics. Multiscale Model. Simul. **9**, 155–182 (2011)

27. Cristiani, E., Piccoli, B., Tosin, A.: Multiscale Modeling of Pedestrian Dynamics. Modeling, Simulation & Applications. Springer (2014)
28. Cristiani, E., Priuli, F.S., Tosin, A.: Modeling rationality to control self-organization of crowds: An environmental approach. SIAM J. Appl. Math. **75**(2), 605–629 (2015). https://doi.org/10.1137/140962413
29. Cucker, F., Smale, S.: Emergent behavior in flocks. IEEE Trans. Autom. Contr. **52**(5), 852–862 (2007)
30. Degond, P., Pulvirenti, M.: Propagation of chaos for topological interactions. Ann. Appl. Probab. **29**(4), 2594–2612 (2019). https://doi.org/10.1214/19-AAP1469
31. Di Francesco, M., Fagioli, S.: Measure solutions for non-local interaction pdes with two species. Nonlinearity **26**(10), 2777 (2013)
32. Di Francesco, M., Markowich, P.A., Pietschmann, J.F., Wolfram, M.T.: On the Hughes' model for pedestrian flow: The one-dimensional case. Journal of Differential Equations **250**(3), 1334–1362 (2011)
33. Duan, H., Sun, C.: Swarm intelligence inspired shills and the evolution of cooperation. Sci. Rep. **4**, 5210 (2014)
34. Düring, B., Markowich, P., Pietschmann, J.F., Wolfram, M.T.: Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. Proc. R. Soc. A **465**, 3687–3708 (2009)
35. Feliciani, C., Nishinari, K.: Measurement of congestion and intrinsic risk in pedestrian crowds. Transportation research part C: emerging technologies **91**, 124–155 (2018)
36. Festa, A., Tosin, A., Wolfram, M.T.: Kinetic description of collision avoidance in pedestrian crowds by sidestepping. Kinet. Relat. Models **11**(3), 491–520 (2018). https://doi.org/10.3934/krm.2018022
37. Fornasier, M., Piccoli, B., Rossi, F.: Mean-field sparse optimal control. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**(2028), 20130400 (2014)
38. Fornasier, M., Solombrino, F.: Mean-field optimal control. ESAIM Control Optim. Calc. Var. **20**(4), 1123–1152 (2014). https://doi.org/10.1051/cocv/2014009
39. Guo, R.Y., Huang, H.J., Wong, S.C.: Route choice in pedestrian evacuation under conditions of good and zero visibility: experimental and simulation results. Transportation Res. B **46**(6), 669–686 (2012). https://doi.org/10.1016/j.trb.2012.01.002
40. Han, J., Li, M., Guo, L.: Soft control on collective behavior of a group of autonomous agents by a shill agent. Jrl. Syst. Sci. & Complexity **19**(1), 54–62 (2006). https://doi.org/10.1007/s11424-006-0054-z
41. Han, J., Wang, L.: Nondestructive intervention to multi-agent systems through an intelligent agent. PLoS ONE **8**(5), e61542 (2013). https://doi.org/10.1371/journal.pone.0061542
42. Haskovec, J.: Flocking dynamics and mean-field limit in the Cucker-Smale-type model with topological interactions. Phys. D **261**, 42–51 (2013). https://doi.org/10.1016/j.physd.2013.06.006
43. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. Nature **407**, 487–490 (2000)
44. Herty, M., Pareschi, L., Steffensen, S.: Mean-field control and Riccati equations. Netw. Heterog. Media **10**(3), 699–715 (2015). https://doi.org/10.3934/nhm.2015.10.699
45. Jin, S., Li, L., Liu, J.G.: Random batch methods (rbm) for interacting particle systems. Journal of Computational Physics **400**, 108877 (2020)
46. Lachapelle, A., Wolfram, M.T.: On a mean field game approach modeling congestion and aversion in pedestrian crowds. Transportation research part B: methodological **45**(10), 1572–1589 (2011)
47. Ma, Y., Yuen, R.K.K., Lee, E.W.M.: Effective leadership for crowd evacuation. Physica A: Statistical Mechanics and its Applications **450**, 333–341 (2016)
48. Motsch, S., Tadmor, E.: A new model for self-organized dynamics and its flocking behavior. J. Stat. Phys. **144**(5), 923–947 (2011). https://doi.org/10.1007/s10955-011-0285-9

49. Pareschi, L., Toscani, G.: Interacting multi-agent systems. Kinetic equations & Monte Carlo methods. Oxford University Press, USA (2013)
50. Parisi, D.R., Dorso, C.O.: Microscopic dynamics of pedestrian evacuation. Physica A **354**, 606–618 (2005). https://doi.org/10.1016/j.physa.2005.02.040. URL http://www.sciencedirect.com/science/article/pii/S0378437105002013
51. Villani, C.: Handbook of Mathematical Fluid Dynamics, vol. 1, chap. A review of mathematical topics in collisional kinetic theory. Elsevier (2002)
52. Wang, J., Zhang, L., Shi, Q., Yang, P., Hu, X.: Modeling and simulating for congestion pedestrian evacuation with panic. Physica A **428**, 396–409 (2015). https://doi.org/10.1016/j.physa.2015.01.057
53. Xie, H., Filippidis, L., Galea, E.R., Blackshields, D., Lawrence, P.J.: Experimental analysis of the effectiveness of emergency signage and its implementation in evacuation simulation. Fire and Materials **36**(5-6), 367–382 (2012)
54. Zhang, Z., Jia, L., Qin, Y.: Optimal number and location planning of evacuation signage in public space. Safety science **91**, 132–147 (2017)
55. Zhou, M., Dong, H., Zhao, Y., Zhang, Y., Ioannou, P.A.: Optimal number and location planning of evacuation leader in subway stations. IFAC-PapersOnLine **51**(9), 410–415 (2018)

The Impact of Physical Distancing on the Evacuation of Crowds

Enrico Ronchi, Daniel Nilsson, Ruggiero Lovreglio, Mikayla Register, and Kyla Marshall

Abstract One of the key implications of COVID-19 is the adoption of physical distancing provisions to minimise the risk of virus transmission. Physical distancing can have significant consequences on crowd movement both in normal conditions and during emergencies. The impact of physical distancing is discussed in this chapter by first presenting an overview of its implications on crowd dynamics and space usage. This is followed by an assessment of expected changes in crowd behaviour, including changes in the fundamental walking speed/density and flow/density relationships. Findings from an experiment investigating the impact of physical distancing on flow rates through doors are presented. In addition, a set of recommendations concerning modifications of the hand calculations currently used for evacuation design (e.g. hydraulic models) are presented alongside a discussion on possible modifications to agent-based crowd models. A verification test to evaluate the results produced by crowd evacuation modelling tools considering physical distancing is also presented. This chapter highlights the importance of considering the increased movement time due to physical distancing in evacuation design and provides insights on how to account for this issue in crowd modelling.

1 Introduction

The COVID-19 pandemic has had a significant impact on any stakeholder dealing with crowds and building usage. The spread of the SARS-CoV-2 virus since December 2019 has pushed many countries all around the world to enforce restrictions

E. Ronchi (✉)
Lund University, Lund, Sweden
e-mail: enrico.ronchi@brand.lth.se

D. Nilsson · M. Register · K. Marshall
University of Canterbury, Christchurch, New Zealand

R. Lovreglio
Massey University, Auckland, New Zealand

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_6

such as cancelling or postponing events involving large crowds and reducing and/or denying access to public buildings [1]. One of the main restrictions adopted is physical distancing (also called social distancing) [2]. Physical distancing has been proven to be an effective way to mitigate the spread of viruses before COVID-19 pandemic [3]. Similar results have shown that SARS-CoV-2 transmission has been reduced by physical distancing when implemented in conjunction with testing and contact tracing of all suspected cases [4].

While the World Health Organization [5] recommended keeping a distance of at least 1m from others, countries around the world have provided different distancing requirements during the pandemic, for instance, 1m (China, Denmark, France, Singapore), 1.4 m (South Korea), 1.5 m (Australia, Germany, Italy, Spain), 1.8m (USA) and 2 m (Canada, UK, NZ) [6]. Physical distancing is raising many challenges for the use of space as it significantly decreases the density of people allowed in the built environment [7, 8]. Compliance with those measures is currently under scrutiny [9], as the type of walking behaviour may impact resulting density and inter-person distance [10]. Physical distancing is creating challenges for crowd management both in normal conditions and during emergencies. In fact, in pandemic times, the safety of built environments needs to account for multiple concurrent threats such as virus transmission and fires or terrorist attacks [11].

To date, several epidemiological modelling studies have been conducted to investigate the impact that different provisions may have on the pandemic at different scales [12]. Nevertheless, only a few studies have investigated the use of crowd modelling for pandemic scenarios. Ronchi and Lovreglio [1] have proposed a modelling solution to retrofit crowd models to assess building occupant exposure to a virus in confined spaces. Ronchi et al. [11] have provided insights on how to use crowd evacuation models in times of pandemic. A risk analysis methodology for the use of crowd modelling tools during the COVID-19 pandemic and its implementation have also been proposed [8]. D'Orazio et al. [13] proposed a probabilistic simulation model based on consolidated proximity and exposure-time-based rules for virus transmission which is applied for university buildings. Crowd modelling was also adopted to estimate the spread of the virus in touristic spaces [14]. Garcia et al. [15] proposed a method for calculating the maximum capacity of public spaces constrained to physical distancing, while Xu and Chraibi [16] used crowd modelling to test the effectiveness of different measures to reduce the contact of customers in supermarkets. Another crowd modelling approach available in the literature includes the implementation of different social and physical pedestrian sizes to investigate their impact on pedestrian behaviour [17].

The existing literature shows how crowd modelling can provide insightful outputs to assess the space usage in normal conditions, as well as assessing the time required to evacuate a building in case of physical distancing. To date, many of the most known and used crowd evacuation models [18] have been modified to consider physical distancing requirements. However, these models generally rely on the use of crowd movement relationships which were generated with data collected without any ongoing pandemic [19, 20]. Similarly, algorithms for route choice are based on pre-pandemic conditions [21–23]. The use of crowd models without a careful re-evaluation of their inputs and assumptions due to physical distancing provisions may lead to misleading results [11]. As such, there is the need for a comprehensive evaluation of *the impact of physical distancing on the evacuation of crowds* as well as a need to provide an answer to the following question: What is the impact of physical distancing on crowd evacuation movement?

To date, there is no research investigating to which extent people comply with physical distancing provisions during an emergency. Nevertheless, it is deemed important to assess what possible implications physical distancing may have, especially given the fact that buildings are currently being designed with the help of crowd models that are calibrated with data obtained in pre-pandemic conditions.

This chapter provides an overview of the implications of physical distancing on crowd evacuation, in which the crowd movement is mainly uni-directional. It starts with a general introduction concerning the issues related to physical distancing provisions, including a brief description of existing efforts aimed at modelling crowd movement in pandemic scenarios as well as measuring pedestrian movement behaviour experimentally. This is followed by the description of a set of issues related to crowd dynamics and physical distancing, such as the changes in local densities and occupant load, changes in crowd movement, route and exit choice and group behaviour. Furthermore, an experiment performed to study the impact of physical distancing on crowd evacuation movement is presented. More specifically, the experiment explored the fundamental walking speed/density and flow/density relationships for different physical distancing recommendations (0 m, 1 m and 2 m). Based on the experimental results presented in the original publication, it is highlighted how crowd evacuation movement may change due to physical distancing and, in turn, how this information can be implemented in evacuation models. This includes changes in fundamental walking speed/density and flow/density relationships [24]. This work also proposes a verification test for evacuation modelling tools used for representing crowd movement under physical distancing restrictions. The chapter is concluded with a general discussion concerning existing and future research on crowd evacuation movement and modelling considering physical distancing.

2 Crowd Dynamics and Physical Distancing

Different assumptions can be used for the calculation of physical distance. They concern the representation of the body size and the reference point used for its estimation. It should also be noted that physical distance calculations are generally performed in two dimensions (i.e. the projection of people on a 2D plane). People may be represented considering a top view of their body and could be approximated with different body shapes. Common shapes include elliptical or circular representations [25, 26]. Simpler approximations of body size would assume common dimensions for all individuals, for instance, circles with a radius in the order of magnitude between 0.25 cm and 0.45 m [27, 28].

Different reference points can be used for physical distance estimations. This can include the distance between different parts of the body (e.g. noses, arms, feet, etc.) and an ideal reference point, for instance, the centre point of two people or the closest points between people. The former is often defined as contact distance, while the latter is called inter-person distance [29]. Under certain assumptions, the orientation of the bodies can impact the calculated distances.

Based on the physical distancing under consideration, several changes may occur concerning crowd evacuation dynamics. Those changes relate primarily to:

1. Local and global densities (i.e. the occupant load expected in a given area)
2. Crowd movement, intended as the fundamental speed/density and flow/density relationships
3. Route and exit choice
4. Group behaviour

In addition, the implementation of a target physical distance for a crowd in motion during an evacuation may require higher distance provisions. This has been investigated experimentally, and it is discussed in Sect. 3.

2.1 Changes in Local Densities and Occupant Load

The need to ensure physical distancing has several consequences on space usage, among which the first key aspect strictly linked to evacuation is the expected occupant load (or global density, considering the static crowd) in a given space (a whole building, area or portion of a building/area) and the maximum local densities during movement. In fact, physical distancing can be considered assuming a static crowd (e.g. sitting, standing) or a moving crowd (e.g. walking, running, queuing with stop-and-go behaviour). In the case of evacuation design, we generally refer to a moving crowd. This implies that the crowd may need more space to keep a given target physical distance. Nevertheless, occupant loads are generally calculated assuming an initial static position of people. Therefore, depending on the assumptions in use, the resulting maximum allowed number of people in a given confined or open space area could greatly impact crowd evacuation.

Another key aspect concerning occupant load and local density estimation is the expected initial location of people. Occupant load can be calculated considering the free area around a person. Individuals can then be assumed to be uniformly distributed (with different patterns, i.e. along a virtual chessboard or in an alternated chevron pattern, as shown in Fig. 1), considering the placing from the corners or not, the impact of physical obstructions avoiding virus transmission and placed individually or in clusters/groups. The resulting distribution of people could greatly affect density estimations, thus modifying the occupant loads in a given space and the maximum achievable local densities. Based on Fig. 1, Eqs. 1 and 2 represent examples of methods to calculate density based on a uniform squared packing or

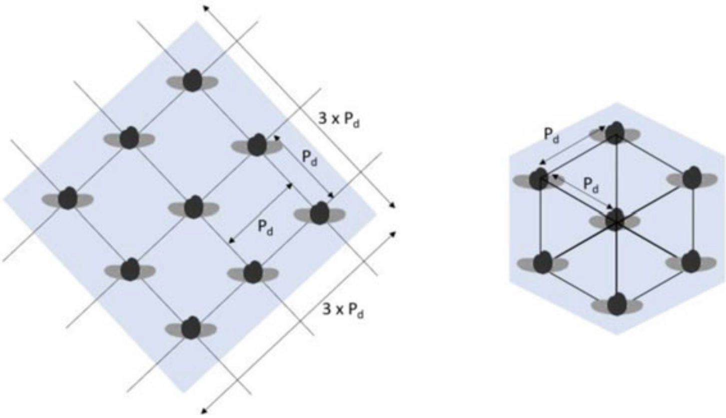


Fig. 1 Possible assumptions for people distributions in occupant load and local density calculations. The figure on the left indicates a hypothetical uniform squared packing, and the figure on the right indicates a hypothetical circular packing approximated with a hexagon

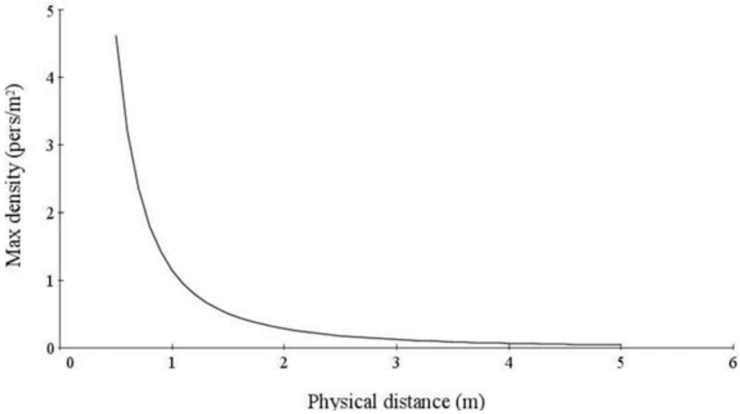


Fig. 2 Relationship between maximum density d_{max} and physical distance according to the circular packing assumption with the approximation presented in Eq. 2

circular packing (here with an approximation to a hexagonal shape), respectively. In addition, the occupant load calculations could be made considering different assumptions regarding the area required by each individual (e.g. circular or squared, as shown in Fig. 2) and the size of the pre-set groups which are considered to be together during their movement being those not required to keep physical distance. Occupant load calculation can then be implemented either considering the area per person (generally in m^2 per person) or the person per area (person per m^2):

$$d_{\max} = \frac{9}{(3P_d)^2} \quad (1)$$

$$d_{\max} = \frac{2}{\sqrt{3}P_d{}^2} \quad (2)$$

Where:

$d_{\max}$ is the maximum density
P_d is the physical distance

In the present work, the use of circular packing has been adopted. The resulting relationship between maximum density $d_{\max}$ (pers/m^2) and physical distance P_d (m) according to Eq. 2 is shown in Fig. 2.

2.2 Changes in Crowd Movement

Physical distancing can have a direct impact on crowd evacuation movement. In fact, each individual may change their speed and/or direction of movement in an attempt to keep a given distance from other people. The movement of an evacuating crowd is generally governed by a set of self-organising rules, which include (but are not limited to) lane formation, which is the tendency to follow people ahead and form lanes when moving in a crowd [30], and clogging effects at bottlenecks [31]. The current understanding is based on experiments or real-world data collected in pre-pandemic conditions, and given the scarcity of data currently available, there is limited knowledge about the self-organisation rules during a pandemic. A clear example of this issue is lane formation, as this rule may change due to the need to maximise physical distance while moving [32].

During an evacuation, crowd movement is expected to mostly happen considering uni-directional flows, so this is the main focus of this work. Nevertheless, there may be conditions in which bi-directional flows may occur (e.g. during rescue service interventions) [33].

An important implication of physical distancing is that the fundamental speed/density and flow/density relationships [24] may need to be modified. Several aspects would be affected by physical distancing, namely, (1) the maximum achievable local density (see Sect. 2.1) and (2) the assumed density range in which walking speed is affected by others. The maximum achievable local density will be directly affected by the maximum physical distance provision (and the compliance with it by the evacuating crowd), and it can be calculated based on the assumptions adopted concerning space usage. The transition between unimpeded walking speed and impeded walking speed would instead depend on the local conditions. It is indeed possible that people may start decreasing their speed at lower-density values

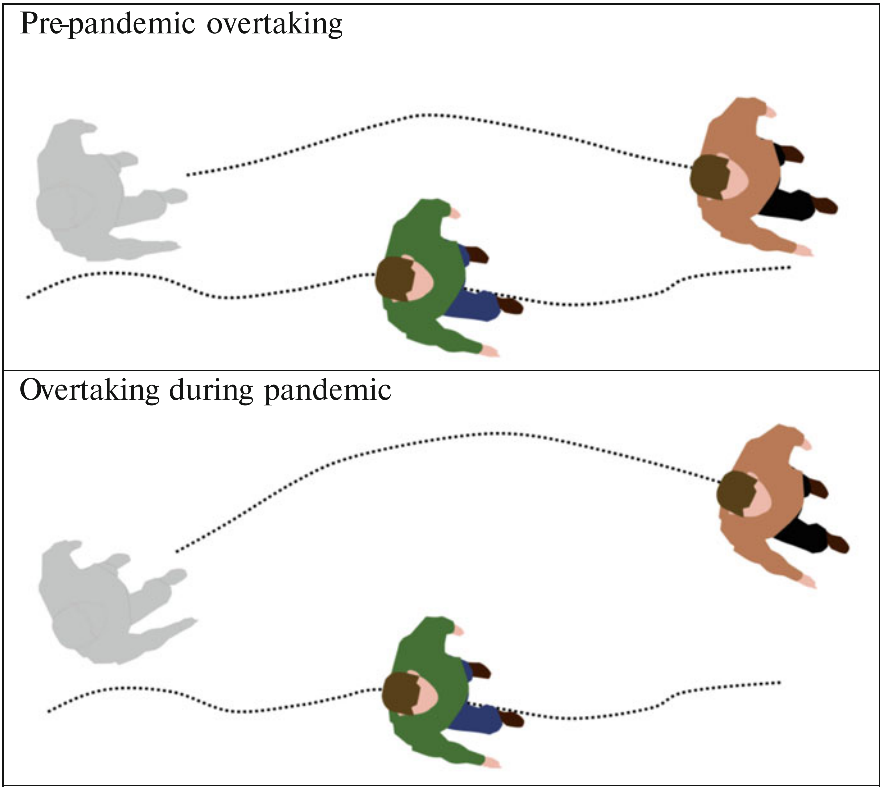


Fig. 3 Hypothetical overtaking movement path in uni-directional flows in pre-pandemic (above) and pandemic (below) conditions

than in pre-pandemic conditions, in order to maintain a larger contact distance from people ahead of them. Possible modifications in the fundamental relationships are discussed in Sect. 4.

At a local crowd movement level, the process of collision avoidance [34] may be impacted by physical distancing. In fact, people may be more prone to avoid face-to-face contact with people (the expected changes would be more visible in bi-directional flows, but this may also impact uni-directional flows), thus altering their movement direction to achieve this goal (see Fig. 3). In general, people may attempt navigating around other people keeping larger contact distances than in a pre-pandemic scenario. A similar issue could be expected during the process of queueing.

2.3 Route and Exit Choice

The understanding of the route and exit choice adopted by crowds during evacuation is currently based on pre-pandemic conditions. This can generally include different aspects such as distance to the exit, expected queuing, route familiarity, route availability, social influence, signage, etc. In pre-pandemic conditions, the analysis of the overall navigation in space does not generally consider the proximity to other people as a deterrent to select a given route/exit. It is therefore important to consider the need to keep physical distancing during the evaluation of the routes possibly adopted by an evacuating crowd.

A complete understanding of route and exit choice during a pandemic would need to take into account people movement and the locations of each individual/groups over time. This is particularly important when attempting an evaluation of queuing. In fact, pedestrians may be more likely to change their initial chosen direction of movement in order to avoid approaching congested areas.

During a pandemic, it is currently not known to which extent the probability of choosing a given route/exit is affected by the presence of other people. This increased uncertainty in route choice decisions may lead to the need to investigate the impact of the variability in route choice decisions to perform a comprehensive assessment of credible crowd evacuation scenarios.

2.4 Group Behaviour

A key aspect to consider when estimating the impact of physical distancing on crowd dynamics is group behaviour [35]. In fact, established groups (e.g. groups of family members, people sharing the same space, etc.) are not required to keep physical distance, thus having a direct impact on all aspects related to crowd dynamics during a pandemic. The assumptions in use concerning the size and characteristics of the groups can therefore have a significant influence on crowd evacuation. Different occupancy types may have different typical groups (e.g. larger groups may be expected in a shopping mall compared to smaller groups or individuals moving in a transient space like a train station). Different variables can be affected by the nature of groups and their behaviour. This includes the maximum achievable local densities, occupant load, space usage/navigation and collision avoidance with other groups/individuals.

3 An Experiment on Physical Distancing

One of the best ways to determine the influence of physical distancing recommendations on crowd flow characteristics is to perform realistic pedestrian movement experiments. However, these types of experiments are arguably unethical to perform

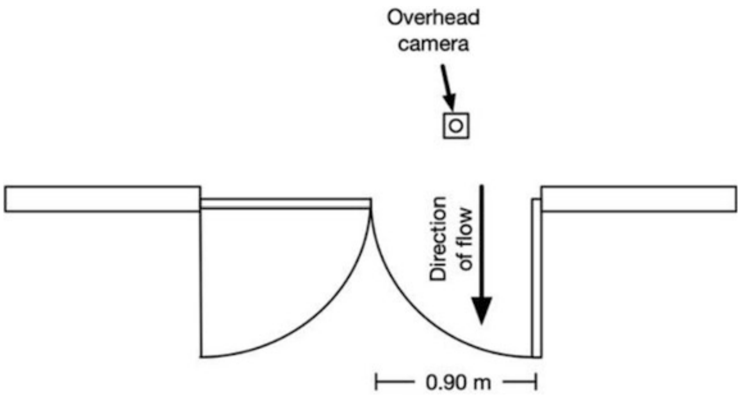


Fig. 4 Schematic representation of the geometry used in the experiment

during a pandemic, as they may lead to the spread of the disease. At the same time, performing experiments in countries that have not been affected by the pandemic may lead to unrealistic behaviour, as participants will have no experience of physical distance keeping.

Given the difficulties highlighted above, New Zealand was in a unique position in 2020. Although it experienced a very limited spread of COVID-19 in 2020 (e.g. 2200 cases over a population lower than 5 million people[1]), the country was in lockdown for a significant time early in the year. People living in New Zealand, therefore, had the experience of physical distancing recommendations, but by September 2020, restriction in most of the country had been relaxed[2]. For this very reason, experiments on crowd movement in simulated pandemic situations were performed at the University of Canterbury on Thursday 17 September 2020 [36].

The experiments [36] were performed in the Engineering Core building at the University of Canterbury (Christchurch, New Zealand). A classroom configuration was chosen, namely, an open plan classroom with an exit leading to a corridor. The area in front of the door was filmed by a total of four cameras, although only the video recording from an overhead camera was eventually used for the analysis (see Fig. 4).

A total of 47 participants took part in the study. Participants were mainly university students and hence represented a relatively young population. The age ranged from 18 to 36 years, with an average age of 21 years. Twenty-seven participants were men, 19 were women and 1 identified as non-binary. All participants were reimbursed with a coffee voucher worth 5 NZD.

[1] https://coronavirus.jhu.edu/region/new-zealand

[2] https://covid19.govt.nz/alert-system/history-of-the-covid-19-alert-system/

A total of four scenarios were evaluated, although only two of the scenarios are deemed relevant for the current publication, namely:

- Scenario 1 m – Participants were told to keep a distance of **1 m** while moving out through the exit wearing a face mask.
- Scenario 2 m – Participants were told to keep a distance of **2 m** while moving out through the exit wearing a face mask.

In both Scenario 1 m and 2 m, the door opened outwards and was held open (see Fig. 4). All scenarios were repeated five times but in a randomised order. The total time for all the participants to walk out from the room through the exit was measured for each iteration of Scenario 1 m and 2 m. In addition, a software called Kinovea[3] was used to measure the inter-person distance between each participant standing in the exit doorway to the participant immediately behind them (l). In addition, the time from the participant standing in the exit doorway until the participant immediately behind had moved to the exit doorway (Δt) was measured. The speed of the person immediately behind (s) was then being calculated according to Eq. 3:

$$v = \frac{l}{\Delta t} \tag{3}$$

This calculation of speed was done for all participants walking through the exit, except for the first participant exiting the room. In addition, the inter-person distance was used to calculate the corresponding occupant density according to the previously mentioned circular packing model (see Fig. 1).

As mentioned previously, each scenario was repeated five times (called iterations 1 to 5). The total time to empty the room varied between 60 and 77 s, with an average time of 68 s, for Scenario 1 m. For Scenario 2 m, the time varied between 87 and 99 s, with an average time of 95 s. The average inter-person distance for each of the iterations of the experiment for the two scenarios can be seen in Table 1.

As shown in Table 1, people kept an average distance of 1.2 m for Scenario 1 m and an average distance of 1.8 for Scenario 2 m. Hence, participants seem to have left a bit more space to the person in front in the 1 m case, as opposed to the 2 m case. The histograms of inter-person distances for scenarios 1m and 2m can be found in Figs. 5 and 6 together with normal distributions obtained with the method of moments considering the average of all the data points (μ = 1.246 and 1.793, respectively) along with the calculated standard deviations (SD = 0.332 and 0.319, respectively).

Figure 7 shows the correlation between speed and density measured in the experiment together with equivalent data curves from Predtechenskii and Milinskii [19] for horizontal movement and movement through openings (exits). For the data curves, a body area of 0.1 m^2 was assumed when converting from dimensionless density. The value of 0.1 m^2 is based on the dimension of people in summer clothes,

[3] https://www.kinovea.org/

Table 1 Average inter-person distance for each iteration for Scenario 1 m and 2 m

Iteration	Inter-person distance (m)	
	Scenario 1 m	Scenario 2 m
1	0.99	1.83
2	1.07	1.77
3	1.29	1.84
4	1.40	1.76
5	1.46	1.92
Average[a]	1.24	1.82

[a]The average is here intended as the value obtained considering aggregated data points for each iteration

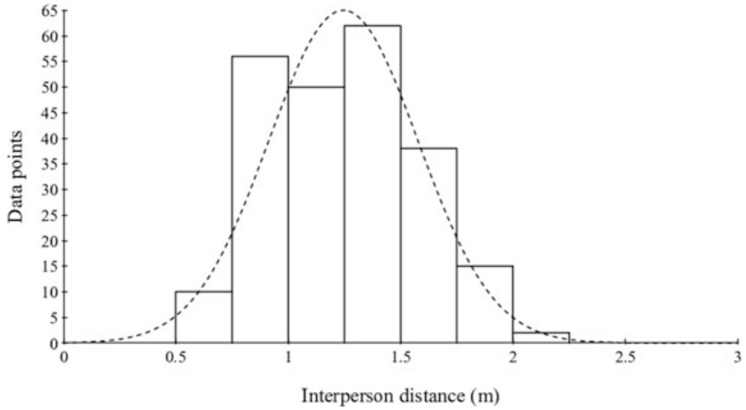


Fig. 5 Distribution of inter-person distance in Scenario 1 m

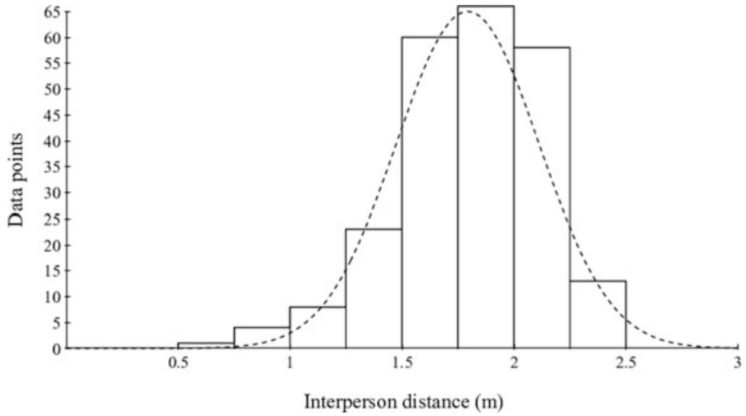


Fig. 6 Distribution of inter-person distance in Scenario 2 m

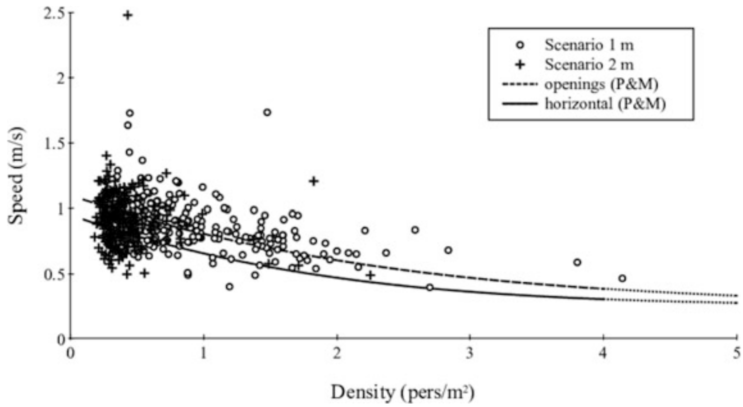


Fig. 7 Participants' data points on speed vs density from the experiments by Marshall and Register [36] and the data correlation curves for horizontal movement and movement through openings from Predtechenskii and Milinskii [19] (P&M in the figure)

according to Predtechenskii and Milinskii [19]. Figure 7 shows that the general trend of the experimental data seems to be in line with the curves proposed by Predtechenskii and Milinskii [19]. This is also confirmed by a residual analysis carried out with the data for both scenarios. The results for Scenario 1 m indicate that the residual average is 0.03 m/s, while for Scenario 2 m, the residual average is 0.04 m/s. This suggests that a cropped version of the design curve proposed in the hydraulic model adopted for engineering design [20] (i.e. stopping at a given density threshold given physical distancing) could be used rather than modifying the shape of the correlation between speed and density.

4 Updated Relationships Between Speed/Flow and Density

This section presents possible modifications to be performed on the simple macroscopic speed/density and flow/density relationships in light of physical distancing. The hand calculation method under consideration is commonly used in evacuation design, namely, the hydraulic model from the *Society of Fire Protection Engineers (SFPE) Handbook* [20]. This calculation method is here updated considering the case of movement along with horizontal egress components (i.e. corridors, doors, etc.).

4.1 The SFPE Hydraulic Model Considering Physical Distancing

The original hydraulic model is a simplified engineering calculation method for evacuation design. It includes a theoretical maximum density (corresponding to an impeded speed of 0 m/s) equal to 3.8 people/m^2. Considering physical distancing, a new maximum density d_{max} would have to be calculated depending on the physical distancing P_d and the assumptions adopted for this calculation. This will result in a value equal to or lower than the original threshold value (see Eq. 4):

$$d_{\max} = f\,(P_d) \leq 3.8\ \frac{\text{people}}{m^2} \tag{4}$$

The SFPE model also includes a so-called minimum density d_m which corresponds to the start of a decrease in the unimpeded speed. This corresponds in the original model to a value of 0.54 people/m^2. The updated density to start the impeded speed can be assumed to be a number between 0 and 0.54 people/m^2 (see Eq. 5):

$$0 \geq d_m \geq 0.54\ \frac{\text{people}}{m^2} \tag{5}$$

Theoretically, the speed reduction in relation to density may then be different than the one considered in the original hydraulic model (see dashed line in Fig. 8), depending on the assumption in use. The first assumption would be to keep the speed unimpeded until the same value of density adopted in the SFPE hydraulic model (see the solid black line in Fig. 8 considering the example of a corridor) and then decrease the speed linearly until $d_{\max}$ (this is equal to 1.15 pers/m^2 in this example; this is the value corresponding to $P_d = 1$ m assuming the theoretical circular packing). The second assumption is to consider the walking speed starting a decrease linearly from the value of unimpeded speed corresponding to density equal to 0 people/m^2 (the unimpeded speed is assumed 1.19 m/s in the original SFPE model; see dotted line in Fig. 8) until $d_{\max}$. The last assumption would be to identify the new value of $d_{\max}$ which should be considered as the maximum threshold, considering the same curve as the original model cropped at $d_{\max}$. Given the experimental data presented in the previous section, the latter approach is recommended. It should be noted that the original SFPE hydraulic model does not differentiate the case of doors/openings from the case of horizontal movement in corridors.

The range of densities in the fundamental relationships between speed and densities in which speed is impeded can therefore be obtained according to Eq. 6:

$$d_m \leq d \leq d_{\max} \tag{6}$$

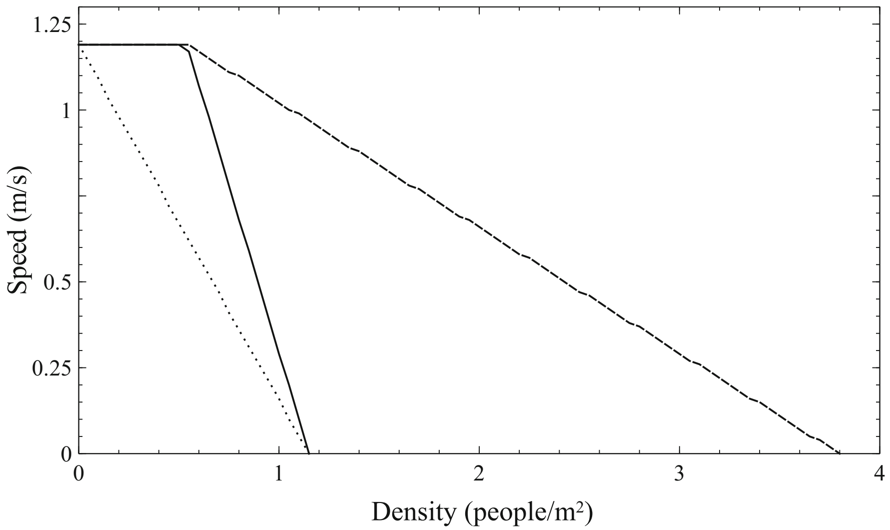


Fig. 8 Examples of changes in speed/density relationship in a corridor depending on the assumptions on density levels corresponding to a start of the decrease in speed. The dashed line represents the original design curve of the hydraulic model, and the solid and dotted lines represent two alternative approaches to modify the speed/density relationship based on different assumptions for d_m and $d_{\max}$ [20]

Considering that the curve will follow the same trend as the original hydraulic model and be applicable only until a certain maximum density $d_{\max}$ based on physical distancing, Equation 4 can be used to calculate the speed:

$$v = k - akd \tag{7}$$

Where:

$a = 0.266$ for metric applications.
$k =$ constant which changes depending on the type of egress component (see [20] for more information).
d is in the interval $[0, d_{\max}]$, where $d_{\max} = \frac{2}{\sqrt{3}{P_d}^2}$ assuming circular packing.

A similar approach can be used to update the corresponding fundamental flow/density relationships, i.e. flowrates can be updated in accordance with physical distancing. The specific flow can be simply obtained by multiplying the speed obtained in Eq. 7 with the corresponding density within the same density interval $[0, d_{\max}]$. The specific flow is therefore calculated as in Eq. 8:

$$F_s = vd \tag{8}$$

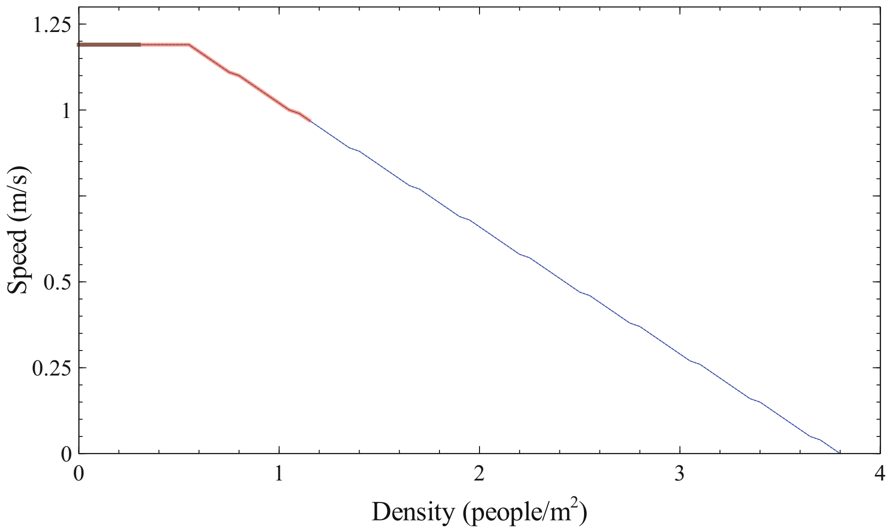


Fig. 9 Examples of cropped speed/density relationships based on physical distances equal to 1m (red line) and 2 m (black line). The dashed blue line represents the original model

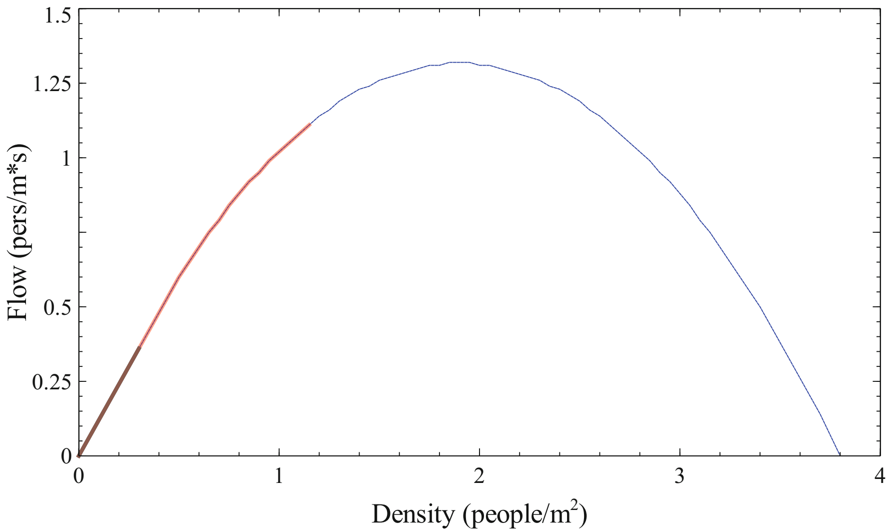


Fig. 10 Examples of cropped flow/density relationships based on physical distances equal to 1m (red line) and 2 m (black line). The dashed blue line represents the original model

An example of resulting cropped speed/density and flow/density relationships corresponding to the physical distances of 1m and 2m (this is calculated based on the theoretical assumption of circular packing) is presented in Figs. 9 and 10.

Once more experimental data become available, it will be possible to confirm the extent to which the shape of the speed/density and flow/density relationships accounting for physical distancing correspond to the pre-pandemic shape. This will allow confirming that the limit for the so-called capacity drop [37] may not be reached due to the lower maximum achievable local density. In addition, data from evacuation drills or actual emergencies could be used to check to which extent people comply with physical distancing provisions, thus being able to identify if the pre-pandemic capacity drop is reached.

5 Crowd Evacuation Modelling

Evacuation modelling can be used to investigate the safety conditions of a confined or open space. In fact, they are useful tools to avoid congestions during the planning of space usage for pedestrians during an evacuation or within a performance-based design approach. The former analysis is used to optimise flows and reduce high-density conditions; the latter is used to address different types of concurrent threats along with the pandemic, such as terrorist attacks [38], toxic releases [39] or fires [40].

A pandemic scenario can impact the whole range of behaviours that can be represented within a crowd evacuation model, including reduced achievable density ranges, modified space usage, collision avoidance, route choice and flow rates. Of particular interest is the representation of the fundamental speed/density and flow/density relationships.

The crowd model developer may need indeed to modify the fundamental modelling methods which govern crowd movement in case of a pandemic. This is linked to the type of crowd evacuation model in use, e.g. coarse network, fine network or continuous model [41]. For instance, in a coarse network model, the required changes include the need to cap the capacity of a node (pers/node) and the need to change the capacity of the links (pers/s). In a fine network model adopting a grid, the cell size would have to be increased to reduce the maximum achievable density. The flow through exits would also need to be modified (i.e. reduced) to account for the impact of physical distancing. In a continuous model, the fundamental speed/density and flow/density relationships would have to be modified, and a minimum distance to others should be defined.

In general terms, the modifications required by a crowd evacuation model can be performed explicitly or implicitly. An explicit representation of the impact of the pandemic at a macroscopic level would lead to adopting dedicated speed/density or flow/density relationships based on physical distance and modified group interactions. These can be imposed within the model or represented through modifications of the underlying algorithms in use for the representation of movement and collision avoidance between pedestrians (in both cases of uni- and bi-directional flows). Commonly adopted microscopic methods for the representation of crowd movement include the social force model [42] or the steering model [43]. The changes needed would therefore result in a modification of the forces and/or rules adopted for modelling the interactions between agents in order to represent the changes due to physical distancing.

Crowd evacuation models would likely have issues (especially network models) in representing the case of groups that are not required to keep physical distancing, as they may require ad hoc modelling implementations which may not be possible to implement just for specific groups.

Existing models are calibrated and validated with collision avoidance rules based on pre-pandemic scenarios. Collision avoidance may depend on several factors [34], such as local density [44], assumed representation of the body size and personal space [45, 46]. During a pandemic, pedestrians may modify their movement to a lower or greater extent in relation to the physical distancing provisions and compliance with them. For this reason, local interactions may greatly be affected during a pandemic in an attempt to increase the distance towards others.

An implicit representation of the impact of the pandemic would require the user to adapt existing or dedicated inputs to consider the impact of the pandemic. This could include, for instance, enforcing a given physical distance between agents during movement and modelling the likelihood of the agents to comply with it. This type of solution may have an impact on several variables related to crowd movement (e.g. walking speed, acceleration, route choice, etc.) and during queuing. In fact, the agents may need to wait until a space gets free from the presence of other people (or groups of people) before proceeding in a given direction or decide to re-route in another direction with lower congestion.

Regardless of the method in use, the model developers/users should make sure that the simulated behaviour corresponds to the intended one. For this reason, it is advisable to run a dedicated set of tests to evaluate the updated fundamental relationships in use by the crowd evacuation model. Different tests and procedures are available in the literature to evaluate the predictive capabilities of crowd evacuation models. In particular, a wide range of verification tests are available to ensure that the conceptual models are correctly implemented and are in line with the intended use of the model [47]. Those procedures include tests from the RIMEA Guidelines for Microscopic Evacuation Analysis [48], the documentation provided by the International Maritime Organization (IMO) [49] and the National Institute of Standards and Technology (NIST) [50]. Most recently, the International Standards Organization (ISO) has released a document providing a comprehensive list of tests that are based on the above-mentioned existing documentation [51].

An example of the use of verification tests to consider the impact of physical distancing is presented here. This concerns the fundamental relationship between walking speed and density for uni-directional flows, and it is based on Test 13 of the ISO document 20414, which is dedicated to this scope (this is a modified version of Test 4 available in the RIMEA Guidelines). The scope of this test is to verify the ability of crowd models to represent the expected uni-directional movement, considering the impact of physical distance. Here there is a description of the test, including the test name, objective, geometry, scenarios expected result, test method and user's actions.

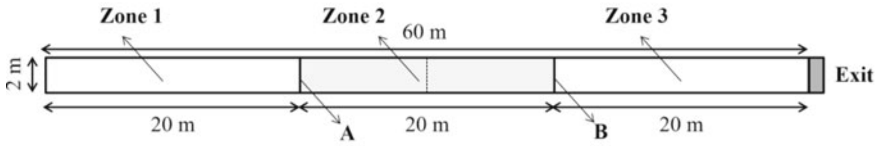


Fig. 11 Schematic geometric layout of the test (top view)

Test name: Relationship between walking speed, uni-directional flow and density considering physical distance.

Objective: Assess qualitative consistency between the relationships of walking speed, uni-directional flow and density assignment and model representation in case of physical distance provisions.

Geometry: A corridor is represented in accordance with Fig. 11, and it is divided into three zones, namely, zone 1 (white), zone 2 (light grey) and zone 3 (white).

Scenarios: Fill in the entire corridor (zones 1, 2 and 3 in Figure 11) with the maximum allowed number of people in accordance with your assumed starting physical distance (people can be placed at random in the space). They have a pre-evacuation time equal to 0 seconds, and a walking speed of 1 m/s is assigned to the entire crowd. Step 1: The occupants move to the right towards the exit of the corridor. Place the last occupant in zone 2 near line A, measure the time that it takes from line A to line B and estimate the associated walking speed. Measure the average occupant flows in line B (with a time interval decided by the tester) starting from the beginning of the simulation until the last occupant in zone 2 arrives at line B. People densities in zone 2 are recorded when the last occupant in zone 2 reaches the centre of zone 2. Step 2: Step 1 is repeated with a number of occupants equal to the double of the original number (i.e. to verify if the model allows an initial density higher than the physical distance provision in use and how people adjust their position to maintain the physical distance), three-quarter of occupants, half the occupants, one-quarter of occupants and one-eight of the occupants.

Expected result: The relationship between walking speeds and people densities in zone 2 as well as the flows in line A vs people densities in zone 2 are plotted and compared with the underlying assumptions used in the evacuation model.

Test method: The test method is a qualitative verification of the crowd movement.

User's actions: The tester may show results in relation to different time intervals adopted for the estimation of flows, people densities and walking speeds. Different methods for the implementation of physical distance in the model can be used (e.g. enforcing distance between agents, setting up the speed/density relationship within the model). Further testing can be done by modifying this test to consider the impact of people with movement disabilities (i.e. some evacuees may have a slower speed and/or occupy a larger space) and attempting to modify the initial number of people further and their initial location.

This test can provide useful information to the crowd evacuation model user concerning the performance of a model considering the physical distance. This test allows, in fact, to investigate maximum achievable flow rates, along with the relationships between the main variables concerning movement (flow, walking speed and density). Similar modified tests could be designed for other key variables which may be impacted by physical distance. This can include testing route choice (i.e. modelling the likelihood of people to re-route due to the presence of congested area), group behaviour (investigating the interactions of people interacting with in-group and out-group members) and flowrates in different components and conditions (e.g. bi-directional flows, flows on vertical components such as stairs).

6 Discussion

This chapter discusses the impact of physical distancing on crowd evacuation movement [36]. Experimental findings concerning people movement at openings assuming two physical distancing provisions (1 m and 2 m) have been used to identify possible implications. The experimental findings indicate that, given our current understanding, a suitable approach is to use the existing design relationships [20] cropped at a given maximum density level (which depends on physical distancing), as shown in [36].

It should be noted that experimental data showed that 1 m of target physical distance might correspond to a slightly higher actual average physical distance (1.24 m), while the 2 m of target physical distance may correspond to a slightly lower actual average distance (1.82 m). This would mean that a user may decide to calibrate the maximum achievable density based on equation 1 or 2 (depending on the assumption on squared or circular packing) but adopting an experimentally observed distance rather than a theoretical one. This would imply that the maximum density at which the curves are cropped would be higher or lower than the theoretical ones. From an evacuation perspective, a conservative assumption would be to assume the higher value of physical distancing between the two of them (e.g. the theoretical or experimental value) since this would yield the lowest density level. This would correspond to a lower value of flow and subsequently higher evacuation times. Nevertheless, it is currently unclear to which extent people would comply with physical distancing provisions in case of emergency and how this is linked to risk perception. For this reason, the current study should be considered applicable for normal circulation conditions, but its use for emergency evacuation scenarios should be the object of further investigation.

A limitation of the experiments in [36] is that they have been conducted with a student population, which is likely to be less compliant to the physical distancing provisions provided [52]. A more compliant population may lead to even higher observed physical distances and, in turn, lower flows. Future studies should investigate a wider variety of population, considering possibly different compliance behaviours, physical and psychological characteristics (including investigating crowd movement at different times of the pandemic and in different countries). Similarly, different physical distancing provisions would need to be investigated.

The experimental data collected also seem to indicate that when people move through doors without reaching the maximum density (e.g. capacity drop), their behaviour tends to be consistent with the pre-pandemic conditions (e.g. the speed/density and flow/density relationships would not need to be completely modified as people tend to decrease their speed at the same density threshold). Nevertheless, the main update in the curve is related to the fact that the capacity drop may not be reached at all [37] due to the lower values of achievable maximum density. Considering, for example, the hydraulic model in the *SFPE Handbook* [20], the maximum flow is achieved at a density level of approximately 2 m, which would not be achievable considering a physical distance equal to or higher than 1 m considering the circular packing assumption (i.e. 2 m of density corresponds to a physical distance of 0.76 m).

On the positive side, lower local densities can be achieved (and possibly lower occupant loads). This would imply that, in principle, a lower number of people may be present in a given space (this being a positive aspect from the crowd evacuation perspective). For example, an area of 100 m^2 would contain only 29 people (0.29 people/m^2), considering a physical distance of 2 m and assuming the hypothetical cylindrical packing. These 29 people would take 83 s to pass through a 1 m door considering the maximum allowed flow of 0.35 pers/m*s of the cropped SFPE curve for a physical distance equal to 2 m. The same space could contain a much larger number of people assuming a higher local density (e.g. 1–2 pers/m^2 would correspond to 100–200 people). This crowd would take 75 s (100 people) or 150 s (200 people), assuming the maximum flow through a door in the SFPE curve (1.33 pers/m*s). This means that there would be a need to revise current evacuation design in case occupant loads would return to normal, but physical distancing provisions would be kept (as this would generate delays in evacuation).

The current experiments and modelling implications discussed in this chapter focus only on uni-directional flows. This is in contrast with previous research studies which have investigated random walks and crossing flows [9, 10]. For this reason, the findings discussed here are deemed relevant for this type of movement. It should also be noted that the experiments and modelling implications described are focused on horizontal movement. The SFPE design curve [20] does not differentiate between doors/openings and corridors, while the experiments seem to indicate that the dataset concerning speed vs inter-person distance (as expected) seems to be more in line with movement at openings rather than in corridors (according to the work by [19]). Therefore, future studies should investigate physical distancing behaviour considering a variety of horizontal and vertical egress components (e.g. staircases).

Crowd models cannot be directly used as they are in the presence of physical distancing, but there is a need by the developers to modify some of their modelling assumptions (e.g. linked to the relationships between speed/density and flow/density, route choice, queuing, etc.), or users would need to modify the existing models to implicitly represent the behaviour that may be observed in a pandemic scenario. In this context, the verification test presented here is deemed useful to evaluate the results produced by evacuation models.

It is important to note that the focus of the current chapter has been on the movement of people in isolation, i.e. the behaviour of pre-existing or emerging groups has not been investigated in detail. This is deemed to have a significant impact on crowd evacuation during a pandemic since physical distancing may be kept by clusters of people rather than individuals. Future research should focus on studying how the nature, type and characteristics of groups can affect evacuation movement.

7 Conclusions

This chapter presents key implications of physical distance related to evacuation design. This includes a discussion on the decrease in the maximum achievable local density, which in turn contributes to decreasing occupant loads, reduced maximum flows and longer evacuation times. This chapter highlights the need for modifying current engineering tools used to design evacuation and the importance of considering the consequences of the impact of physical distancing on the tools adopted (either hand calculation methods or evacuation simulators).

Acknowledgments Enrico Ronchi wishes to acknowledge Cantene srl for providing financial support to Lund University for the work concerning crowd evacuation modelling. He also thanks Rugiada Scozzari and Michele Fronterrè for the fruitful discussions on the topic during the COVID-19 pandemic.

References

1. E. Ronchi and R. Lovreglio, “EXPOSED: An occupant exposure model for confined spaces to retrofit crowd models during a pandemic,” *Saf. Sci.*, vol. 130, p. 104834, Oct. 2020, doi: https://doi.org/10.1016/j.ssci.2020.104834.
2. R. M. Anderson, H. Heesterbeek, D. Klinkenberg, and T. D. Hollingsworth, “How will country-based mitigation measures influence the course of the COVID-19 epidemic?,” *The Lancet*, vol. 395, no. 10228, pp. 931–934, Mar. 2020, doi: https://doi.org/10.1016/S0140-6736(20)30567-5.
3. F. Ahmed, N. Zviedrite, and A. Uzicanin, “Effectiveness of workplace social distancing measures in reducing influenza transmission: a systematic review,” *BMC Public Health*, vol. 18, no. 1, p. 518, Dec. 2018, doi: 10.1186/s12889-018-5446-1.
4. L. Matrajt and T. Leung, “Evaluating the Effectiveness of Social Distancing Interventions to Delay or Flatten the Epidemic Curve of Coronavirus Disease,” *Emerg. Infect. Dis.*, vol. 26, no. 8, pp. 1740–1748, Aug. 2020, doi: https://doi.org/10.3201/eid2608.201093.
5. World Health Organization, “Coronavirus disease (COVID-19) advice for the public.” WHO, 2020, Accessed: Oct. 13, 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.

6. L. Movement Strategies, “Comparison of social distancing guidance across countries amid the COVID-19 crisis,” London (UK), 2020.
7. J. Honey-Roses *et al.*, “The Impact of COVID-19 on Public Space: A Review of the Emerging Questions,” Open Science Framework, preprint, Apr. 2020. doi: 10.31219/osf.io/rf7xa.
8. E. Ronchi, R. Scozzari, and M. Fronterrè, “A risk analysis methodology for the use of crowd models during the Covid-19 pandemic,” Lund, Sweden, Technical Report 3235, 2020.
9. C. A. Pouw, F. Toschi, F. van Schadewijk, and A. Corbetta, “Monitoring physical distancing for crowd management: Real-time trajectory and group analysis,” *PloS One*, vol. 15, no. 10, p. e0240963, 2020.
10. I. Echeverría-Huarte, A. Garcimartín, R. C. Hidalgo, C. Martín-Gómez, and I. Zuriguel, “Estimating density limits for walking pedestrians keeping a safe interpersonal distancing,” *Sci. Rep.*, vol. 11, no. 1, p. 1534, Dec. 2021, doi: 10.1038/s41598-020-79454-0.
11. E. Ronchi, R. Lovreglio, R. Scozzari, and M. Fronterrè, “Use of crowd evacuation models in times of pandemic,” *SFPE European Magazine*, 2020.
12. N. Bellomo *et al.*, “A multi-scale model of virus pandemic: Heterogeneous interactive entities in a globally connected world,” *ArXiv Prepr. ArXiv200603915*, 2020.
13. M. D’Orazio, G. Bernardini, and E. Quagliarini, “A probabilistic model to evaluate the effectiveness of main solutions to COVID-19 spreading in university buildings according to proximity and time-based consolidated criteria,” 2020.
14. M. D’Orazio, G. Bernardini, and E. Quagliarini, “Sustainable and resilient strategies for touristic cities against COVID-19: an agent-based approach,” *ArXiv Prepr. ArXiv200512547*, 2020.
15. Á. G. García, A. Caciagli, and others, “Crowd Control in Plazas Constrained to Social Distancing,” *ArXiv Prepr. ArXiv200507038*, 2020.
16. Q. Xu and M. Chraibi, “On the Effectiveness of the Measures in Supermarkets for Reducing Contact among Customers during COVID-19 Period,” *Sustainability*, vol. 12, no. 22, p. 9385, 2020.
17. J. Vacková and M. Bukáček, “Social and physical pedestrian sizes and their impact on decision-based modeling,” presented at the FEMTC2020, Online, 2020.
18. R. Lovreglio, E. Ronchi, and M. J. Kinsey, “An Online Survey of Pedestrian Evacuation Model Usage and Users,” *Fire Technol.*, vol. 56, no. 3, pp. 1133–1153, May 2020, doi: https://doi.org/10.1007/s10694-019-00923-8.
19. V. M. Predtechenskii and A. I. Milinskii, *Planning for foot traffic flow in buildings*. Amerind Publishing, 1978.
20. S. M. V. Gwynne and E. R. Rosenbaum, “Employing the Hydraulic Model in Assessing Emergency Movement,” in *SFPE Handbook of Fire Protection Engineering*, M. J. Hurley, D. T. Gottuk, J. R. Hall, K. Harada, E. D. Kuligowski, M. Puchovsky, J. L. Torero, J. M. Watts, and C. J. Wieczorek, Eds. New York, NY: Springer New York, 2016, pp. 2115–2151.
21. K. Bladström, “Route choice modelling in fire evacuation simulators,” *LUTVDG/TVBB*, 2017.
22. R. Lovreglio, A. Fonzone, and L. dell’Olio, “A mixed logit model for predicting exit choice during building evacuations,” *Transp. Res. Part Policy Pract.*, vol. 92, pp. 59–75, Oktober 2016, doi: https://doi.org/10.1016/j.tra.2016.06.018.
23. E. Ronchi, “Developing and validating evacuation models for fire safety engineering,” *Fire Saf. J.*, p. 103020, May 2020, doi: 10.1016/j.firesaf.2020.103020.
24. A. Seyfried, B. Steffen, W. Klingsch, T. Lippert, and M. Boltes, “The Fundamental Diagram of Pedestrian Movement Revisited — Empirical Results and Modelling,” in *Traffic and Granular Flow’05*, A. Schadschneider, T. Pöschel, R. Kühne, M. Schreckenberg, and D. E. Wolf, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 305–314.
25. T. Korhonen and S. Hostikka, “Fire Dynamics Simulator with Evacuation: FDS+Evac Technical Reference and User‘s Guide,” VTT Technical Research Center of Finland, Working paper 119, 2009.
26. P. A. Thompson and E. W. Marchant, “Testing and application of the computer model ‘SIMULEX,’” *Fire Saf. J.*, vol. 24, no. 2, pp. 149–166, Jan. 1995, doi: https://doi.org/10.1016/0379-7112(95)00020-T.

27. S. Hostikka, T. Korhonen, T. Paloposki, T. Rinne, K. Matikainen, and S. Heliövaara, *Development and validation of FDS+Evac for evacuation simulations*. VTT Technical Research Center of Finland, 2007.
28. J. J. Fruin, *Pedestrian Planning and Design*, (Revised Edition). Elevator World, Inc, Mobile, AL, 1987.
29. P. Thompson, D. Nilsson, K. Boyce, M. Molloy, and D. McGrath, "Exploring the biomechanics of walking and crowd 'flow,'" *Fire Mater.*, vol. 44, no. 6, pp. 879–893, Oct. 2020, doi: https://doi.org/10.1002/fam.2889.
30. C. Feliciani and K. Nishinari, "Empirical analysis of the lane formation process in bidirectional pedestrian flow," *Phys. Rev. E*, vol. 94, no. 3, Sep. 2016, doi: 10.1103/PhysRevE.94.032304.
31. D. Helbing, L. Buzna, A. Johansson, and T. Werner, "Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions," *Transp. Sci.*, vol. 39, no. 1, pp. 1–24, Feb. 2005, doi: https://doi.org/10.1287/trsc.1040.0108.
32. H. Murakami, C. Feliciani, and K. Nishinari, "Lévy walk process in self-organization of pedestrian crowds," *J. R. Soc. Interface*, vol. 16, no. 153, p. 20180939, Apr. 2019, doi: 10.1098/rsif.2018.0939.
33. J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, "Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram," *J. Stat. Mech. Theory Exp.*, vol. 2012, no. 02, p. P02002, 2012.
34. K. Kitazawa and T. Fujiyama, "Pedestrian Vision and Collision Avoidance Behavior: Investigation of the Information Process Space of Pedestrians Using an Eye Tracker," in *Pedestrian and Evacuation Dynamics 2008*, W. W. F. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 95–108.
35. N. Bellomo, D. Clarke, L. Gibelli, P. Townsend, and B. J. Vreugdenhil, "Human behaviours in evacuation crowd dynamics: From modelling to 'big data' toward crisis management," *Phys. Life Rev.*, vol. 18, pp. 1–21, Sep. 2016, doi: https://doi.org/10.1016/j.plrev.2016.05.014.
36. K. J. Marshall and M. L. Register, "Crowd Movement in a Pandemic Situation." University of Canterbury, 2020.
37. E. M. Cepolina, "Phased evacuation: An optimisation model which takes into account the capacity drop phenomenon in pedestrian flows," *Fire Saf. J.*, vol. 44, no. 4, pp. 532–544, Maj 2009, doi: 10.1016/j.firesaf.2008.11.002.
38. E. D. Kuligowski, *Terror Defeated: Occupant Sensemaking, Decision-making and Protective Action in the 2001 World Trade Center Disaster*. University of Colorado at Boulder, 2011.
39. R. Lovreglio, E. Ronchi, G. Maragkos, T. Beji, and B. Merci, "A dynamic approach for the impact of a toxic gas dispersion hazard considering human behaviour and dispersion modelling," *J. Hazard. Mater.*, vol. 318, pp. 758–771, Nov. 2016, doi: https://doi.org/10.1016/j.jhazmat.2016.06.015.
40. B. J. Meacham and R. L. Custer, *Performance-based fire safety engineering: an introduction of basic concepts*. Sage Publications Sage CA: Thousand Oaks, CA, 1995.
41. E. Ronchi and D. Nilsson, "Basic Concepts and Modelling Methods," in *Evacuation Modeling Trends*, A. Cuesta, O. Abreu, and D. Alvear, Eds. Cham: Springer International Publishing, 2016, pp. 1–23.
42. D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E*, vol. 51, no. 5, pp. 4282–4286, Maj 1995, doi: 10.1103/PhysRevE.51.4282.
43. C. W. Reynolds, "Steering Behaviors For Autonomous Characters," 1999, pp. 763–782.
44. M. Plaue, G. Bärwolff, and H. Schwandt, "On measuring pedestrian density and flow fields in dense as well as sparse crowds," Zurich, 2012.
45. L. A. Hayduk, "Personal space: An evaluative and orienting overview.," *Psychol. Bull.*, vol. 85, no. 1, pp. 117–134, 1978, doi: https://doi.org/10.1037/0033-2909.85.1.117.
46. H. Hecht, R. Welsch, J. Viehoff, and M. R. Longo, "The shape of personal space," *Acta Psychol. (Amst.)*, vol. 193, pp. 113–122, Feb. 2019, doi: 10.1016/j.actpsy.2018.12.009.
47. E. Ronchi, "Developing and validating evacuation models for fire safety engineering," *Fire Saf. J.*, 2020.

48. T. Meyer-König, N. Waldau, and H. Klüpfel, "The RiMEA Project — Development of a new Regulation," in *Pedestrian and Evacuation Dynamics 2005*, N. Waldau, P. Gattermann, H. Knoflacher, and M. Schreckenberg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 309–313.
49. I. M. Organization, *Revised guidelines on Evacuation Analysis for New and Existing Passenger Ships, MSC.1/Circ.1533*. International Maritime Organization, 2016.
50. E. Ronchi, E. D. Kuligowski, D. Nilsson, R. D. Peacock, and P. A. Reneke, "Assessing the Verification and Validation of Building Fire Evacuation Models," *Fire Technol.*, vol. 52, pp. 197–219, Sep. 2014, doi: https://doi.org/10.1007/s10694-014-0432-3.
51. International Standards Organization, "Fire Safety Engineering - Verification and validation protocol for building fire evacuation models ISO/DIS 20414." ISO, 2020.
52. A. Nivette *et al.*, "Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study," *Soc. Sci. Med.*, vol. 268, p. 113370, 2021.

A Kinetic Theory Approach to Model Crowd Dynamics with Disease Contagion

Daewa Kim and Annalisa Quaini

Abstract We present some ideas on how to extend a kinetic-type model for crowd dynamics to account for an infectious disease spreading. We focus on a medium size crowd occupying a confined environment where the disease is easily spread. The kinetic theory approach we choose uses tools of game theory to model the interactions of a person with the surrounding people and the environment, and it features a parameter to represent the level of stress. It is known that people choose different walking strategies when subjected to fear or stressful situations. To demonstrate that our model for crowd dynamics could be used to reproduce realistic scenarios, we simulate passengers in one terminal of Hobby Airport in Houston. In order to model disease spreading in a walking crowd, we introduce a variable that denotes the level of exposure to people spreading the disease. In addition, we introduce a parameter that describes the contagion interaction strength and a kernel function that is a decreasing function of the distance between a person and a spreading individual. We test our contagion model on a problem involving a small crowd walking through a corridor.

1 Introduction

We are interested in studying the early stage of an infectious disease spreading in an intermediate size population occupying a confined environment, such as an airport terminal or a school, for a short period of time (minutes or hours). Classical models in epidemiology use mean-field approximations based on averaged large population behaviors over a long time span (typically weeks or months). Obviously, such models fail when population size is small to medium. Our overarching goal is

D. Kim
Department of Mathematics, West Virginia University, Morgantown, WV, USA
e-mail: daewa.kim@mail.wvu.edu

A. Quaini (✉)
Department of Mathematics, University of Houston, Houston, TX, USA
e-mail: quaini@math.uh.edu

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_7

to model the spreading of a disease in a walking crowd by extending a kinetic theory approach for crowd dynamics that compares favorably with experimental data for a medium-sized population [22]. In this chapter, we present the key features of our crowd dynamics model and preliminary ideas for the extension, which are tested in 1D cases. We assume that the disease is such that it spreads with close proximity of individuals, like, e.g., measles, influenza, or COVID-19.

The reason why we focus on a mesoscopic model for crowd dynamics is related to our interest in confined environments and medium-sized crowds. In broad terms, the large variety of models proposed over the years can be divided into three main categories depending on the (microscopic, mesoscopic, or macroscopic) scale of observation [4]. Macroscopic models (see, e.g., [16, 20, 27]) treat the crowd as a continuum flow, which is well suited for large-scale dense crowds. Microscopic models (see, e.g., [3, 5, 14, 15, 17, 19, 25, 30], and the references therein) use Newtonian mechanics to interpret pedestrian movement as the physical interaction between the people and the environment. Mesoscale models (see, e.g., [1, 6–8, 10–12]) use a Boltzmann-type evolution equation for the statistical distribution function of the position and velocity of pedestrians, in a framework close to that of the kinetic theory of gases. There is one key difference though: the interactions in Boltzmann equations are conservative and reversible, while the interactions in the kinetic theory of active particles are irreversible, non-conservative, and, in some cases, non-local and nonlinearly additive. An important consequence is that often for active particles the Maxwellian equilibrium does not exist [2]. Another reason why we choose to work with a kinetic-type model is the flexibility in accounting for multiple interactions (hard to achieve in microscopic models) and heterogeneous behavior in people (hard to achieve in macroscopic models). Finally, we would like to mention that multiscale approaches are possible as well. See, e.g., [4] for a multiscale vision to human crowds which provides a consistent description at the three possible modeling scales.

The first part of this manuscript is dedicated to a description of a crowd dynamics model. The model, first presented in [22], is based on earlier works [1, 10] and is capable of handling evacuation from a room with one or more exits of variable size and in presence of obstacles. The main ingredients of the model are the following: (i) discrete velocity directions to take into account the granular feature of crowd dynamics, (ii) interactions modeled using tools of stochastic games, and (iii) heuristic, deterministic modeling of the speed corroborated by experimental evidence [26]. In [22], we show that for groups of 40 to 138 people the average people density and flow rate computed with our kinetic model are in great agreement with the respective measured quantities reported in a recent empirical study focused on egressing from a facility [28].

To demonstrate that our model for crowd dynamics could be used to reproduce realistic scenarios, we simulate passengers in one terminal of Hobby Airport in Houston (USA). In a first set of tests, the passengers from two planes at the opposite ends of the terminal walk through the terminal to reach the exit. In the second set of tests, we add a group of passengers that enters the terminal through the entrance at the same time as the other two groups deplane and is directed to a gate. The aim of both sets of tests is to understand how the presence of obstacles in the terminal affects the egress time. Obviously, the longer one stays in the terminal in close proximity with other individuals the more likely he/she gets infected. Thus, the egress time from a potentially crowded confined environment, such as an airport terminal, is a key factor in the early spreading of an airborne disease. This is why we chose these tests and how they are connected to the second part of the chapter.

In order to model disease spreading in a walking crowd, we take inspiration from the work on emotional contagion (i.e., spreading of fear or panic) in [29]. We introduce a variable that denotes the level of exposure to people spreading the disease, with the underlying idea that the more a person is exposed the more likely they are to get infected. The model features a parameter that describes the contagion interaction strength and a kernel function that is a decreasing function of the distance between a person and a spreading individual. As a simplification, we assume that walking speed and direction are given. We will show preliminary results for a problem involving a small crowd walking through a corridor. The simplifying assumption will be removed in a follow-up paper, where the people dynamics will be provided by the complex pedestrian model described in the first part of the chapter. The approach we have in mind is different from what we used in [23]. Therein, we coupled the pedestrian dynamics model to a disease contagion model, while in the future we intend to add to the pedestrian dynamics model terms that account for disease spreading.

For related work on coupled dynamics of virus infection and healthy cells, see, e.g., [13], and the references therein. A multiscale model of virus pandemic accounting for the interaction of different spatial scales (from the small scale of the virus itself and cells to the large scale of individuals and further up to the collective behavior of populations) is presented in [9].

The chapter is organized as follows. Section 2 describes the crowd dynamics model and its full discretization and shows numerical results in an airport terminal. In Sect. 3, we introduce our simplified contagion model. The discretization and preliminary results are also shown in Sect. 3. Conclusions are drawn in Sect. 4.

2 A Kinetic Model for Crowd Dynamics

Let $\Omega \subset \mathbb{R}^2$ denote a bounded domain where people are walking to reach an exit E that is either within the domain or belongs to the boundary $\partial\Omega$. The case of multiple exits (i.e., E is the finite union of disjoint sets) can be easily handled as well. The rest of the boundary is made of walls, denoted with W. Walls and other kinds of obstacles could be present within the domain. Let $\boldsymbol{x} = (x, y)$ denote position and $\boldsymbol{v} = v(\cos\theta, \sin\theta)$ denote velocity, where v is the velocity modulus and θ is the velocity direction. For a large group of people inside Ω, let

$$f = f(t, \boldsymbol{x}, v, \theta) \quad \text{for all } t \geq 0,\ \boldsymbol{x} \in \Omega,\ v \in [0, V_M],\ \theta \in [0, 2\pi),$$

where V_M is the largest speed a person can reach in low density and optimal environmental conditions. Under suitable integrability conditions, $f(t, \boldsymbol{x}, v, \theta) d\boldsymbol{x} dv d\theta$ represents the number of individuals who, at time t, are located in the infinitesimal rectangle $[x, x + dx] \times [y, y + dy]$ and have a velocity belonging to $[v, v + dv] \times [\theta, \theta + d\theta]$.

For simplicity and following [1], we make two simplifying assumptions on the velocity vector:

1. Variable θ is discrete, i.e., it can take values in the set:

$$I_\theta = \left\{ \theta_i = \frac{i-1}{N_d} 2\pi : i = 1, \ldots, N_d \right\},$$

   where N_d is the maximum number of possible directions.
2. People adjust their walking speed v depending on the level of congestion around them, i.e., v is treated as a deterministic variable.

The second assumption is corroborated by experimental studies that show that the walking speed mainly depends on the local level of congestion. Given the deterministic nature of the variable v, the distribution function can be written as

$$f(t, \boldsymbol{x}, \theta) = \sum_{i=1}^{N_d} f^i(t, \boldsymbol{x}) \delta(\theta - \theta_i), \tag{1}$$

where $f^i(t, \boldsymbol{x}) = f(t, \boldsymbol{x}, \theta_i)$ represents the people that, at time t and position $\boldsymbol{x}$, move with direction θ_i. In Eq. (1), δ denotes the Dirac delta function.

In the rest of the chapter, we will work with dimensionless variables. To this purpose, we introduce the following reference quantities:

- D: the largest distance a pedestrian can cover in domain Ω
- T: a reference time given by D/V_M (recall that V_M is the largest speed a person can reach)
- ρ_M: the maximal admissible number of pedestrians per unit area

The dimensionless variables are then: position $\hat{\boldsymbol{x}} = \boldsymbol{x}/D$, time $\hat{t} = t/T$, velocity modulus $\hat{v} = v/V_M$, and distribution function $\hat{f} = f/\rho_M$. For ease of notation, the hats will be omitted with the understanding that all variables are dimensionless from now on.

Due to the normalization of f and the f_i, $i = 1, \ldots, N_d$, the dimensionless local density is obtained by summing the distribution functions over the set of directions:

$$\rho(t, \boldsymbol{x}) = \sum_{i=1}^{N_d} f^i(t, \boldsymbol{x}). \tag{2}$$

Following assumption 2 mentioned above, the walking speed is given by $v = v[\rho](t, \boldsymbol{x})$, where square brackets are used to denote that v depends on ρ in a functional way. For instance, v can depend on ρ and on its gradient.

In order to define the walking speed, we introduce a parameter $\alpha \in [0, 1]$ to represent the quality of the walkable domain: where $\alpha = 1$ people can walk at the desired speed (i.e., V_M) because the domain is clear, while where $\alpha = 0$ people are forced to slow down or stop because an obstruction is present. For simplicity, we assume that the maximum dimensionless speed a person can reach is equal to α. Let ρ_c be a critical density value such that for $\rho < \rho_c$ we have free flow regime (i.e., low density condition), while for $\rho > \rho_c$, we have a slowdown zone (i.e., high density condition). Following the experimentally measured values of ρ_c reported in [26], we set $\rho_c = \alpha/5$. Then, we set the walking speed v equal to α in the free flow regime and equal to a heuristic third-order polynomial in the slowdown zone:

$$v = v(\rho) = \begin{cases} \alpha & \text{for} \quad \rho \leq \rho_c(\alpha) = \alpha/5 \\ a_3\rho^3 + a_2\rho^2 + a_1\rho + a_0 & \text{for} \quad \rho > \rho_c(\alpha) = \alpha/5, \end{cases} \tag{3}$$

where a_i is constant for $i = 0, 1, 2, 3$. To set the value of these constants, we impose the following conditions: $v(\rho_c) = \alpha$, $\partial_\rho v(\rho_c) = 0$, $v(1) = 0$, and $\partial_\rho v(1) = 0$, which lead to

$$\begin{cases} a_0 &= (1/(\alpha^3 - 15\alpha^2 + 75\alpha - 125))(75\alpha^2 - 125\alpha) \\ a_1 &= (1/(\alpha^3 - 15\alpha^2 + 75\alpha - 125))(-150\alpha^2) \\ a_2 &= (1/(\alpha^3 - 15\alpha^2 + 75\alpha - 125))(75\alpha^2 + 375\alpha) \\ a_3 &= (1/(\alpha^3 - 15\alpha^2 + 75\alpha - 125))(-250\alpha). \end{cases} \tag{4}$$

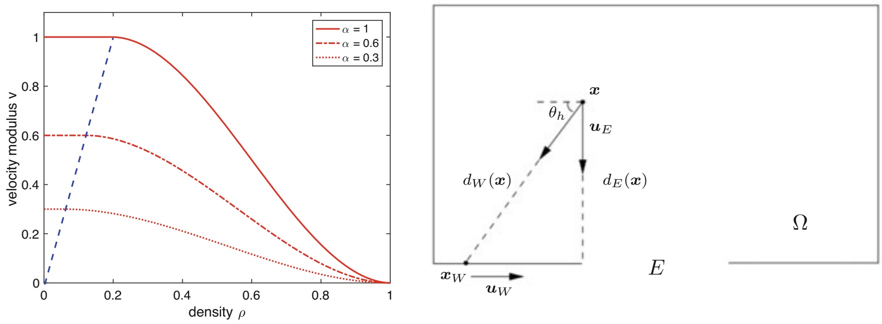
Figure 1a shows v as a function of ρ for $\alpha = 0.3, 0.6, 1$.

2.1 Modeling Interactions

Let us consider the scenario depicted in Fig. 1b, where there is a person Located at a point $\boldsymbol{x}$ that needs to reach exit E. We model the path this person takes as the result of four factors:

(F1) The goal to reach the exit
(F2) The desire to avoid collisions with the walls
(F3) The tendency to look for less congested areas
(F4) The tendency to follow the stream or herding

Factors F1 and F2 are related to geometric aspects of the domain, while factors F3 and F4 consider that people's behavior is strongly affected by the surrounding crowd. These last two factors are dominant in different situations: F4 emerges in



(a) v as a function of ρ

(b) Notation in a sample computational domain

Fig. 1 (**a**) Dependence of the dimensionless walking speed v on the dimensionless density ρ for different values of the parameter α, which represents the quality of the walkable domain. (**b**) Sketch of computational domain Ω with exit E and a pedestrian located at $\boldsymbol{x}$, moving with direction θ_h. The pedestrian should choose direction $\boldsymbol{u}_E$ to reach the exit, while direction $\boldsymbol{u}_W$ is to avoid collision with the wall. The distances from the exit and from the wall are d_E and d_W, respectively

stressful situations, while F3 characterizes normal behavior. To weight between F3 and F4, we use parameter $\varepsilon \in [0, 1]$ with $\varepsilon = 0$ (respectively, $\varepsilon = 1$) if F3 (respectively, F4) prevails.

In order to explain how the four factors are modeled, we need to introduce some terminology. Interactions involve three types of people:

- *Test people* with state $(\boldsymbol{x}, \theta_i)$: they are representative of the whole system.
- *Candidate people* with state $(\boldsymbol{x}, \theta_h)$: they can reach, in probability, the state of the test people after individual-based interactions with the environment or with field people.
- *Field people* with state $(\boldsymbol{x}, \theta_k)$: their interactions with candidate people trigger a possible change of state.

We note that while the candidate person modifies their state, in probability, into that of the test person due to interactions with field people, the test person loses their state as a result of these interactions.

Next, we introduce some notation. Given a candidate person at point $\boldsymbol{x}$ in the walkable domain Ω, we define the distance to the exit as

$$d_E(\boldsymbol{x}) = \min_{\boldsymbol{x}_E \in E} ||\boldsymbol{x} - \boldsymbol{x}_E||,$$

and we consider the unit vector $\boldsymbol{u}_E(\boldsymbol{x})$, pointing from $\boldsymbol{x}$ to the exit; see Fig. 1b. Both d_E and $\boldsymbol{u}_E$ will be used to model (F1).

Assume that the candidate person at $\boldsymbol{x}$ is moving with direction θ_h. We define the distance $d_W(\boldsymbol{x}, \theta_h)$ from the person to a wall at a point $\boldsymbol{x}_W(\boldsymbol{x}, \theta_h)$ where the person is expected to collide with the wall if they do not change direction. The unit tangent vector $\boldsymbol{u}_W(\boldsymbol{x}, \theta_h)$ to $\partial\Omega$ at $\boldsymbol{x}_W$ points to the direction of the exit; see Fig. 1b. Vector $\boldsymbol{u}_W$ is used to avoid a collision with the walls, i.e., to model (F2).

In order to model (F3), i.e., the decision of candidate person $(\boldsymbol{x}, \theta_h)$ to change direction in order to avoid congested areas, we use the direction that gives the minimal directional derivative of the density at the point $\boldsymbol{x}$. We denote such direction by unit vector $\boldsymbol{u}_C(\theta_h, \rho)$.

Finally, we introduce unit vector $\boldsymbol{u}_F = (\cos\theta_k, \sin\theta_k)$ to model (F4), i.e., the decision of candidate person with direction θ_h to follow a field person with direction θ_k with whom they came into contact.

2.1.1 Interaction with the Walls

We assume that people change direction, in probability, only to an adjacent clockwise or counterclockwise direction in set I_θ. This means a candidate person with walking direction θ_h may choose directions $\theta_{h-1}, \theta_{h+1}$ or keep direction θ_h. For $h = 1$, we set $\theta_{h-1} = \theta_{N_d}$, and for $h = N_d$, we set $\theta_{h+1} = \theta_1$. Let $\mathcal{A}_h(i)$ be the *transition probability*, i.e., the probability that a candidate person with direction θ_h adjusts their direction to θ_i (the direction of the test person) due to the presence of walls and/or an exit. The following constraint for $\mathcal{A}_h(i)$ has to be satisfied:

$$\sum_{i=1}^{N_d} \mathcal{A}_h(i) = 1 \quad \text{for all } h \in \{1, \ldots, N_d\}.$$

The set of all transition probabilities $\mathcal{A} = \{\mathcal{A}_h(i)\}_{h,i=1,\ldots,N_d}$ forms the so-called *table of games* that models the game played by active people interacting with the walls.

We define the vector

$$\boldsymbol{u}_G(\boldsymbol{x}, \theta_h) = \frac{(1 - d_E(\boldsymbol{x}))\boldsymbol{u}_E(\boldsymbol{x}) + (1 - d_W(\boldsymbol{x}, \theta_h))\boldsymbol{u}_W(\boldsymbol{x}, \theta_h)}{||(1 - d_E(\boldsymbol{x}))\boldsymbol{u}_E(\boldsymbol{x}) + (1 - d_W(\boldsymbol{x}, \theta_h))\boldsymbol{u}_W(\boldsymbol{x}, \theta_h)||} = (\cos\theta_G, \sin\theta_G). \tag{5}$$

Here, θ_G is the *geometrical preferred direction*, which is the ideal direction that a person should take in order to reach the exit (factor F1) and avoid the walls (factor F2) in an optimal way. Notice that the closer a person is to an exit (respectively, a wall), the more direction $\boldsymbol{u}_E$ (respectively, $\boldsymbol{u}_W$) weights.

A candidate person with direction θ_h will change their direction by choosing the angle closest to θ_G among the three allowed directions θ_{h-1}, θ_h, and θ_{h+1}. The transition probability is given by

$$\mathcal{A}_h(i) = \beta_h(\alpha)\delta_{s,i} + (1 - \beta_h(\alpha))\delta_{h,i}, \quad i = h-1, h, h+1, \tag{6}$$

where

$$s = \underset{j \in \{h-1, h+1\}}{\arg\min} \{d(\theta_G, \theta_j)\},$$

with

$$d(\theta_p, \theta_q) = \begin{cases} |\theta_p - \theta_q| & \text{if } |\theta_p - \theta_q| \leq \pi, \\ 2\pi - |\theta_p - \theta_q| & \text{if } |\theta_p - \theta_q| > \pi. \end{cases} \tag{7}$$

In (6), δ denotes the Kronecker delta function. Coefficient β_h is defined by

$$\beta_h(\alpha) = \begin{cases} \alpha & \text{if } d(\theta_h, \theta_G) \geq \Delta\theta, \\ \alpha \dfrac{d(\theta_h, \theta_G)}{\Delta\theta} & \text{if } d(\theta_h, \theta_G) < \Delta\theta, \end{cases}$$

where $\Delta\theta = 2\pi/N_d$. The role of β_h is to allow for a transition to θ_{h-1} or θ_{h+1} even in the case that the geometrical preferred direction θ_G is closer to θ_h. Such a transition is more likely to occur the more distant θ_h and θ_G are.

2.1.2 Interaction with Obstacles

In [22], we introduced a strategy to handle obstacles within domain Ω. This strategy uses three ingredients to exclude the real obstacle area from the walkable domain:

1. An effective area: an enlarged area that encloses the real obstacle
2. A definition of $\boldsymbol{u}_W$ to account for the effective area
3. A setting of the parameter α in the effective area depending on the shape of the obstacle

The effective area is necessary especially if the obstacle is close to an exit: it allows to define $\boldsymbol{u}_W$ with respect to a larger area than the obstacle area itself to achieve the goal of having no people walk on the real obstacle area. In [22], we found that the goal is successfully achieved with an effective area four times bigger than the real obstacle area.

Since some pedestrians will walk on part of the effective area, one needs to set parameter α in a suitable way. For a discussion on how to set α to realize different obstacle shapes, we refer to [22].

2.1.3 Interactions Between Pedestrians

As a candidate person with direction θ_h walks, they interact with a field person that moves with direction θ_k. As a result of this interaction, the candidate person can change their direction to θ_i (direction of the test person) in the search for less congested areas if their stress level is low or change to θ_k (direction of the field person) if their stress level is high. The *transition probability* is given by $\mathcal{B}_{hk}(i)[\rho]$.

The following constrain for $\mathcal{B}_{hk}(i)$ has to be satisfied:

$$\sum_{i=1}^{N_d} \mathcal{B}_{hk}(i)[\rho] = 1 \quad \text{for all } h, k \in \{1, \ldots, N_d\},$$

where again the square brackets denote the dependence on the density ρ. Of course, we are still under the assumption that people change direction, in probability, only to an adjacent clockwise or counterclockwise direction in set I_θ.

To take into account the search for a less congested area (factor F3) and the tendency to herd (factor F4), for a candidate person with direction θ_h interacting with a field person with direction θ_k, we define the vector

$$\boldsymbol{u}_P(\theta_h, \theta_k, \rho) = \frac{\varepsilon \boldsymbol{u}_F + (1-\varepsilon)\boldsymbol{u}_C(\theta_h, \rho)}{||\varepsilon \boldsymbol{u}_F + (1-\varepsilon)\boldsymbol{u}_C(\theta_h, \rho)||} = (\cos\theta_P, \sin\theta_P),$$

where the subscript P stands for *people*. Direction θ_P is the *interaction-based preferred direction*, obtained as a weighted combination between the direction of the field person (i.e., $\boldsymbol{u}_F = (\cos\theta_k, \sin\theta_k)$) and the direction pointing to a less crowded area (i.e., $\boldsymbol{u}_C$). The latter direction can be computed for a candidate pedestrian with direction θ_h and located at $\boldsymbol{x}$, by taking

$$C = \underset{j \in \{h-1, h, h+1\}}{\arg\min} \{\partial_j \rho(t, \boldsymbol{x})\},$$

where $\partial_j \rho$ denotes the directional derivative of ρ in the direction given by angle θ_j. We have $\boldsymbol{u}_C(\theta_h, \rho) = (\cos\theta_C, \sin\theta_C)$.

The transition probability is given by

$$\mathcal{B}_{hk}(i)[\rho] = \beta_{hk}(\alpha)\rho\delta_{r,i} + (1 - \beta_{hk}(\alpha)\rho)\delta_{h,i}, \quad i = h-1, h, h+1,$$

where r and β_{hk} are defined by

$$r = \underset{j \in \{h-1, h+1\}}{\arg\min} \{d(\theta_P, \theta_j)\},$$

$$\beta_{hk}(\alpha) = \begin{cases} \alpha & \text{if } d(\theta_h, \theta_P) \geq \Delta\theta \\ \alpha \dfrac{d(\theta_h, \theta_P)}{\Delta\theta} & \text{if } d(\theta_h, \theta_P) < \Delta\theta. \end{cases}$$

We recall that $d(\cdot, \cdot)$ is defined in (7).

2.2 Mathematical Model

Two last ingredients are needed before we can state the mathematical model. These are:

- The *interaction rate* with geometric features $\mu[\rho]$: it models the frequency of interactions between candidate people and the walls and/or obstacles. If the local density is lower, it is easier for pedestrians to see the walls and doors. Thus, we set $\mu[\rho] = 1 - \rho$.
- The *interaction rate* with people $\eta[\rho]$: it defines the number of binary encounters per unit time. If the local density increases, then the interaction rate also increases. For simplicity, we take $\eta[\rho] = \rho$.

The mathematical model is derived from a suitable balance of people in an elementary volume of the space of microscopic states, considering the net flow into such volume due to transport and interactions. We obtain

$$\begin{aligned}
&\frac{\partial f^i}{\partial t} + \nabla \cdot \left(\boldsymbol{v}^i[\rho](t,\boldsymbol{x}) f^i(t,\boldsymbol{x}) \right) \\
&\quad = \mathcal{J}^i[f](t,\boldsymbol{x}) \\
&\quad = \mathcal{J}^i_G[f](t,\boldsymbol{x}) + \mathcal{J}^i_P[f](t,\boldsymbol{x}) \\
&\quad = \mu[\rho] \left(\sum_{h=1}^{n} \mathcal{A}_h(i) f^h(t,\boldsymbol{x}) - f^i(t,\boldsymbol{x}) \right) \\
&\qquad + \eta[\rho] \left(\sum_{h,k=1}^{n} \mathcal{B}_{hk}(i)[\rho] f^h(t,\boldsymbol{x}) f^k(t,\boldsymbol{x}) - f^i(t,\boldsymbol{x}) \rho(t,\boldsymbol{x}) \right) \qquad (8)
\end{aligned}$$

for $i = 1, 2, \ldots, N_d$. Functional $\mathcal{J}^i[f]$ represents the net balance of people that move with direction θ_i due to interactions. Since we consider both the interaction with the environment and with the surrounding people, we can write $\mathcal{J}^i$ as $\mathcal{J}^i = \mathcal{J}^i_G + \mathcal{J}^i_P$, where $\mathcal{J}^i_G$ is an interaction between candidate people and the geometry of the environment and $\mathcal{J}^i_P$ is an interaction between candidate and field people.

Equation (8) is completed with Eq. (2) for the density and eqs. (3) and (4) for the velocity. In the next section, we will discuss a numerical method for the solution of problem (2), (3), (4), (8).

2.3 Full Discretization

The approach we consider is based on a splitting method that decouples the treatment of the transport term and the interaction term in Eq. (8). As usual with splitting methods, the idea is to split the model into a set of subproblems that are easier to solve and for which practical algorithms are readily available. Among the available operator splitting methods, we chose the Lie splitting scheme because it provides a good compromise between accuracy and robustness, as shown in [18].

Let $\Delta t > 0$ be a time discretization step for the time interval $[0, T]$. Denote $t^k = k\Delta t$, with $k = 0, \ldots, N_t$, and let ϕ^k be an approximation of $\phi(t^k)$. Given an initial condition $f^{i,0} = f^i(0, \boldsymbol{x})$, for $i = 1, \ldots, N_d$, the Lie operator splitting scheme applied to problem (8) reads: for $k = 0, 1, 2, \ldots, N_t - 1$, perform the following steps:

– **Step 1**: Find f^i, for $i = 1, \ldots, N_d$, such that

$$\begin{cases} \dfrac{\partial f^i}{\partial t} + \dfrac{\partial}{\partial x}\big((v[\rho]\cos\theta_i) f^i(t, \boldsymbol{x})\big) = 0 \ \text{ on}(t^k, t^{k+1}), \\ f^i(t^k, \boldsymbol{x}) = f^{i,k}. \end{cases} \tag{9}$$

Set $f^{i,k+\frac{1}{3}} = f^i(t^{k+1}, \boldsymbol{x})$.

– **Step 2**: Find f^i, for $i = 1, \ldots, N_d$, such that

$$\begin{cases} \dfrac{\partial f^i}{\partial t} + \dfrac{\partial}{\partial y}\big((v[\rho]\sin\theta_i) f^i(t, \boldsymbol{x})\big) = 0 \ \text{ on}(t^k, t^{k+1}), \\ f^i(t^k, \boldsymbol{x}) = f^{i,k+\frac{1}{3}}. \end{cases} \tag{10}$$

Set $f^{i,k+\frac{2}{3}} = f^i(t^{k+1}, \boldsymbol{x})$.

– **Step 3**: Find f_i, for $i = 1, \ldots, N_d$, such that

$$\begin{cases} \dfrac{\partial f^i}{\partial t} = \mathcal{J}^i[f](t, \boldsymbol{x}) \ \text{ on}(t^k, t^{k+1}), \\ f^i(t^k, \boldsymbol{x}) = f^{i,k+\frac{2}{3}}. \end{cases} \tag{11}$$

Set $f^{i,k+1} = f^i(t^{k+1}, \boldsymbol{x})$.

Once $f^{i,k+1}$ is computed for $i = 1, \ldots, N_d$, we use Eq. (2) to get the density ρ^{k+1} and Eqs. (3) and (4) to get the velocity magnitude at time t^{k+1}.

To complete the numerical method, we need to pick an appropriate numerical scheme for each subproblem.

For simplicity, we present space discretization for computational domain $[0, L] \times [0, H]$, with L and H given. We mesh the domain by choosing Δx and Δy to partition interval $[0, L]$ and $[0, H]$, respectively. Let $N_x = L/\Delta x$ and $N_y = H/\Delta y$. We define the discrete mesh points $\boldsymbol{x}_{pq} = (x_p,\ y_q)$ by

$$x_p = p\Delta x \text{ with } p = 0, 1, \ldots, N_x, \quad y_q = q\Delta y \text{ with } q = 0, 1, \ldots, N_y.$$

It is also useful to define

$$x_{p+1/2} = x_p + \Delta x/2 = \Big(p + \frac{1}{2}\Big)\Delta x, \quad y_{q+1/2} = y_q + \Delta y/2 = \Big(q + \frac{1}{2}\Big)\Delta y.$$

In order to simplify notation of the fully discrete steps 1–3, let us set $\phi = f^i$, $\theta = \theta_i$, $t_0 = t^k$, and $t_f = t^{k+1}$. Let M be a positive integer (≥ 3, in practice). We associate with M a time discretization step $\tau = (t_f - t_0)/M$ and set $t^m = t_0 + m\tau$. The fully discretized version of the Lie splitting algorithm is as follows.

Discrete Step 1
Let $\phi_0 = f^{i,k}$. Problem (9) can be rewritten as

$$\begin{cases} \dfrac{\partial \phi}{\partial t} + \dfrac{\partial}{\partial x}\left((v[\rho]\cos\theta)\phi(t, \boldsymbol{x})\right) = 0 \ \ \text{on}(t_0, t_f), \\ \phi(t_0, \boldsymbol{x}) = \phi_0. \end{cases} \tag{12}$$

We adopt a finite difference method that produces an approximation $\Phi^m_{p,q} \in \mathbb{R}$ of the cell average:

$$\Phi^m_{p,q} \approx \frac{1}{\Delta x\, \Delta y} \int_{y_{q-1/2}}^{y_{q+1/2}} \int_{x_{p-1/2}}^{x_{p+1/2}} \phi(t^m, x, y) dx\, dy,$$

where $m = 1, \ldots, M$, $1 \leq p \leq N_x - 1$, and $1 \leq q \leq N_y - 1$. Given an initial condition ϕ_0, function ϕ^m will be approximated by Φ^m with

$$\Phi^m \Big|_{[x_{p-1/2}, x_{p+1/2}] \times [y_{q-1/2}, y_{q+1/2}]} = \Phi^m_{p,q}.$$

The Lax–Friedrichs method for problem (12) can be written in conservative form as follows:

$$\Phi^{m+1}_{p,q} = \Phi^m_{p,q} - \frac{\tau}{\Delta x}\Big(\mathcal{F}(\Phi^m_{p,q}, \Phi^m_{p+1,q}) - \mathcal{F}(\Phi^m_{p-1,q}, \Phi^m_{p,q})\Big),$$

where

$$\begin{aligned} \mathcal{F}(\Phi^m_{p,q}, \Phi^m_{p+1,q}) = &\frac{\Delta x}{2\tau}(\Phi^m_{p,q} - \Phi^m_{p+1,q}) \\ &+ \frac{1}{2}\Big((v[\rho^m_{p,q}]\cos\theta)\Phi^m_{p,q} + (v[\rho^m_{p+1,q}]\cos\theta)\Phi^m_{p+1,q}\Big). \end{aligned}$$

Discrete Step 2
Let $\phi_0 = f^{i,k+\frac{1}{3}}$. Problem (10) can be rewritten as

$$\begin{cases} \dfrac{\partial \phi}{\partial t} + \dfrac{\partial}{\partial y}\left((v[\rho]\sin\theta)\phi(t, \boldsymbol{x})\right) = 0 \ \ \text{on}(t_0, t_f), \\ \phi(t_0, \boldsymbol{x}) = \phi_0. \end{cases}$$

Similarly to step 1, we use the conservative Lax–Friedrichs scheme:

$$\Phi_{p,q}^{m+1} = \Phi_{p,q}^{m} - \frac{\tau}{\Delta y}\Big(\mathcal{F}(\Phi_{p,q}^{m}, \Phi_{p,q+1}^{m}) - \mathcal{F}(\Phi_{p,q-1}^{m}\Phi_{p,q}^{m})\Big),$$

where

$$\mathcal{F}(\Phi_{p,q}^{m}, \Phi_{p,q+1}^{m}) = \frac{\Delta y}{2\tau}(\Phi_{p,q}^{m} - \Phi_{p,q+1}^{m}) \\ +\frac{1}{2}\Big((v[\rho_{p,q}^{m}]\sin\theta)\Phi_{p,q}^{m} + (v[\rho_{p,q+1}^{m}]\sin\theta)\Phi_{p,q+1}^{m}\Big).$$

Discrete Step 3

Let $\mathcal{J} = \mathcal{J}^i$ and $\phi_0 = f^{i,k+\frac{2}{3}}$. Problem (11) can be rewritten as

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = \mathcal{J}[f](t, \boldsymbol{x}) \ \ \text{on}(t_0, t_f), \\ \phi(t_0, \boldsymbol{x}) = \phi_0. \end{cases}$$

For the approximation of the above problem, we use the forward Euler scheme:

$$\Phi_{p,q}^{m+1} = \Phi_{p,q}^{m} + \tau\Big(\mathcal{J}^m[F^m]\Big),$$

where F^m is the approximation of the reduced distribution function (1) at time t^m.

For stability, the subtime step τ is chosen to satisfy the Courant–Friedrichs–Lewy (CFL) condition (see, e.g., [24]):

$$\max\left\{\frac{\tau}{\Delta x}, \ \frac{\tau}{\Delta y}\right\} \le 1.$$

2.4 Numerical Results

We consider a part of Houston's William P. Hobby Airport as the walkable domain. The terminal has an upside down V shape with eight gates (four per wing) and an entrance/exit at the top, which is 6.8 m wide; see Fig. 2 (top left panel). The shape and the size of the terminal (each wing is about 136 m long and 20 m wide) are realistic, while the number of gates is reduced for simplicity. We consider the following geometries:

- Configuration a: no obstacle in the terminal; see Fig. 2 (top left panel).
- Configuration b: waiting area chairs are located near each terminal; see Fig. 3 (top left panel).

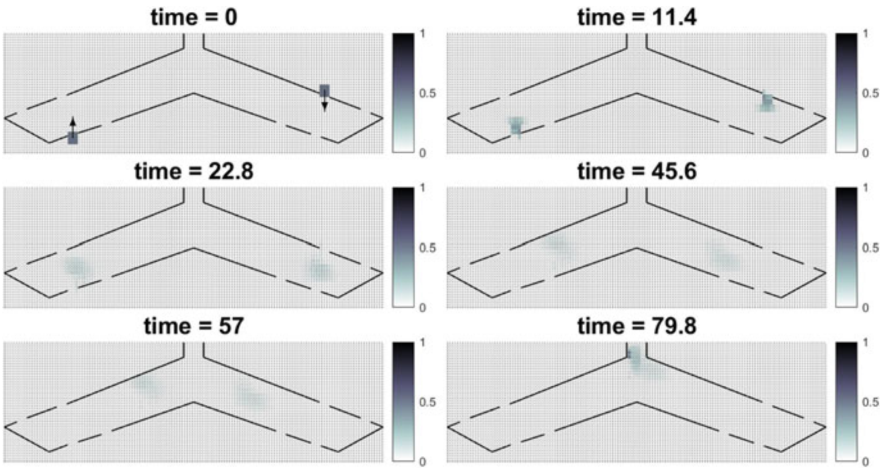


Fig. 2 Test 1a: evacuation process of 404 people grouped into two clusters with initial directions θ_3 (group in the left wing) and θ_7 (group in the right wing)

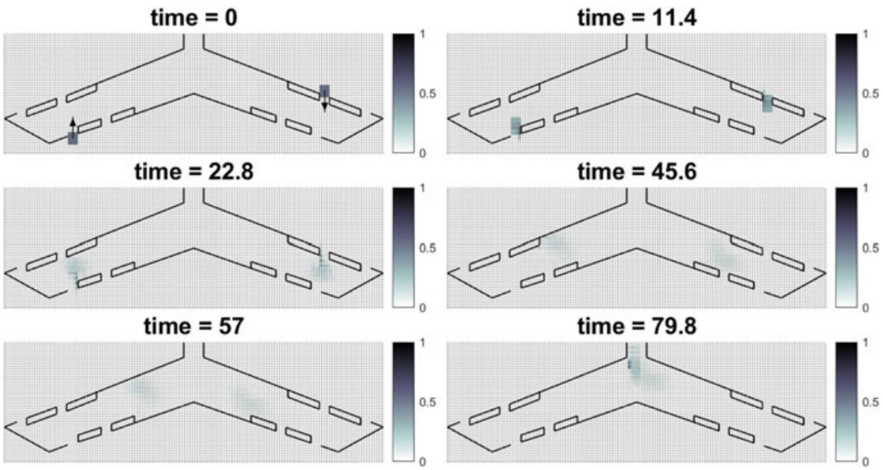


Fig. 3 Test 1b: evacuation process of 404 people grouped into two clusters with initial directions θ_3 (group in the left wing) and θ_7 (group in the right wing)

– Configuration c: in addition to the waiting area chairs, a large obstacle, like a temporary store, is located at the intersection of the two wings; see Fig. 4 (top left panel).

In these configurations, we run two sets of simulations:

– Test 1: a total of 404 passengers from two planes at the opposite ends of the terminal walk through the terminal to reach the exit.

– Test 2: the 404 passengers have the same target as in test 1, but there is an additional group of 202 passengers that enter the terminal through the entrance and are directed to a gate.

The aim is to compute the egress time, i.e., the total time it takes all the passengers to leave the terminal through either the exit or a gate.

For all the simulations, we consider eight different velocity directions $N_d = 8$ in the discrete set:

$$I_\theta = \left\{ \theta_i = \frac{i-1}{8} 2\pi : i = 1, \ldots, 8 \right\}.$$

In order to work with dimensionless quantities, we define the following reference quantities: $D = 137.5$ m, $V_M = 2$ m/s, and $\rho_M = 7$ people/m^2. Once the results are computed, we convert them back to dimensional quantities.

We consider a mesh with $\Delta x = \Delta y = 1.9$ m. The time step is set to $\Delta t = 5.7$ s, and we choose $M = 3$. Figures 2, 3, and 4 show the density computed at different times for tests 1a, 1b, and 1c, respectively. For the large obstacle in configuration c, we use an effective area that is a square with side 15.2 m, while the actual obstacle is a rectangle with dimensions of 9.5 m in length and 4.75 m in width. The reader interested in learning more about how obstacles are handled is referred to [22]. In configuration a, we observe a denser crowd only when the several passengers reach the exit, as shown in Fig. 2 (lower right panel). Configuration b creates dense gatherings also when passengers deplane and their motion is restricted by the waiting area chairs. See Fig. 3 for times $t = 11.4, 22.8$ s. Nonetheless, we observe a similar evacuation dynamics between configurations a and b, indicating that the waiting area chairs do not hinder the evacuation process. Compare Fig. 2 with Fig. 3. This is confirmed by Fig. 5, which shows the number of passengers inside the terminal over time for all the tests. The curves for tests 1a and 1b are either superimposed or very close to each other over the entire time interval. On the other hand, we see that the presence of a large obstacle at the intersection of the two terminal wings increases the egress time by over 10 s. Also, compare Fig. 3 with Fig. 4, bottom right panels.

Because of the similarities in the evacuation process for tests 1a and 1b, we decided to run test 2 only in configurations a and c. The density computed at different times for these two tests is shown in Figs. 6 and 7. In test 2, the exit size is halved because half of the top corridor is used as an entrance. From Figs. 6 and 7 (second row, right panel), we see that by time $t = 45.6$ s, the two groups of passengers with opposite directions (heading to the exit vs to the gate) have met. As expected, halving the exit size leads to a longer evacuation process (for example, compare Figs. 2 and 6) and creates a dense crowd at the exit; see Figs. 6 and 7 (bottom right panel). From Fig. 5, we see that the increase in ingress time from test 1a to 2a is about 30s, while it is about 40s from test 1c to 2c.

The results presented in this subsection corroborate the effectiveness of some strategies adopted in airports during the COVID-19 pandemic: dedicated, distant

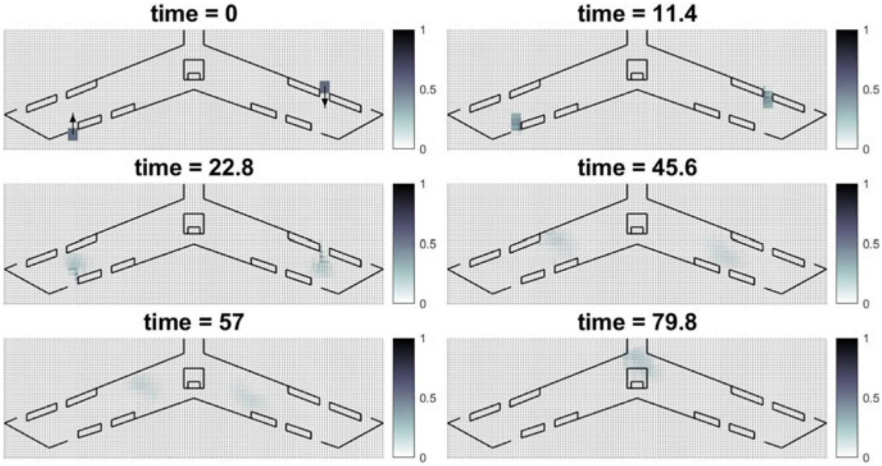
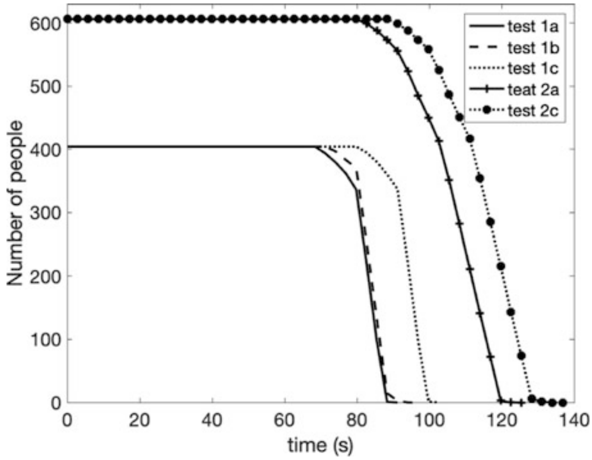


Fig. 4 Test $1c$: evacuation process of 404 people grouped into two clusters with initial directions θ_3 (group in the left wing) and θ_7 (group in the right wing)

Fig. 5 Number of passengers inside the terminal over time for tests $1a$, $1b$, $1c$, $2a$, and $2c$



sites for entrances and exits, and minimization of the obstacles inside the terminal. These strategies are conducive to short egress times and limit congregation points, thereby containing the spreading of COVID-19.

3 Contagion Model in One Dimension

We start from an agent-based model at the microscopic level. We consider a group of N people, N_h of whom are healthy or not spreading the disease yet, while the remaining $N_s = N - N_h$ are in the spreading phase of the disease. If person n

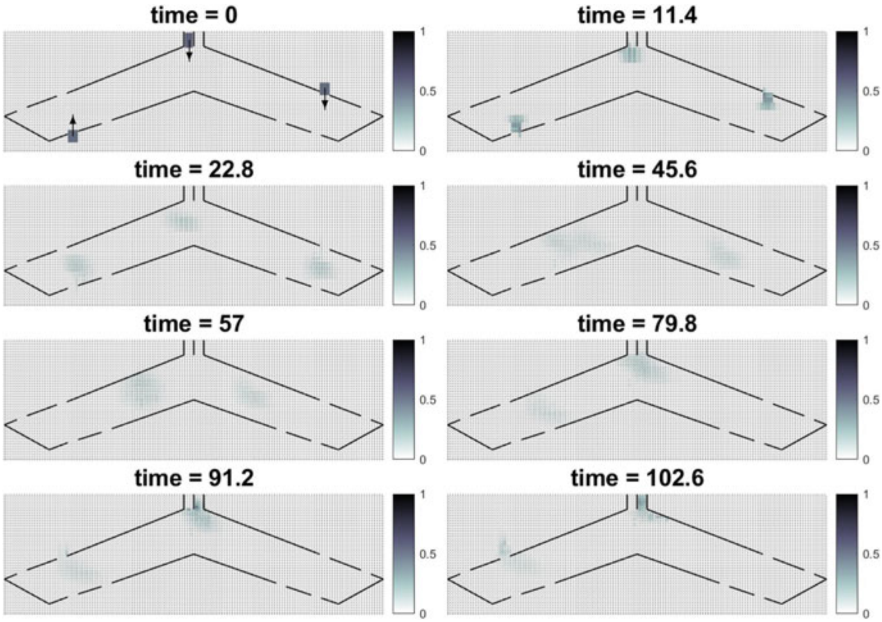


Fig. 6 Test 2a: evacuation process of 404 people grouped into two clusters with initial directions θ_3 (group in the left wing) and θ_7 (group in the right wing) at the same time as a third group of 202 people with initial direction θ_7 enters the airport and is directed to a gate in the left wing

belongs to the former group, we denote with $q_n \in [0, 1)$ their level of exposure to people spreading the disease, with the underlying idea that the more a person is exposed the more likely they are to get infected. If person n belongs to the latter group, then $q_n = 1$, and this value stays constant throughout the entire simulation time. In addition, let $x_n(t)$ and $v_n(t)$ denote the position and speed of person n.

The microscopic model reads, for $n = 1, 2, 3, \ldots, N$,

$$\frac{dx_n}{dt} = v_n \cos\theta_n, \quad \frac{dq_n}{dt} = \gamma \max\{(q_n^* - q_n), 0\}, \quad q_n^* = \frac{\sum_{m=1}^{N} \kappa_{n,m} q_m}{\sum_{m=1}^{N} \kappa_{n,m}}, \tag{13}$$

where the walking speed v_n and walking direction θ_n are given. In the future, we will combine the model in this section with the model presented in Sect. 2 that will provide walking speed and direction. In model (13), q_n^* corresponds to a weighted average "level of sickness" surrounding person n, with $\kappa_{n,m}$ that serves as the weight in the average. We define $\kappa_{n,m}$ as follows:

$$\kappa_{n,m} = \kappa(|x_n - x_m|) = \frac{R}{(|x_n - x_m|^2 + R^2)\pi}. \tag{14}$$

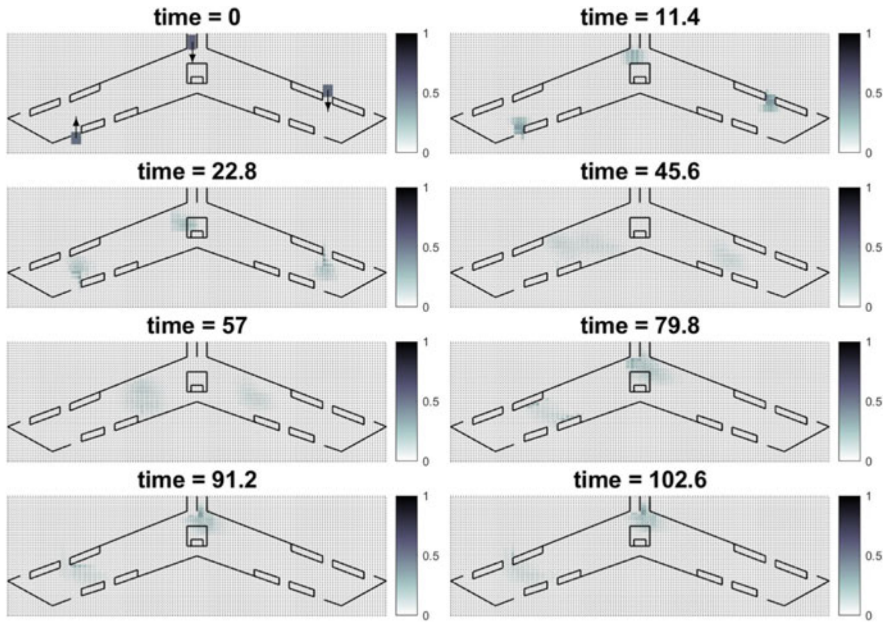


Fig. 7 Test 2*c*: evacuation process of 404 people grouped into two clusters with initial directions θ_3 (group in the left wing) and θ_7 (group in the right wing) at the same time as a third group of 202 people with initial direction θ_7 enters the airport and is directed to a gate in the left wing

Notice that the interaction kernel is a decreasing function of mutual distance between two people and is parametrized by an interaction distance R, set so that the value of $\kappa_{n,m}$ is "small" at about 6 ft or 2 m. Parameter γ in (13) describes the contagion interaction strength: for $\gamma = 0$ there is no contagion, while for $\gamma \neq 0$ the larger the value of γ the faster the contagion. Note that obviously the level of exposure can only increase over time. The second equation in (13) also ensures that the people spreading the disease will constantly have $q_n = 1$ in time.

From the agent-based model (13), we derive a model at the kinetic level. Denote the empirical distribution by

$$h^N = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x_n(t))\delta(q - q_n(t)),$$

where δ is the Dirac delta measure. We assume that the people remain in a fixed compact domain $(x_n(t), q_n(t)) \in \Omega \subset \mathbb{R}^2$ for all n and for the entire time interval under consideration. Prohorov's theorem implies that the sequence $\{h^N\}$ is relatively compact in the weak* sense. Therefore, there exists a subsequence $\{h^{N_k}\}_k$ such that h^{N_k} converges to h with weak*-convergence in $\mathcal{P}(\mathbb{R}^2)$ and pointwise convergence in time as $k \to \infty$. Here, $\mathcal{P}(\mathbb{R}^2)$ denotes the space of probability measures on $\mathbb{R}^2$.

Let $\psi \in C_0^1(\mathbb{R}^2)$ be a test function. We have

$$\begin{aligned}
\frac{d}{dt}\langle h^N, \psi\rangle_{x,q} &= \frac{d}{dt}\left\langle \frac{1}{N}\sum_{n=1}^{N}\delta(x-x_n(t))\delta(q-q_n(t)), \psi\right\rangle_{x,q} \\
&= \frac{d}{dt}\frac{1}{N}\sum_{n=1}^{N}\psi(x_n(t), q_n(t)) \\
&= \frac{1}{N}\sum_{n=1}^{N}\left(\psi_x v_n \cos\theta_n + \psi_q \gamma \max\{(q_n^* - q_n), 0\}\right) \\
&= \langle h^N, \psi_x v \cos\theta_n\rangle_{x,q} + \frac{\gamma}{N}\sum_{n=1}^{N}\psi_q \max\left\{\left(\frac{\sum_{m=1}^{N}\kappa_{n,m}q_n}{\sum_{m=1}^{N}\kappa_{n,m}} - q_n\right), 0\right\}, \qquad (15)
\end{aligned}$$

where $\langle\cdot\rangle_{x,q}$ means integration against both x and q.

Let us define

$$\rho(x) = \frac{1}{N}\sum_{n=1}^{N}\delta(x-x_n)$$

and

$$m(x) = \left\langle q, \frac{1}{N}\sum_{m=1}^{N}\delta(x-x_m)\delta(q-q_m)\right\rangle_{x,q} = \frac{1}{N}\sum_{m=1}^{N}\delta(x-x_m)q_m.$$

We have

$$\frac{1}{N}\sum_{m=1}^{N}\kappa(|x_n-x_m|) = \left\langle \kappa(|x_n-\tilde{x}|), \frac{1}{N}\sum_{m=1}^{N}\delta(\tilde{x}-x_m)\right\rangle_x = \kappa * \rho(x_n),$$

$$\frac{1}{N}\sum_{m=1}^{N}\kappa(|x_n-x_m|)q_m = \left\langle \kappa(|x_n-\tilde{x}|), \frac{1}{N}\sum_{m=1}^{N}\delta(\tilde{x}-x_m)q_m\right\rangle_x = \kappa * m(x_n),$$

where $\langle\cdot\rangle_x$ means integration only in x. Then, we can rewrite Eq. (15) as

$$\frac{d}{dt}\langle h^N, \psi\rangle_{x,q} = \langle h^N, \psi_x v\cos\theta\rangle_{x,q} + \gamma\left\langle h^N, \psi_q \max\left\{\frac{\kappa * m}{\kappa * \rho} - q, 0\right\}\right\rangle_{x,q}. \qquad (16)$$

Via integration by parts, Eq. (16) leads to

$$h_t^N + (v\cos\theta\, h^N)_x + \gamma(\max\{(q^* - q), 0\}h^N)_q = 0, \qquad (17)$$

where q^* is the local *average* sickness level weighted by (14):

$$q^*(t,x) = \frac{\iint \kappa(|x-\overline{x}|)h(t,\overline{x},q)q\,dq\,d\overline{x}}{\iint \kappa(|x-\overline{x}|)h(t,\overline{x},q)\,dq\,d\overline{x}}. \tag{18}$$

Sick people that are in the spreading phase of the disease weight more in the average since they have the highest value of q, nonetheless exposed people contribute to the average level of sickness too since they might spread the virus they recently got exposed to (recall we are simulating short periods of time), e.g., by close contact.

Now letting $k \to \infty$, the subsequence h^{N_k} formally leads to the limiting kinetic equation

$$h_t + (v\cos\theta\ h)_x + \gamma(\max\{(q^*-q),0\}h)_q = 0, \tag{19}$$

where $h(t,x,q)$ is the probability of finding at time t and position $\boldsymbol{x}$ a person with level of exposure q if $q \in [0,1)$ or a person spreading the disease if $q = 1$.

Finally, we note that while modeling motion and disease spreading in one dimension (spatial variable x), Eq. (19) is a 2D problem in variables x and q. Modeling pedestrian motion in two dimensions would lead to a 3D problem that requires a carefully designed numerical scheme to contain the computational costs. This is currently under investigation.

3.1 Full Discretization

We present a space and time discretization for Eq. (19). Let $x \in [0,D]$ and $q \in [0,1]$. Given $N_x = D/\Delta x$, the discrete mesh points x_p are given by

$$x_p = p\Delta x, \quad x_{p+1/2} = x_p + \frac{\Delta x}{2} = \left(p+\frac{1}{2}\right)\Delta x, \tag{20}$$

for $p = 0,1,\dots,N_x$. We partition $[0,1]$ into subintervals $[q_{l-\frac{1}{2}}, q_{l+\frac{1}{2}}]$, with $l \in 1,2,\dots,N_q$, where

$$q_l = l\Delta q, \quad q_{l+1/2} = q_l + \frac{\Delta q}{2} = \left(l+\frac{1}{2}\right)\Delta q.$$

For simplicity, we assume that all subintervals have equal length Δq. The two partitions induce a partition of domain $[0,D]\times[0,1]$ into cells. The time step Δt is chosen as

$$\Delta t \le \min\left\{\frac{\Delta x}{\max_p v_p}, \frac{\Delta q}{2\gamma \max_l q_l}\right\}$$

to satisfy the Courant–Friedrichs–Lewy (CFL) condition.

Let us denote $h_{j,l} = h(t, x_j, q_l)$ and $q_j^* = q^*(t, x_j)$. We consider a first-order semi-discrete upwind scheme for Eq. (19) adapted from one of the methods used in [29], which reads

$$\partial_t h_{j,l} + \frac{\eta_{j,l} - \eta_{j-1,l}}{\Delta x} + \gamma \frac{\xi_{j,l+\frac{1}{2}} - \xi_{j,l-\frac{1}{2}}}{\Delta q} = 0, \tag{21}$$

where

$$\eta_{j,l} = v_j \cos\theta_j \; h_{j,l},$$
$$\xi_{j,l+\frac{1}{2}} = \max\left\{\left(q_j^* - q_{l+\frac{1}{2}}\right), 0\right\} h_{j,l}.$$

For the time discretization of problem (21), we use the forward Euler scheme:

$$h_{j,l}^{m+1} = h_{j,l}^m - \Delta t \left(\frac{\eta_{j,l}^m - \eta_{j-1,l}^m}{\Delta x} + \gamma \frac{\xi_{j,l+\frac{1}{2}}^m - \xi_{j,l-\frac{1}{2}}^m}{\Delta q} \right). \tag{22}$$

The discretization scheme in this section is only first order in space and time. The numerical errors are expected to introduce significant dissipation in the numerical solution. Extension to higher order discretization schemes is possible (see, e.g., [21, 23, 29]) but will not be considered for this chapter.

3.2 Numerical Results

We test the approach presented in Sect. 3.1 on a series of 1D problems, corresponding to unidirectional pedestrian flow in a narrow corridor. For all the problems, the computational domain in the xq-plane is [0, 10] × [0, 1], and it is occupied by a group of 40 people. We set $R = 1$ m since this choice makes the value of the kernel function relatively small at a distance of 2 m (or about 6 ft); see Fig. 8. The dimensionless quantities are obtained by using the following reference quantities: $D = 10$ m, $V_M = 1$ m/s, $T = 10$ s, and $\rho_M = 4$ people/m. In all the tests, we take the initial density to be constant in space and equal to ρ_M.

We take $\Delta x = 0.1$ m and $\Delta q = 0.01$. We will consider two values for contagion strength $\gamma = 100$ and $\gamma = 50$, with the associated respective time steps $\Delta t = 0.00005$ s and $\Delta t = 0.0001$ s. First, we keep the group of people still (i.e., $v = 0$) to observe how the level of exposure to the disease evolves. Then, in a second set of tests, we change to $v = 1$ m/s and see how the motion affects the spreading. We run each simulation for $t \in (0, 10]$ s.

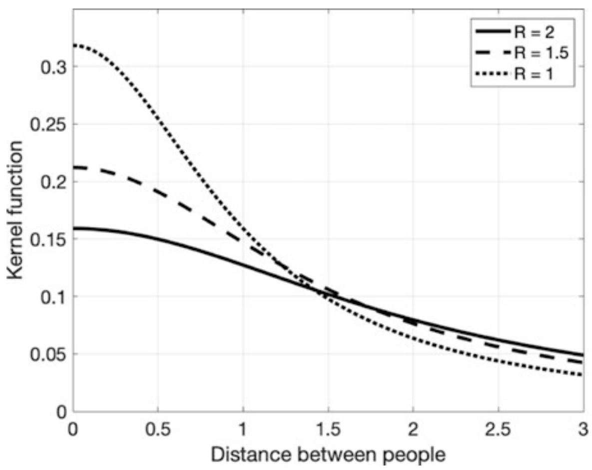


Fig. 8 Kernel function vs the distance between people for interaction radius $R = 1, 1.5, 2$

Tests with $v = 0$ We consider two initial conditions:

– IC1: people that are certainly spreading (i.e., $q = 1$) are located at $x \in [0, 4]$ m and $x \in [6, 10]$ m, while in $x \in (4, 6)$ m we place people that have certainly not been exposed (i.e., $q = 0$).
– IC2: people that are certainly spreading (i.e., $q = 1$) are located at $x \in [0, 2]$ m and $x \in [8, 10]$ m, while the rest of the people located in $x \in (2, 8)$ m have certainly not been exposed (i.e, $q = 0$).

All the healthy people in IC1 are exposed to both groups of spreading people, while in IC2 some healthy people are exposed to one group of spreading people and the centrally located healthy people are not exposed.

Figure 9 shows the evolution of the distribution density h for initial condition IC1 with $\gamma = 100, 50$ and for initial condition IC2 with $\gamma = 50$. We see that the level of exposure of the central group of healthy people in IC1 increases quickly. It increases faster the closer people are to the group of sick people and the larger γ is. Parameter γ plays a central role in the spreading of the disease and would have to be carefully tuned in the future for more realistic applications. The rise in the level of exposure is much slower for the simulation with initial condition IC2. Compare center and bottom rows in Fig. 9. In particular, we notice that the increase in q is very small for the centrally located group of healthy people, as we expected.

This first set of tests was meant to verify our implementation of method described in Sect. 3.1 and to check that the disease spreading term in Eq. (17) (i.e., the third term on the left-hand side) produced the expected outcomes. Next, we are going to get people in motion.

Tests with $v = 1$ m/s We assign to all people walking direction $\theta = 0$, as if they were headed to an exit located at $x = 10$ m. Once spreading people have left the domain, we assume they cannot spread the disease to the people in the domain anymore. We consider IC1 and IC2 and set $\gamma = 50$.

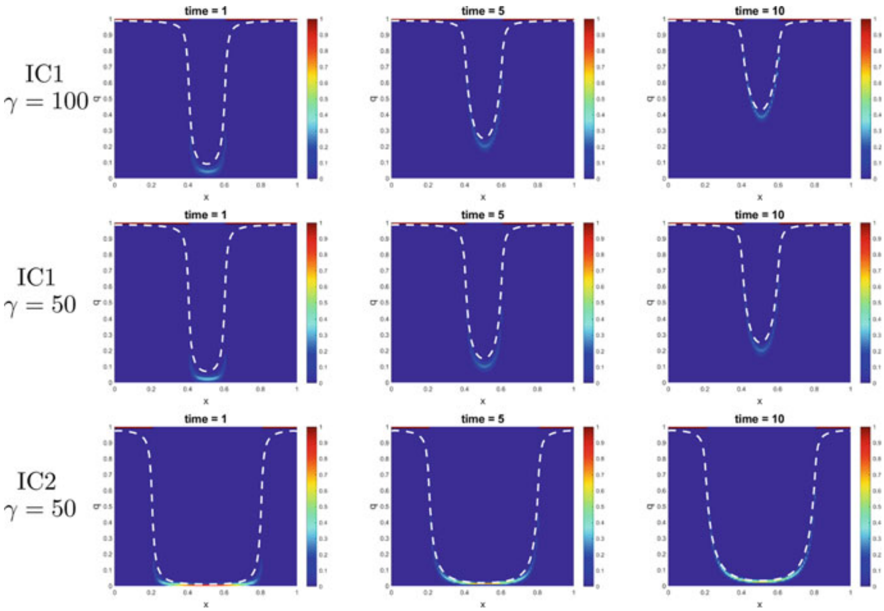


Fig. 9 Tests with $v = 0$: evolution of the distribution density h for initial condition IC1 with $\gamma = 100$ (top) and $\gamma = 50$ (center) and for initial condition IC2 with $\gamma = 50$ (bottom). The white dashed line represents q^*

Figure 10 shows the evolution of the distribution density h for initial conditions IC1 and IC2. We observe that the motion contributes to lowering the exposure level in both the cases, since some of the spreading people leave the domain first. Compare the top and bottom rows of Fig. 10 with the central and bottom rows of Fig. 9.

Finally, we experiment with a slight modification of the initial conditions to show that our model can handle scenarios with uncertainty. The initial conditions are changed:

– IC1-bis: people are positioned like in IC1, but the probabilities of finding people with $q = 1$ and $q = 0$ are reduced from 100% to 60%, and another value of q for a given x is assigned; see Fig. 11 (left panel).
– IC2-bis: people are positioned like in IC2, but the probabilities of finding people with $q = 1$ and $q = 0$ are reduced from 100% to 60%, and another value of q for a given x is assigned; see Fig. 11 (right panel).

Figure 12 shows the evolution of the distribution density h for initial conditions IC1-bis and IC2-bis.

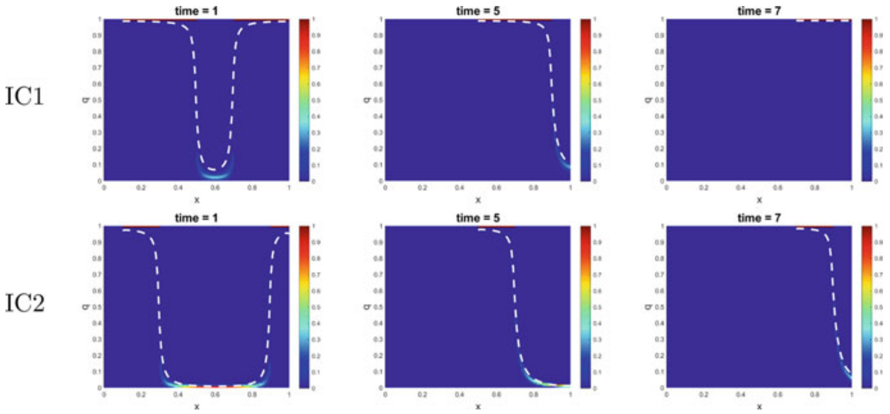


Fig. 10 Tests with $v = 1$ m/s: evolution of the distribution density h for initial condition IC1 (top) and IC2 (bottom). In both cases, we set $\gamma = 50$. The white dashed line represents q^*

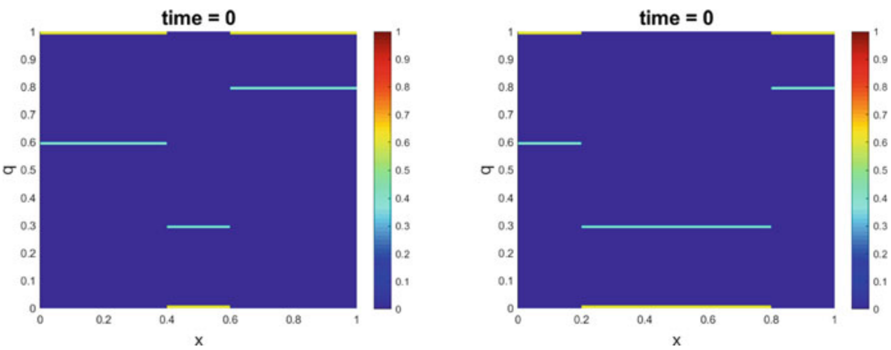


Fig. 11 Tests with $v = 1$ m/s: initial conditions IC1-bis (left) and IC2-bis (right)

4 Conclusions

This chapter is divided into two parts. In the first part, we presented a kinetic-type model for crowd dynamics, while in the second part we introduced a simplified model for disease contagion in a crowd walking through a confined environment.

Kinetic (or mesoscopic) approaches to simulate the motion of medium-sized crowds are appealing because of their flexibility in accounting for multiple interactions (hard to achieve in microscopic models) and heterogeneous behavior in people (hard to achieve in macroscopic models). The particular kinetic model we chose was also shown to compare favorably with experimental data for a medium-sized population. Previously, this model had been used to simulate simple scenarios such as evacuation with a room. In this chapter, we showed that realistic scenarios, such as passengers walking in an airport terminal, can be handled as well.

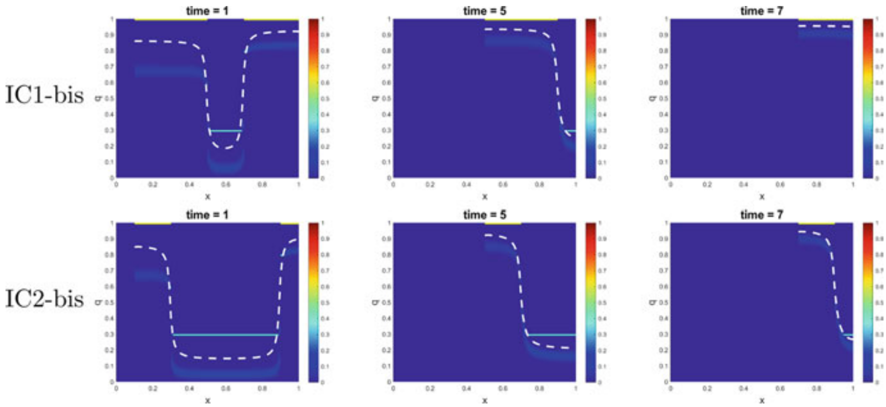


Fig. 12 Tests with $v = 1$ m/s: evolution of the distribution density h for initial conditions IC1-bis (top) and IC2-bis (bottom). In both cases, we set $\gamma = 50$. The white dashed line represents q^*

The simplifying assumptions that we used in the model for disease contagion are that people's walking speed and direction are given. The disease spreading is modeled using three main ingredients: an additional variable that denotes the level of exposure to people spreading the disease, a parameter that describes the contagion interaction strength, and a kernel function that is a decreasing function of the distance between a person and a spreading individual. We tested the proposed contagion model and numerical approach on simple 1D problems.

The obvious next step is to combine the kinetic-type model for crowd dynamics in the first part of the chapter with the disease contagion model in order to drop the simplifying assumption, i.e., walking speed and direction are provided by the model instead of being given.

Acknowledgments This work has been partially supported by NSF through grant DMS-1620384.

References

1. J. P. Agnelli, F. Colasuonno, and D. Knopoff. A kinetic theory approach to the dynamics of crowd evacuation from bounded domains. *Mathematical Models and Methods in Applied Sciences*, 25(01):109–129, 2015.
2. V. V. Aristov. Biological systems as nonequilibrium structures described by kinetic methods. *Results in Physics*, 13:102232, 2019.
3. M. Asano, T Iryo, and M. Kuwahara. Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour. *Transportation Research Part C: Emerging Technologies*, 18(6):842–855, 2010.
4. B. Aylaj, N. Bellomo, L. Gibelli, and A. Reali. A unified multiscale vision of behavioral crowds. *Mathematical Models and Methods in Applied Sciences*, 30(1):1–22, 2020.
5. S. Bandini, S. Manzoni, and G. Vizzari. Agent based modeling and simulation: An informatics perspective. *Journal of Artificial Societies and Social Simulation*, 12(4):4, 2009.

6. N. Bellomo and A. Bellouquid. On the modeling of crowd dynamics: Looking at the beautiful shapes of swarms. *Networks and Heterogeneous Media*, 6(3):383–399, 2011.
7. N. Bellomo, A. Bellouquid, L. Gibelli, and N Outada. *A quest towards a mathematical theory of living systems*. Modeling and Simulation in Science, Engineering and Technology. (Birkhäuser, 2017).
8. N. Bellomo, A. Bellouquid, and D. Knopoff. From the microscale to collective crowd dynamics. *SIAM Multiscale Modeling & Simulation*, 11(3):943–963, 2013.
9. N. Bellomo, R. Bingham, M. K. Chaplain, G. Dosi, G. Forni, D. A. Knopoff, J. Lowengrub, R. Twarock, and M. E. Virgillito. A multiscale model of virus pandemic: Heterogeneous interactive entities in a globally connected world. *Mathematical Models and Methods in Applied Sciences*, 30(8):1591–1651, 2020.
10. N. Bellomo and L. Gibelli. Toward a mathematical theory of behavioral-social dynamics for pedestrian crowds. *Mathematical Models and Methods in Applied Sciences*, 25(13):2417–2437, 2015.
11. N. Bellomo and L. Gibelli. Behavioral crowds: Modeling and Monte Carlo simulations toward validation. *Computers & Fluids*, 141:13–21, 2016.
12. N. Bellomo, L. Gibelli, and N. Outada. On the interplay between behavioral dynamics and social interactions in human crowds. *Kinetic and Related Models*, 12(2):397–409, 2019.
13. N. Bellomo, K. J. Painter, Y. Tao, and M. Winkler. Occurrence vs. absence of taxis-driven instabilities in a May–Nowak model for virus infection. *SIAM Journal on Applied Mathematics*, 79(5):1990–2010, 2019.
14. M. Chraibi, U. Kemloh, A. Schadschneider, and A. Seyfried. Force-based models of pedestrian dynamics. *Networks and Heterogeneous Media*, 6(3):425–442, 2011.
15. M. Chraibi, A. Tordeux, A. Schadschneider, and A. Seyfried. *Modelling of Pedestrian and Evacuation Dynamics*, pages 649–669. (Springer-New York, 2019).
16. E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*. (Springer-Cham, 2014).
17. J. Dai, X. Li, and L. Liu. Simulation of pedestrian counter flow through bottlenecks by using an agent-based model. *Physica A: Statistical Mechanics and its Applications*, 392(9):2202–2211, 2013.
18. R. Glowinski. *Finite element methods for incompressible viscous flow, in Handbook of numerical analysis, P. G. Ciarlet, J. L. Lions (Eds)*, volume 9. (North-Holland, 2003).
19. D. Helbing and P. Molnár. Social force model for pedestrian dynamics. *Physical review. E*, 51:4282–4286, 1998.
20. R. L. Hughes. The flow of human crowds. *Annual Review of Fluid Mechanics*, 35(1):169–182, 2003.
21. D. Kim, K. O'Connell, W. Ott, and A. Quaini. A kinetic theory approach for 2D crowd dynamics with emotional contagion. *Submitted*, 2020. https://arxiv.org/abs/2012.08108.
22. D. Kim and A. Quaini. A kinetic theory approach to model pedestrian dynamics in bounded domains with obstacles. *Kinetic & Related Models*, 12(6):1273–1296, 2019.
23. D. Kim and A. Quaini. Coupling kinetic theory approaches for pedestrian dynamics and disease contagion in a confined environment. *Mathematical Models and Methods in Applied Sciences*, 30(10):1893–1915, 2020.
24. R. J. LeVeque. *Numerical Methods for Conservation Laws*. Lectures in Mathematics ETH Zürich, Department of Mathematics Research Institute of Mathematics. (Springer, 1992).
25. S. Liu, S. Lo, J. Ma, and W. Wang. An agent-based microscopic pedestrian flow simulation model for pedestrian traffic problems. *IEEE Transactions on Intelligent Transportation Systems*, 15(3):992–1001, 2014.
26. A. Schadschneider and A. Seyfried. Empirical results for pedestrian dynamics and their implications for modeling. *Networks and Heterogeneous Media*, 6(3):545–560, 2011.
27. A. Shende, M. P. Singh, and P. Kachroo. Optimization-based feedback control for pedestrian evacuation from an exit corridor. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1167–1176, 2011.

28. A. U. K Wagoum, A. Tordeux, and W. Liao. Understanding human queuing behaviour at exits: An empirical study. *Royal Society Open Science*, 4(1):160896, 2017.
29. L. Wang, M. B. Short, and A. L. Bertozzi. Efficient numerical methods for multiscale crowd dynamics with emotional contagion. *Mathematical Models and Methods in Applied Sciences*, 27(1):205–230, 2017.
30. B. Zhou, X. Wang, and X. Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2871–2878, 2012.

Toward a Quantitative Reduction of the SIR Epidemiological Model

Matteo Colangeli and Adrian Muntean

Abstract Motivated by our intention to use SIR-type epidemiological models in the context of dynamic networks, we investigate in this framework possibilities to reduce the classical SIR model to a representative evolution model for a suitably chosen observable. For selected scenarios, we provide practical a priori error bounds between the approximate and the original observables. Finally, we illustrate numerically the behavior of the reduced models compared to the original ones. As a long-term goal, we would like to apply such techniques in the context of large-scale highly interacting inhomogeneous human crowds.

1 Introduction

The quest of a reduced description from a microscopic dynamics characterized by a large number of degrees of freedom is one of the classical problems of statistical and many-body physics, where model reduction and coarse-graining techniques proved to be a central tool underpinning renormalization group methods [26, 33]. Recently, model reduction techniques also found relevant applications in meteorology [16] and in physical and chemical kinetics [21]. One example is represented by the derivation of the hydrodynamic laws, described in terms of a restricted set of fields (e.g., density, momentum, and temperature), from a kinetic description based on an extended set of moments or on the Boltzmann equation [10–12, 27]. With this occasion, we shall discuss the application of one such method of reduced description, called invariant manifold (IM) method [19], to the SIR model, which

M. Colangeli
Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, L'Aquila, Italy
e-mail: matteo.colangeli1@univaq.it

A. Muntean (✉)
Department of Mathematics and Computer Science, Centre for Societal Risk Research (CSR), Karlstad University, Karlstad, Sweden
e-mail: adrian.muntean@kau.se

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_8

stands as one of the historical benchmarks in the field of epidemic modelling. The approach can be adapted to our epidemiological models described by coupled systems of nonlinear differential equations. As discussed in the sequel, the rationale behind the IM method is based on the identification of a restricted set of fields whose evolution, when observed in the appropriate timescale, captures some distinctive features of the microscopic dynamics of the system.

This work belongs to the recent attempts of the applied mathematics community to understand, from a more fundamental perspective, the spread of viruses, like COVID-19, which are drastically affecting the well-being of our society; see, e.g., [1, 2, 23, 24, 28, 31, 34, 35], to cite but a few. Our own motivation stems from the potential use of SIR-type epidemiological models in the context of dynamic networks as provided by large-scale highly interacting inhomogeneous human crowds. In this framework, we identify possibilities to reduce the classical SIR model to a representative evolution model for a suitably chosen observable. It is worth remarking that, in general, the identification of the relevant observables is a central step of the model reduction procedure, which necessarily requires some physical intuition and some insight into the properties of the system under investigation [30]. One basic guiding principle, in nonequilibrium thermodynamics, relies on distinguishing between slow and fast variables and on retaining only the slowest ones.

The chapter is organized as follows: after a brief discussion of the SIR model and ideas for a reduction is done in Sect. 2, our derivation of simple quantitative estimates allows us to bound from above in Sect. 3.3 the a priori error produced by the proposed reduction strategy. We present in Sects. 4.1 and 4.2 how the Mori–Zwanzig formalism and, respectively, the invariant manifold method work for a simple linear case. Inspired by such an analysis, we propose in Sect. 4.3 an improved reduction method that we test numerically. We close the chapter with a few observations listed in Sect. 5.

2 Basic SIR Model and Its Quantitative Reduction

We focus our attention on the structure of the celebrated SIR model. We refer the reader, for instance, to [8] for a nice description of the modelling ideas behind SIR as well as to [34, 35] for a number of qualitative properties of the solutions to SIR, SIRD, SEIR, and closely related models.

We consider three populations of individuals belonging to a larger population whose total number of individuals is fixed. We shall denote by S, I, and R the fraction of *susceptible, infected,* and *removed* individuals, respectively, such that $S + I + R = 1$.

The model equations entering the SIR model are

$$\frac{dS}{dt} = -b\ S\ I, \tag{1}$$

$$\frac{dI}{dt} = b\ S\ I - \gamma\ I, \tag{2}$$

$$\frac{dR}{dt} = \gamma\ I, \tag{3}$$

where the initial conditions are prescribed as $S(0) = S_0$, $I(0) = I_0$, and $R(0) = R_0$ such that $S_0+I_0+R_0 = 1$. Furthermore, b and γ are here strictly positive parameters that refer to as constant averaged infection rate and constant averaged recovery rate, respectively. As a natural consequence, we see that if $S_0 + I_0 + R_0 = 1$ holds, then we have that also the mass conservation law $S(t) + I(t) + R(t) = 1$ holds for any $t \in (0, T)$, where $T > 0$ is arbitrarily fixed.

We refer to the system of ODEs (1)–(3) as the *original dynamics*.

In this framework, we will offer a couple of reduced variants of this SIR model. In all cases, we rely on the existence and uniqueness of classical positive solutions to the used models.

To obtain a reduced description from the original dynamics, we make the following ansatz: we assume that a suitable timescale exists, in which the time evolution of the *driven* observables is ruled by the dynamics of the *leading* observable.

For the SIR model, we may identify the leading observable either with $\hat{I}(t)$ or with $\hat{S}(t)$, and hence we have the two options:

$$\hat{S}(t) = \Phi[\hat{I}(t)] \quad \text{and} \quad \hat{R}(t) = \Xi[\hat{I}(t)] \tag{4}$$

or

$$\hat{I}(t) = \Psi[\hat{S}(t)] \quad \text{and} \quad \hat{R}(t) = \Omega[\hat{S}(t)]. \tag{5}$$

The discussion of the reduction method based on Eq. (4) is given in Sect. 3, while the analysis of the model corresponding to Eq. (5) is deferred to Sect. 4.3.

3 Using the Constitutive Law $\hat{S}(t) = \Phi[\hat{I}(t)]$

The time evolution of the observables $\hat{S}(t)$, $\hat{I}(t)$, and $\hat{R}(t)$ is dictated by the original dynamics, Eqs. (1)–(3), complemented by the ansatz (4). In particular, the dynamics of $\hat{I}(t)$ reads

$$\frac{d\hat{I}}{dt} = b\ \Phi[\hat{I}]\ \hat{I} - \gamma\ \hat{I}. \tag{6}$$

Equation (7) corresponds to the desired *reduced description* of the original SIR model, in which an expression for the constitutive law $\Phi[\hat{I}]$ is yet to be found. The dynamics of the driven observables $\hat{S}(t)$ and $\hat{R}(t)$ is given by

$$\frac{d\hat{S}}{dt} = -b\ \Phi[\hat{I}]\ \hat{I} \tag{7}$$

$$\frac{d\hat{R}}{dt} = \gamma\ \hat{I}. \tag{8}$$

The initial value problem for the system (6)–(8) is defined by fixing the values $\hat{S}(0) = \hat{S}_0$, $\hat{I}(0) = \hat{I}_0$, and $\hat{R}(0) = \hat{R}_0$. Furthermore, we also set $\Phi[\hat{I}(0)] = \Phi_0$ and $\Xi[\hat{I}(0)] = \Xi_0$.

We look for the exact expression of the functional $\Phi[\hat{I}]$ or, at least, a good approximate version thereof. To this aim, using the ansatz (4), we may also write the time derivative of the observable $\hat{S}(t)$ by relying on the chain rule, namely we have

$$\frac{d\hat{S}}{dt} = \frac{d\hat{I}}{dt}\ \Phi'[\hat{I}], \tag{9}$$

with $\Phi'[\hat{I}] := d\Phi[\hat{I}]/d\hat{I}$, whereas the time derivative of $\hat{I}$ is given by (6).

The IM reduction method stipulates the equality of the two expressions of the time derivative of $\hat{S}(t)$ given in Eqs. (7) and (9). This procedure thus leads to the *invariance equation* [19, 20], which reads

$$-\,b\ \Phi[\hat{I}] = \Phi'[\hat{I}](b\ \Phi[\hat{I}]\ - \gamma\ ). \tag{10}$$

While Eq. (10) attains an exact, although not explicit, solution in terms of the Lambert W function [18], we wish to follow here another route and look for approximate, possibly explicit, solutions to Eq. (10). We can then integrate the latter by separation of variables, thus obtaining

$$\log \frac{\Phi[\hat{I}(t)]}{\Phi_0} = \frac{b}{\gamma}(\Phi[\hat{I}(t)] - \Phi_0]) + (\hat{I}(t) - \hat{I}_0)). \tag{11}$$

We note in passing that, by proceeding in the same manner with the observable $\hat{R}$, we obtain the corresponding invariance equation

$$\Xi'[\hat{I}(t)]\left(\frac{b}{\gamma}\ \Phi[\hat{I}(t)] - 1\right) = 1. \tag{12}$$

Using now (10), we can rewrite (12) in the form

$$\Xi'[\hat{I}] = -\frac{\gamma \Phi'[\hat{I}(t)]}{b \, \Phi[\hat{I}(t)]}, \tag{13}$$

which yields the expression

$$\log \frac{\Phi[\hat{I}(t)]}{\Phi_0} = -\frac{b}{\gamma} \left[\Xi[\hat{I}(t)] - \Xi_0 \right]. \tag{14}$$

Finally, letting

$$\Phi_0 = \hat{S}_0 \quad \text{and} \quad \Xi_0 = \hat{R}_0 \tag{15}$$

and using (11) and (14), we obtain the consistency relation

$$\hat{S(t)} + \hat{I(t)} + \hat{R(t)} = \hat{S}_0 + \hat{I}_0 + \hat{R}_0, \tag{16}$$

which shows that the conservation of the total number of individuals in the population is inherited by the reduced description.

Next, in order to determine an explicit approximate expression of the functional $\Phi[\hat{I}]$, we seek approximate solutions of Eq. (11) obtained via an iteration method. As a possible choice of the iteration method, we propose

$$\log \frac{\Phi^{(i+1)}[\hat{I}(t)]}{\Phi_0} = \frac{b}{\gamma} \left[\left(\hat{I}(t) - \hat{I}_0 \right) + \left(\Phi^{(i)}[\hat{I}(t)] - \Phi_0 \right) \right]. \tag{17}$$

For instance, by setting the initial condition for the recurrence equation (17) equal to $\Phi^{(0)}[\hat{I}(t)] = \hat{I}(t)$, we find, when setting $i = 0$, the solution

$$\Phi^{(1)}[\hat{I}(t)] = \Phi_0 \exp \left\{ \frac{b}{\gamma} (2\hat{I}(t) - \hat{I}_0 - \Phi_0) \right\}. \tag{18}$$

A natural question thus arises: what do we learn from approximate solutions? It will turn out that the better we can approximate the exact solution of Eq. (11), given in terms of the Lambert W function, the better we can reduce the SIR model; see Claim 3 later on.

3.1 A Direct Short-Time Estimate

We now aim at estimating the nearness of the solution $S(t)$ to Eq. (1) with initial datum $S(0) = S_0$, and the *constitutive law* $\Phi[\hat{I}(t)]$, defined in Eq. (4) with initial datum $\Phi[\hat{I}(0)] = \Phi_0$. It will turn out that our bound is meaningful only for a small time interval of observation and for a convenient parameter regime.

For an arbitrarily fixed value $\delta > 0$ with $t^* \in (0, \delta)$, we can derive for any $t \in (0, \delta)$ the next upper bound:

$$\begin{aligned}
|S(t) - \Phi[I(t)]| &= |S_0 + \int_0^t (-bSI)\, d\tau - \Phi[I(t)]| \\
&= |S_0 + \int_0^t (-bSI)\, d\tau - \Phi\left[I_0 + \int_0^t (bSI - \gamma I)\, d\tau\right]| \\
&\leq \mathcal{O}(\delta^2) + |S_0 - \Phi[I_0]| + |\left(-bS(t^*)I(t^*)\right)\delta| \\
&+ |\Phi'[I_0]\left(bS(t^*)I(t^*) - \gamma I(t^*)\right)\delta|. \\
&\leq |S_0 - \Phi[I_0]| + c^*\delta, \qquad (19)
\end{aligned}$$

where $c^* > 0$ is a constant depending on the parameters of the model, as well as on a priori uniform bounds on S, I and on the smoothness of Φ.

The last term comes from the Taylor expansion of $\Phi[I(t)]$, while the term involving the point evaluation in t^* is the result of the application of the mean-value theorem. Based on the rough estimate (19), we observe that the quantity

$$e(t) := |S(t) - \Phi[I(t)]|$$

can be made small if $\delta > 0$ is sufficiently small, Φ is at least twice differentiable, and $S(t)$ and $I(t)$ are bounded positive continuous functions. Note, however, that the smallness of $e(t)$ strongly depends on the choice of the parameters b and γ. For instance, in the limit of large values of γ, most likely the quantity $e(t)$ will grow. On the other hand, if γ takes moderate values, then we expect $e(t) \sim \mathcal{O}(\delta)$ for sufficiently small t and $e(t) \sim \mathcal{O}(\delta T)$ for $t \in (0, T)$.

Remark 1

(i) The structure of the original ODE system indicates that if the functionals $\Phi[\cdot]$ and $\Xi[\cdot]$ are suitable exponentials (obtained by integrating the equations for S and R), then $e(t) = 0$ for any $t \in (0, T)$.
(ii) We expect that instead of estimating from above the quantity $e(t)$, it is more practical to bound the quantity $|S(t) - \Phi[\hat{I}(t)]|$ for $t \in (0, \tau)$, with τ fixed.

3.2 An Indirect Large-Time Estimate

The dynamics of $I(t)$ is governed by

$$\frac{dI}{dt} = b\ S\ I - \gamma\ I, \qquad (20)$$

where $I(0) = I_0$. The starting point of this discussion is the fact that besides

$$\frac{d\hat{I}}{dt} = b\ \Phi[\hat{I}]\ \hat{I} - \gamma\ \hat{I}, \tag{21}$$

we may also consider

$$\frac{d\tilde{I}}{dt} = b\Phi^{(n)}[\tilde{I}]\tilde{I} - \gamma\tilde{I}, \tag{22}$$

where $\Phi^{(n)}$ is the solution to our iterative method at the step $n \in \mathbb{N}$. We provide also the information on the initial data $I(0)$, $\hat{I}(0)$, and $\tilde{I}(0)$.

Claim 1 The iteration method works such that for any $r \in [0, ||I||_\infty]$, it holds

$$|\Phi(r) - \Phi^{(n)}(r)| \le \epsilon_n, \tag{23}$$

with $\lim_{n\to\infty} \epsilon_n = 0$. Here, $||\cdot||_\infty$ denotes the standard uniform norm on $C[0, T]$.

The previous work done in the existing literature on the rigorous numerical approximation of the Lambert function gives trust in this claim; see, for instance, [18] and the references cited therein. Therefore, we omit to prove here this statement.

Claim 2 Under the assumptions for which *Claim 1* holds, there exist constants $\hat{c}_1 > 0$ and $\hat{c}_2 > 0$ such that

$$|I(t) - \tilde{I}(t)| \le e^{\hat{c}_1 T}\left(|I(0) - \tilde{I}(0)| + \hat{c}_2 T\epsilon_n\right) \tag{24}$$

holds for any $t \in (0, T)$. Here, $\hat{c}_1$ and $\hat{c}_2$ are independent of t and T.

Proof of Claim 2 We observe firstly that

$$|\Phi(I) - \Phi^{(n)}(\tilde{I})| = |\Phi(I) - \Phi(\tilde{I}) + \Phi(\tilde{I}) - \Phi^{(n)}(\tilde{I})| \le |\Phi(I) - \Phi(\tilde{I})|$$
$$+ |\Phi(\tilde{I}) - \Phi^{(n)}(\tilde{I})| \le$$
$$\le ||\Phi'||_\infty |I - \tilde{I}| + \epsilon_n.$$

Subtracting (22) from (21) gives

$$\frac{d}{dt}(I - \tilde{I}) = \gamma(\tilde{I} - I) + b\left[\Phi(I)I - \Phi^{(n)}(\tilde{I})\tilde{I}\right].$$

Noting that

$$\Phi(I)I - \Phi^{(n)}(\tilde{I})\tilde{I} = I\left(\Phi(I) - \Phi^{(n)}(\tilde{I})\right) + \Phi^{(n)}(\tilde{I})(I - \tilde{I})$$

leads to

$$\frac{d}{dt}|I-\tilde{I}| \leq (\gamma + |\Phi^{(n)}(\tilde{I})|)|\tilde{I}-I| + ||I||_\infty \Big(||\Phi'||_\infty |I-\tilde{I}| + \epsilon_n\Big) \leq$$

$$\leq \epsilon_n ||\Phi'||_\infty + (\gamma + |\Phi^{(n)}(\tilde{I})|(1+||I||_\infty))|\tilde{I}-I|.$$

Now, using Grönwall's inequality (cf., e.g., Appendix B in [17]) gives

$$|I(t)-\tilde{I}(t)| \leq e^{\int_0^t \left(\gamma + |\Phi^{(n)}(\tilde{I}(\tau))|(1+||I||_\infty)\right)d\tau} \Big[|I(0)-\tilde{I}(0)| + ||\Phi'||_\infty T\epsilon_n\Big].$$

Choosing now $\hat{c}_1 := \gamma + ||\Phi^{(n)}||_\infty|(1+||I||_\infty)$ and $\hat{c}_2 := ||\Phi'||_\infty$ leads to the desired estimate proposed by Claim 2.

3.3 Estimate on the Error of the Reduction Method

In this section, we aim to bound from above the error produced by the reduction method proposed within this framework.

Claim 3 Assume the hypothesis of Claim 2 to be true. Let $\tau > 0$ be arbitrarily fixed. Then, there exist strictly positive constants $\hat{c}_1$, $\hat{c}_2$, and $\hat{c}_3$ such that the following a priori estimate holds:

$$\int_0^\tau |S(s) - \Phi(\hat{I}(s))|ds \leq \hat{c}_1|I-\hat{I}| + \hat{c}_2\int_0^\tau |I(s)-\hat{I}(s)|ds + \hat{c}_3|I(0)-\hat{I}(0)|, \tag{25}$$

where I and $\hat{I}$ satisfy (2) and, respectively, (51).

Proof of Claim 3 We have

$$bS\hat{I} - b\Phi(\hat{I})\hat{I} + Sb(I-\hat{I}) = \frac{d}{dt}(I-\hat{I}) + \gamma(I-\hat{I}).$$

By a direct manipulation of the structure of equations (2) and (51), we obtain

$$S - \Phi(\hat{I}) = \frac{1}{b}\frac{1}{\hat{I}}\left[\frac{d}{dt}(I-\hat{I}) + (\gamma - Sb)(I-\hat{I})\right],$$

which by integration on $[0,\tau]$ leads to

$$\int_0^\tau |S(s) - \Phi(\hat{I}(s))|ds \leq \frac{1}{b}\int_0^\tau |\frac{1}{\hat{I}(s)}(\gamma - S(s)b)|ds$$

$$+\frac{1}{b}|\int_0^{\tau}\frac{1}{\hat{I}(s)}\frac{d}{dt}(I(s)-\hat{I}(s))ds|. \tag{26}$$

As the first term on the right-hand side of the last inequality can be bounded above by $(\gamma+||S||_{\infty}b)\frac{1}{\beta}\int_0^{\tau}|I(s)-\hat{I}(s)|ds$ (with $0<\beta\leq I(t)$) and, respectively, the last term by $\frac{2}{b}\left(|I(\tau)-\hat{I}(\tau)|+|I(0)-\hat{I}(0)|\right)$, the claim is now proven by choosing correspondingly the constants $\hat{c}_1$, $\hat{c}_3$, and $\hat{c}_3$.

Remark 2 Combining the statements of Claim 3 and Claim 2, we note that there exists a constant $c>0$ such that

$$\int_0^{\tau}|S(s)-\Phi(\hat{I}(s))|ds\leq c\left(\epsilon_n+|I(0)-\hat{I}(0)|\right). \tag{27}$$

This is obtained by adding and subtracting an $\tilde{I}$ in each term on the right-hand side of the estimate provided by Claim 3 and employing conveniently the statement of Claim 2. Note that the estimate (27) is quite practical. It basically tells that if $I(0)=\hat{I}(0)$, then the quality of the reduction method depends mostly on the quality of the numerical approximation of the constitutive law Φ^n.

The inequality (27) is an a priori bound on the error produced by the reduction strategy. This simply gives confidence that the reduction question makes sense in the SIR context.

4 Ideas for an Improved Reduced Description

The analysis of the SIR model, developed in Sect. 2, has shed light on the conditions under which one may hope to quantitatively capture the features of the original dynamics by using a reduced description based on the IM method. We shall now turn our attention to another model, amenable to an analytical solution, which will also clarify the strengths and limitations of the IM method.

4.1 An Instance of the Mori–Zwanzig Method

The Mori–Zwanzig method, in its essence, performs a partition of the dynamical variables into two subsets corresponding to the "relevant" and the "irrelevant" variables. Suitable projection operators are then employed to project the original dynamics onto the subspace of the relevant variables [32]. The simplest case in which this method can be discussed is a linear system of coupled first-order ODEs:

$$\dot{x}=L_{11}x+L_{12}y \tag{28}$$

$$\dot{y} = L_{21}x + L_{22}y \tag{29}$$

with $x(0) = x_0$ and $y(0) = y_0$, and where L_{ij}, $i, j = 1, 2$, are real parameters. Here, $x(t)$ and $y(t)$ are regarded, respectively, as the "relevant" and the "irrelevant" variables.

Using the set-up of Sect. 2, we may also regard $y(t)$ as the dynamical variable whose time evolution is driven by $x(t)$.

In this case, the original dynamics, given by Eqs. (28)–(29), is amenable to an analytical solution; namely, we can first solve (29) for $y(t)$:

$$y(t) = \exp\{L_{22}t\}y_0 + \int_0^t \exp\{L_{22}(t-s)\}L_{21}x(s)ds, \tag{30}$$

which represents the exact constitutive law linking $y(t)$ to $x(t)$. We can plug, next, (30) into (28) to obtain a closed ODE for $x(t)$, which reads

$$\dot{x} = L_{11}x + L_{12}\int_0^t \exp\{L_{22}(t-s)\}L_{21}x(s)ds + L_{12}\exp\{L_{22}t\}y_0\ . \tag{31}$$

Equation (31) represents the exact reduced description in terms of the relevant variable $x(t)$. The price we paid to get a reduced description, in this example, amounts to the presence of a memory term in the evolution equation for $x(t)$, which echoes the dynamics of the irrelevant variables.

We point out that, in the modelling of multiscale phenomena, the choice of the relevant dynamical variables is not always supported by guiding thermodynamic principles [22]. In fact, an improper choice of the relevant variables may not lead, eventually, to a successful reduced description [29]. On the other hand, a meaningful selection of the relevant variables proved to be extremely important, in statistical mechanics, to establish general results such as the Fluctuation–Dissipation Relations and the Fluctuation Relations in nonequilibrium systems [13, 14].

4.2 Using the Invariant Manifold Method

We shall now discuss the application of the IM method to the system (28)–(29). To get started, we focus on a perturbative method known as the Chapman–Enskog expansion, which stems from the geometrical theory of singular perturbations [25]; namely, we introduce a singular perturbation into the equation of $y(t)$ as follows:

$$\dot{y} = L_{21}x + \frac{1}{\varepsilon}L_{22}y, \tag{32}$$

where $\varepsilon > 0$ is a small parameter. Proceeding as in Sect. 2, we introduce the new variable $\hat{y}$ such that

$$\hat{y}(t) = \Phi[\hat{x}(t)], \tag{33}$$

where $\hat{x}(t)$ is defined via its ODE

$$\dot{\hat{x}} = L_{11}\,\hat{x} + L_{12}\,\Phi[\hat{x}]. \tag{34}$$

The dynamics of $\hat{y}(t)$ can be described equivalently using the following two equations:

$$\frac{d\hat{y}}{dt} = L_{21}\,x + \frac{1}{\varepsilon}L_{22}\,\Phi[\hat{x}] \tag{35}$$

$$\frac{d\hat{y}}{dt} = \frac{d\hat{x}}{dt}\,\Phi'[\hat{x}]. \tag{36}$$

We then impose the equality of two expressions of the time derivative of $\hat{y}(t)$, (35) and (36), thus obtaining the invariance equation

$$L_{21}\,\hat{x} + \frac{1}{\varepsilon}L_{22}\,\Phi[\hat{x}(t)] = \Phi'[\hat{x}]\,\left(L_{11}\,\hat{x} + L_{12}\,\Phi[\hat{x}]\right). \tag{37}$$

In the Chapman–Enskog method, the solution of Eq. (37) is sought by expanding the variable $\hat{y}$ in a form of a series in powers of the number ε, i.e.,

$$\Phi[\hat{x}] = \Phi^{(0)}[\hat{x}] + \sum_{i=1}^{\infty}\varepsilon^i\,\Phi^{(i)}[\hat{x}]. \tag{38}$$

Thus, Eq. (37) takes the form

$$\begin{aligned} &L_{21}\hat{x} + \frac{1}{\varepsilon}L_{22}(\Phi^{(0)} + \varepsilon\Phi^{(1)} + \varepsilon^2\Phi^{(2)} + \ldots) \\ &= \frac{d(\Phi^{(0)} + \varepsilon\Phi^{(1)} + \varepsilon^2\Phi^{(2)} + \ldots)}{d\hat{x}} \times \\ &\times[L_{11}\hat{x} + L_{12}(\Phi^{(0)} + \varepsilon\Phi^{(1)} + \varepsilon^2\Phi^{(2)} + \ldots)], \end{aligned} \tag{39}$$

which must be solved order by order by equating terms on both sides of the equation. At the lowest orders of ε, one finds the following sequence of constitutive laws:

$$\text{order}\varepsilon^{-1}: \qquad \Phi^{(0)}[\hat{x}] = 0 \tag{40}$$

$$\text{order}\varepsilon^{0}: \qquad \Phi^{(1)}[\hat{x}] = -\frac{L_{21}}{L_{22}}\hat{x} \tag{41}$$

$$\text{order}\varepsilon: \qquad \Phi^{(2)}[\hat{x}] = -\frac{L_{21}L_{11}}{L_{22}^2}\hat{x} \tag{42}$$

$$\text{order}\varepsilon^2: \qquad \Phi^{(3)}[\hat{x}] = \frac{L_{21}}{L_{22}^3}(L_{12}L_{21} - L_{11}^2)\hat{x}. \tag{43}$$

An inspection of the structure of the approximated solutions in Eqs. (40)–(43) gives us a hint on the structure of the solution of the invariance equation (37) when no singular perturbation is introduced in Eq. (32) (i.e., when setting $\varepsilon = 1$ in Eq. (32)) [20]. Therefore, we seek for a constitutive law (33) endowed with the linear structure

$$\Phi[\hat{x}] = A\ \hat{x}, \tag{44}$$

where $A(L_{11}, L_{12}, L_{21}, L_{22})$ is an unknown function of the parameters L_{ij}, $i, j = 1, 2$, yet to be determined. Hence, Eq. (37) takes the form

$$L_{21}\ \hat{x} + L_{22}\ A\ \hat{x} = A\ (L_{11}\ \hat{x} + L_{12}\ A\ \hat{x}). \tag{45}$$

Equation (45) becomes the quadratic equation

$$L_{12}\ A^2 + (L_{11} - L_{22})\ A - L_{21} = 0, \tag{46}$$

whose roots are

$$A^* = \frac{-(L_{11} - L_{22}) \pm \sqrt{(L_{11} - L_{22})^2 + 4L_{12}L_{21}}}{2L_{12}}. \tag{47}$$

Thus, the IM method leads to the following constitutive law:

$$\hat{y} = A^*\hat{x}, \tag{48}$$

which should be compared with the exact constitutive law (30).

Using (48), we can now rewrite (34) as follows:

$$\dot{\hat{x}} = \mathcal{A}^*\hat{x}, \tag{49}$$

where we have set $\mathcal{A}^* = L_{11} + L_{12}A^*$.

Figure 1 shows the behavior of $x(t)$, obtained by integration of Eq. (31), and of the real part $\Re(\hat{x}(t))$ of the solution of the reduced system (49), with an initial datum $x_0 = \hat{x}_0 = 1$, for two different values of the coupling parameter L_{21}. In particular, Fig. 1 evidences that fixing a larger value of L_{21}, while keeping the other parameters fixed, leads to a better performance of the IM reduction method. The reason is that a larger value of L_{21}, in (29), makes the time derivative of $y(t)$ more strongly affected by the behavior of $x(t)$. This, hence, fits nicely with the ansatz (33), which requires the dynamics of $\hat{y}(t)$ to be driven by $\hat{x}(t)$.

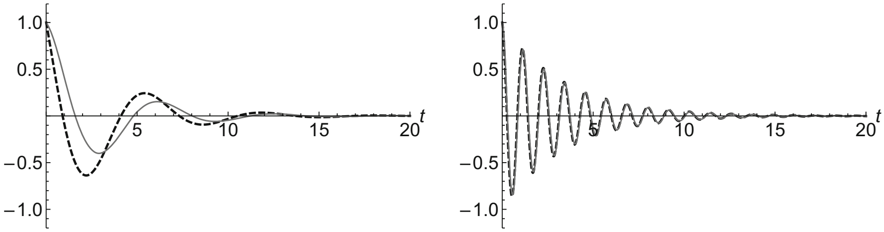


Fig. 1 *Left panel:* behavior of $x(t)$ (black dashed line), obtained from Eq. (31), and $\Re(\hat{x}(t))$ (gray solid line), obtained from Eq. (49), with $x_0 = \hat{x}_0 = 1$, $L_{11} = -0.1$, $L_{12} = -1$, and $L_{22} = -0.5$ and with $L_{21} = 1$ (left panel) and $L_{21} = 30$ (right panel)

4.3 Back to the SIR Model

In this concluding section, we return to the SIR model of Sect. 2 and discuss the application of the IM method by using the ansatz (5). It will turn out that, in this case, the IM method may yield an exact reduced description.

The dynamics of the field $\hat{S}(t)$ now reads

$$\frac{d\hat{S}}{dt} = -b\ \hat{S}\ \Psi[\hat{S}(t)] \tag{50}$$

with $\hat{S}(0) = \hat{S}_0$. We also fix $\Psi[\hat{S}(0)] = \Psi_0$ and $\Omega[\hat{S}(0)] = \Omega_0$.

In the present case, Eq. (50) corresponds to the reduced description of the original SIR model, Eqs. (1)–(3). To find an explicit expression for the constitutive law $\Psi[\hat{S}]$, we write, first, the dynamics of the driven observable $\hat{I}(t)$ as

$$\frac{d\hat{I}}{dt} = b\ \hat{S}\ \Psi[\hat{S}(t)] - \gamma\ \Psi[\hat{S}(t)]. \tag{51}$$

Next, as in Sect. 3, we also write the time derivative of $\hat{I}(t)$ by using the chain rule, i.e.,

$$\frac{d\hat{I}}{dt} = \frac{d\hat{S}}{dt}\ \Psi'[\hat{S}]. \tag{52}$$

We thus obtain the invariance equation:

$$\Psi'[\hat{S}(t)] = -1 + \frac{\gamma}{b\ \hat{S}(t)}. \tag{53}$$

We can then integrate Eq. (53) by separation of variables, thus obtaining

$$\Psi[\hat{S}(t)] = \Psi_0 - (\hat{S}(t) - \hat{S}_0) + \frac{\gamma}{b} \log \frac{\hat{S}(t)}{\hat{S}_0}. \tag{54}$$

In a similar fashion, we find

$$\Omega[\hat{S}(t)] = \Omega_0 - \frac{\gamma}{b} \log \frac{\hat{S}(t)}{\hat{S}_0}. \tag{55}$$

Finally, by setting

$$\Psi_0 = \hat{I}_0 \quad \text{and} \quad \Omega_0 = \hat{R}_0 \tag{56}$$

and by summing up (54) and (55), we obtain

$$\hat{S}(t) + \hat{I}(t) + \hat{R}(t) = \hat{S}_0 + \hat{I}_0 + \hat{R}_0, \tag{57}$$

which yields, again, the conservation of the total number of individuals in the reduced dynamics. In Fig. 2, the behavior of $S(t)$, obtained by integration of the original SIR model, Eq. (1)–(3), is compared with the solution $\hat{S}(t)$ of Eq. (50), equipped with the constitutive law (54). Figure 2 shows that the behavior of $\hat{S}(t)$ recovers with striking accuracy that of $S(t)$. Moreover, the two panels of Fig. 3 show the parametric plots of the $I(t)$ vs. $S(t)$ (left panel) and $R(t)$ vs. $S(t)$ (right panel) for the original SIR model and for the reduced description.

We would like, however, to point out that the nice agreement between original and reduced dynamics outlined above might easily be lost when considering time-dependent parameters $b = b(t)$ and $\gamma = \gamma(t)$, for $t > 0$. A careful rewriting of the invariance equation is demanded to handle such a case. We will discuss this scenario elsewhere.

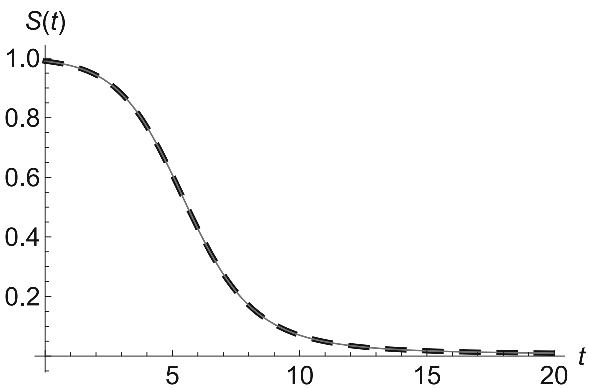


Fig. 2 Behavior of $S(t)$ for the original dynamics, Eq. (1) (black dashed line), and for the reduced description (gray solid line), obtained from Eqs. (50) and (54). We fixed $S_0 = \hat{S}_0 = 0.99$ and $\Psi_0 = \hat{I}_0 = 0.01$, with $b = 1$ and $\gamma = 0.2$

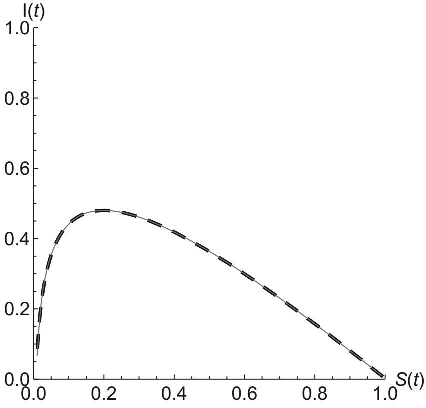


Fig. 3 *Left panel:* parametric plots of $I(t)$ vs $S(t)$ for the original dynamics, Eqs. (1)–(2) (black dashed line) and for the reduced description (gray solid line), obtained from Eq. (54). *Right panel:* parametric plot of $R(t)$ vs $S(t)$ for the original dynamics (black dashed line) and for the reduced description (gray solid line), obtained from Eq. (55). We fixed $I_0 = \hat{I}_0 = \Psi_0 = 0.01$ and $R_0 = \hat{R}_0 = \Omega_0 = 0$

5 Conclusion

We have succeeded to identify constitutive laws to reduce the presence of either the fraction of the susceptible or the infected individuals in the standard SIR model. The reduced descriptions, obtained using the IM method, agree via numerical simulations and practical a priori error bounds with what is expected from the original SIR dynamics.

Our work opens the possibility to use the reduced SIR dynamics for reading off data available, for instance, on demonstrated COVID-19 infections and deaths and, based on a parameter identification approach done at this level, produce a new forecast on the effects of the pandemic evolution.

From a long-term research perspective, the method discussed in these notes indicates new routes to be exploited to obtain reduced descriptions in yet uncharted, or only partially explored, territories, such as the mathematical modelling of crowd dynamics [5–7, 15], electronic transport [3], and uphill diffusions [4, 9], in the framework of interacting particle systems.

References

1. L. Almeida, P.-A. Bliman, G. Nandin, B. Perthame, N. Vauchelet, *Final size and convergence rate for an epidemic in heterogeneous populations*. M3AS, 2021.
2. T. Britton, F. Ball, P. Trapman, *A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2*. Science **369**, 6505: 846–849 (2020).

3. E.N.M. Cirillo, M. Colangeli, L. Rondoni, *Transport in quantum multi-barrier systems as random walks on a lattice*, J. Stat. Phys. **176**(3), 692–709 (2019).
4. E.N.M. Cirillo, M. Colangeli, *Stationary uphill currents in locally perturbed zero-range processes*, Phys. Rev. E **96**(5) 052137 (2017).
5. E.N.M. Cirillo, M. Colangeli, A. Muntean, *Effects of communication efficiency and exit capacity on fundamental diagrams for pedestrian motion in an obscure tunnel – a particle system approach*, Multiscale Model. Sim. **14**(2), 906–922 (2016).
6. E.N.M. Cirillo, M. Colangeli, A. Muntean, T.K.T. Thieu, *A lattice model for active - Passive pedestrian dynamics: A quest for drafting effects*, Math. Biosci. Eng., **17**(1), 460–477 (2020).
7. E.N.M. Cirillo, M. Colangeli, A. Muntean, T.K.T. Thieu, *When diffusion faces drift: Consequences of exclusion processes for bi-directional pedestrian flows*, Physica D **413**, 132651 (2020).
8. M. Choisy, J.-F. Guégan, P. Rohani, *Mathematical modeling of infectious diseases dynamics*. In *Encyclopedia of Infectious Diseases*, M. Tibayrenc (Ed.), 379–404, (2007).
9. M. Colangeli, C. Giardinà, C. Giberti, C. Vernia, *Nonequilibrium two-dimensional Ising model with stationary uphill diffusion*, Phys. Rev. E **97**(3) 030103 (2018).
10. M. Colangeli, I. V. Karlin, M. Kröger, *From hyperbolic regularization to exact hydrodynamics for linearized Grad's equations*, Phys. Rev. E **75**, 051204 (2007).
11. M. Colangeli, I. V. Karlin, M. Kröger, *Hyperbolicity of exact hydrodynamics for three-dimensional linearized Grad's equations*, Phys. Rev. E **76**, 022201 (2007).
12. M. Colangeli, M. Kröger, H. C. Öttinger, *Boltzmann equation and hydrodynamic fluctuations*, Phys. Rev. E **80**, 051202 (2009).
13. M. Colangeli, L. Rondoni, A. Vulpiani, *Fluctuation-dissipation relation for chaotic non-Hamiltonian systems*, J. Stat. Mech. L04002 (2012).
14. M. Colangeli, L. Rondoni, *Equilibrium, fluctuation relations and transport for irreversible deterministic dynamics*, Physica D **241**, 681–691 (2012).
15. M. Colangeli, A. Muntean, O. Richardson, T.K.T. Thieu, *Modelling interactions between active and passive agents moving through heterogeneous environments*, Modeling and Simulation in Science, Engineering and Technology, 211–257, Springer Basel (2018).
16. D. T. Crommelin, A. J. Majda, *Strategies for model reduction: Comparing different optimal bases*, Journal of the Atmospheric Sciences **61**, 2206–2217 (2004).
17. L. C. Evans, *Partial Differential Equations*, 2nd Edition, Graduate Studies in Mathematics, vol. **19**, American Mathematical Society (2010).
18. M. Fasi, N.J. Higham and B. Iannazzo, *An algorithm for the matrix Lambert W function*, SIAM J. Matrix Anal. Appl., **36**, No. 2, 669–685 (2015).
19. A. N. Gorban, I. V. Karlin, *Invariant Manifolds for Physical and Chemical Kinetics*, Lect. Notes Phys. **660**, Springer, Berlin (2005).
20. A. N. Gorban, I. V. Karlin, *Hilbert's* 6*th problem: Exact and approximate manifolds for kinetic equations*, Bulletin of the American Mathematical Society, Volume **51** (2), 187–246 (2013).
21. A. N. Gorban, I. G. Kevrekidis, C. Theodoropoulos, N. K. Kazantzis, H. C. Öttinger, *Model Reduction and Coarse-Graining Approaches for Multiscale Phenomena*, Springer, Berlin (2006).
22. M. Grmela, *Role of thermodynamics in multiscale physics*, Comp. Math. App. **65**, no. 10, 1457–1470 (2013).
23. H. Jardón-Kojakhmetov, C. Kuehn, A. Pugliese, M. Sensi, *A geometric analysis of the SIRS epidemiological model on a homogeneous network*, Nonlinear Analysis RWA **58**, 103220, (2021).
24. H. Jardón-Kojakhmetov, C. Kuehn, A. Pugliese, M. Sensi, *A geometric analysis of the SIR, SIRS and SIRWS epidemiological models*, preprint. arXiv:2002.00354v1 (2020).
25. C. K. R. T. Jones, *Geometric Singular Perturbation Theory*, Dynamical systems (Montecatini Terme, 1994), Lecture Notes in Math., vol. 1609, Springer, Berlin, 1995.
26. L.P. Kadanoff, *Scaling laws for Ising models near* T_c, Physics **2**, 263–272 (1966).
27. I. V. Karlin, M. Colangeli, M. Kröger, *Exact linear hydrodynamics from the Boltzmann equation*, Phys. Rev. Lett. **100**, 214503 (2008).

28. W. R. KhudaBukhsh, S. K. Khalsa, E. Kenah, G. A. Rempala, J. H. Tien, *COVID-19 dynamics in an Ohio prison*, medRxiv 2021.01.14.21249782.
29. H. C. Öttinger, *Beyond Equilibrium Thermodynamics*, John Wiley & Sons (2005).
30. H. C. Öttinger, *On the stupendous beauty of closure*, Journal of Rheology **53**, 1285 (2009).
31. F. J. Vermolen, *A spatial Markov chain cellular automata model for the spread of viruses*. preprint. arXiv:2004.05635 (2020)
32. R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford University Press (2001).
33. K.G. Wilson, *Renormalization group and critical phenomena I. Renormalization group and the Kadanoff scaling picture*, Phys. Rev. B **4**, 3174–3183 (1971).
34. M. Wolff, *On build-up of epidemiologic models–Development of a SEI^3RSD model for the spread of SARS-CoV-2*, Zeitschrift für angewandte Mathematik und Mechanik (ZAMM) **100** (11), 1–36 (2020).
35. M. Wolff, *Mathematische Bemerkungen zu einem epidemiologischen Modell im Zusammenhang mit der Corona-Pandemie*, preprint, May 2020. https://doi.org/10.13140/RG.2.2.20564.35209.

An Agent-Based Model of COVID-19 Diffusion to Plan and Evaluate Intervention Policies

Gianpiero Pescarmona, Pietro Terna, Alberto Acquadro, Paolo Pescarmona, Giuseppe Russo, Emilio Sulis, and Stefano Terna

Abstract A model of interacting agents, following plausible behavioral rules into a world where the Covid-19 epidemic is affecting the actions of everyone. The model works with (i) infected agents categorized as symptomatic or asymptomatic and (ii) the places of contagion specified in a detailed way. The infection transmission is related to three factors: the characteristics of both the infected person and the susceptible one, plus those of the space in which contact occurs. The model includes the structural data of Piedmont, an Italian region, but we can easily calibrate it for other areas. The micro-based structure of the model allows factual, counterfactual, and conditional simulations to investigate both the spontaneous or controlled development of the epidemic.

The model is generative of complex epidemic dynamics emerging from the consequences of agents' actions and interactions, with high variability in outcomes and stunning realistic reproduction of the successive contagion waves in the reference region. There is also an inverse generative side of the model, coming

G. Pescarmona · A. Acquadro · E. Sulis
University of Torino, Torino, Italy
e-mail: gianpiero.pescarmona@unito.it; alberto.acquadro@unito.it; emilio.sulis@unito.it

P. Terna (✉)
University of Torino, Torino, Italy

Fondazione Collegio Carlo Alberto, Turin, Italy
e-mail: pietro.terna@unito.it

P. Pescarmona
University of Groningen, Groningen, The Netherlands
e-mail: p.p.pescarmona@rug.nl

G. Russo
Centro Einaudi, Torino, Italy
e-mail: russo@centroeinaudi.it

S. Terna
TomorrowData, Turin, Italy
e-mail: stefano.terna@tomorrowdata.io
https://tomorrowdata.io

N. Bellomo, L. Gibelli (eds.), *Crowd Dynamics, Volume 3*, Modeling and Simulation in Science, Engineering and Technology, https://doi.org/10.1007/978-3-030-91646-6_9

from the idea of using genetic algorithms to construct a meta-agent to optimize the vaccine distribution. This agent takes into account groups' characteristics—by age, fragility, work conditions—to minimize the number of symptomatic people.

1 A Quick Introduction to Our Agent-Based Epidemic Model

The starting point is a compartmental model with Susceptible, Infected, and Recovered people (S.I.R.), but adding both a more detailed breakdown of the subjects involved in the contagion process [1] and a multi-scale framework to account for the interaction at different dimensional, and spatial levels [2]. From the virus micro-level, we move to individuals and up to the collective behavior of the population.

Following [3], we know that the analysis based on the assumption of heterogeneity strongly differs from S.I.R. compartmental structures modeled by differential equations. The authors of this work argue when it is best to use agent-based models and when it would be better to use differential equation models ponder when it is better to use agent-based models and when it would be better to use differential equation models. Differential equation models assume homogeneity and perfect mixing of characteristics within compartments, while agent-based models can capture heterogeneity in agent attributes and the structure of their interactions. We follow the second approach (about agent-based approach, see Sect. 1.1).

- Our model takes into consideration:
  - (i) infected agents categorized as symptomatic or asymptomatic and
  - (ii) the places of contagion specified in a detailed way, thanks to agent-based modeling capabilities.
- The infection transmission is related to three factors: the infected person's characteristics and those of the susceptible one, plus those of the space in which a contact occurs.

Finally, we subscribe the call of [4] to «cover the full behavioural and social complexity of societies under pandemic crisis» and we work arguing that «the study of collective behavior must rise to a "crisis discipline" just as medicine, conservation, and climate science have, with a focus on providing actionable insight to policymakers and regulators for the stewardship of social systems», as in [5].

A look at the structure of the whole presentation. In Sect. 1.1, we discuss models and specifically agent-based models; in Sect. 1.2, the molecular support to agents' intrinsic susceptibility construction; in Sect. 1.3, the structure of the model, with the daily sequence of the agents' actions. Section 2 introduces a detailed description of the internal model mechanisms, with: conditional actions in Sect. 2.1, parameters in Sect. 2.2 and agents' interaction in Sect. 2.3.

A technique for contagion representation is introduced in Sect. 3. Then we explore simulation cases in Sect. 4, building several batches of runs and comparing extreme situations in Sect. 4.1.

Section 4.2 reports the actual epidemic data in the reference region. With those data, we verify factual and counterfactual analyses in Sect. 5. Considering the possibility of calculating infection indicators without delays (Sect. 5.4), we experiment with the effect of adopting the control measure with 20 days of anticipation (Sect. 5.5). In Sect. 5.6 we verify another counterfactual policy, that of concentrating the efforts uniquely in defense of fragile persons. Section recap 5.7 summarizes these results.

The final application of the model is dedicated to a planning exercise on vaccination campaigns (Sect. 7). We introduce an analysis of the vaccine mechanism in the perspective of our model (Sect. 7.1), using both planned strategies (Sects. 7.4 and 7.5) and genetic algorithms (Sect. 7.6). The GAs goal is to optimize the behavior of a meta-agent, deciding the sequence of the vaccinations.

1.1 Why Models? Why Agents? Why Another Model?

Why another model, and most of all, why models? With [6]:

> The choice (...) is not whether to build models; it's whether to build explicit ones. In explicit models, assumptions are laid out in detail, so we can study exactly what they entail. On these assumptions, this sort of thing happens. When you alter the assumptions that is what happens. By writing explicit models, you let others replicate your results.

With even more strength:

> I am always amused when these same people challenge me with the question,"Can you validate your model?" The appropriate retort, of course, is,"Can you validate yours?" At least I can write mine down so that it can, in principle, be calibrated to data, if that is what you mean by "validate" a term I assiduously avoid (good Popperian that I am).

To reply to "why agents?", with [7] we define in short what an agent-based model is:

> An agent-based model consists of individual agents, commonly implemented in software as objects. Agent objects have states and rules of behavior. Running such a model simply amounts to instantiating an agent population, letting the agents interact, and monitoring what happens. That is, executing the model—spinning it forward in time—is all that is necessary in order to "solve" it.

More in detail:

> There are, ostensibly, several advantages of agent-based computational modeling over conventional mathematical theorizing. First, [...] it is easy to limit agent rationality in agent-based computational models. Second, even if one wishes to use completely rational agents, it is a trivial matter to make agents heterogeneous in agent-based models. One simply instantiates a population having some distribution of initial states, e.g., preferences. That is, there is no need to appeal to representative agents. [...] Finally, in most social processes either physical space or social networks matter. These are difficult to account for

> mathematically except in highly stylized ways. However, in agent-based models it is usually quite easy to have the agent interactions mediated by space or networks or both.

In [8] we have a relevant step ahead, considering *inverse generative social science*:

> The agent-based model (ABM) is the principal scientific instrument for understanding how individual behaviors and interactions, the micro-world, generates change and stasis in macroscopic social regularities. So far, agents have been iterated forward to generate such explananda as settlement patterns, scaling laws, epidemic dynamics, and many other phenomena [6]. But these are all examples of the forward problem: we design agents and grow the target phenomenon. The motto of generative social science is: "If you didn't grow it, you didn't explain it." [9] But there may be many ways to grow it! How do we find 'all' the non-trivial generators? This is inverse generative social science—agent architectures as model outputs not model inputs—and machine learning can enable it.

And now, "why another?" As a commitment to our creativity, using our knowledge to understand what is happening. Indeed, with arbitrariness: it is up to others and time to judge.

As any model, also this one is based on assumptions: time will tell whether these were reasonable hypotheses. Modeling the Covid-19 pandemic requires a scenario and the actors. As in a theater play, the author defines the roles of the actors and the environment. The characters are not real, they are prebuilt by the author, and they act according to their peculiar constraints. If the play is successful, it will run for a long time, even centuries. If not, we will rapidly forget it. Shakespeare's Hamlet is still playing after centuries, even if the characters and the plot are entirely imaginary. The same holds for our simulations: we are the authors, we arbitrarily define the characters, we force them to act again and again in different scenarios. However, in our model, the micro-micro assumptions are not arbitrary but based on scientific hypotheses at the molecular level, the micro agents' behaviors are modeled in an explicit and realistic way. In both plays and simulations, we compress the time: a whole life to two or three hours on the stage. In a few seconds, we run the Covid-19 pandemic spread in a given regional area.

1.2 The Molecular Basis of SARS-CoV-2 Infection

To fully understand what the word infection means, we have previously to define the scenario where life takes place.

We start with the properties of life on earth's surface [10]. Making a long story short, basically life is a dissipative process fueled by energy supplied by the sun. As the sun has been shining for billions of years the biological systems on earth expanded exponentially in a finite environment, and they became limited in their growth due to the shortage of available atoms/molecules ("nutrients"). The competition for the limiting nutrients in each local environment ("niche") will locally drive the selection and will explain the complexity of the interactions of the different organisms in any environment, from the microscopic to the social level.

The ground, the soil, and the ponds are overcrowded with bacteria, algae, molds, insects, and so on. They help to keep the ground healthy and ready for cultivation. We too are populated by microorganisms, the gut, the mucosae, the skin. But when we feel healthy we do not realize they are there; but sometimes we do not feel well, we are sick. We have a disease, we need a culprit: somebody different from us, a virus, a bacterium, a protozoan that infected us.

In most cases, the same agent is shared by people surrounding us, but most of them are healthy, few are sick. The coexistence/cooperation between organisms sharing the same "niche" is the rule after billions of years of evolution. The disease is the exception. The asymptomatic infection is the rule, the symptomatic infection is the disease. The simplest explanation for the rise of the symptoms is the competition of different organisms for a limiting "nutrient."

In the case of the Herpesvirus family and man, the limiting "nutrient" is Iron. Virus Ribonucleotide reductase is an enzyme with an affinity for iron higher than human cells. Infected cells survive quite well until iron availability covers the needs of both host and virus. In the case of iron shortage (evaluated as the level of serum ferritin) the infected cells are forced to reduce heme synthesis, necessary for the respiratory chain, and hence ATP synthesis. Less ATP, loss of many cellular functions, symptoms. In our experience, most of the people seropositive to HSV had no symptoms, provided they had serum ferritin levels $\geq$90 ug/dl. The lower the ferritin level, the higher the frequency of the symptoms. The level of ferritin depends on genetic, dietary, environmental factors, explaining the variability of the clinical manifestations [11].

In the case of HSV, the virus metabolism is well known and studied for tens of years. In the case of SARS-CoV-2 our experience is in the range of months and the identification of the limiting "nutrient" is only speculative.

On the basis of the data collected up to now, Cysteine could be the most relevant. One of the co-authors here, G. Pescarmona, with other contributors, has recently developed a software able to easily compare the amino acids (AA) percentage and some selected ratios between couples of them, using Uniprot proteins repository as data source [12]. Using this software, it has been possible to compare the AA percentage in different tissues [13] demonstrating the limiting role of AA local availability on the synthesis of specific proteins.

From the beginning of the pandemic, it has been clear that ACE2 was the preferred ligand for the Spyke protein and that cells expressing it were the perfect host for the fast synthesis of viral protein [14]. Our working hypothesis is that the best host cell is one producing a protein with a similar AA percentage. From Fig. 1 we can extract the following info: most AA percentages of the viral proteins are similar, with the exception of Cysteine (lower) and Methionine and Tryptophan (higher) to the ACE2. Higher methionine associates with faster protein synthesis, higher Tryptophan with higher nucleic acid synthesis. A perfect environment for replication of both RNA and virus proteins. Expression of viral proteins with high Cysteine decreases free cysteine and therefore Glutathione (GSH) synthesis, with impaired antioxidant defense and increased ROS activity. Increased ROS activity has been one of the first well-identified mechanisms of viral infection [15] and their

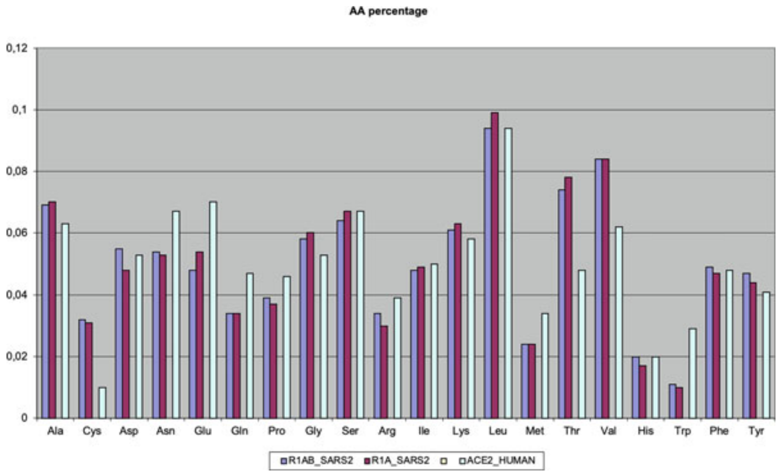


Fig. 1 Comparison of the Amino Acids percentage in the most representative SARS-Cov-2 proteins and the human ACE2, receptor on the surface of host cells

scavenging by GSH has been proposed as a preventive/therapeutic approach to the symptomatic disease [16].

Also, the ratio between AA couples, in Fig. 2, shows good similarity with the exception of the Spyke protein, as far as the ratios including glutamate are involved, but the almost perfect coincidence between the catalytic proteins of the virus and ACE2 explains the reproductive advantage of entering a cell expressing it. The interesting information that we get from this approach is that the cysteine deprivation of the naturally infected cells is shared also by cells induced to produce Spyke protein, independently by the vector used. Moreover, whilst the full virus enters the cells expressing on the outer surface ACE2, and we can identify them and try to imagine the long-term effects of infection, in the case of vaccines the synthetic vectors should allow the entry in any kind of cells.

In conclusion, we can expect oxidative damage (ROS increase and inflammation) in any kind of cell in our body. The extent of the inflammation will vary according to so many variables: age, diet, drugs, previous silent sites of inflammation, to make almost impossible the prediction of the localization and gravity of the side effects.

The Cytokine Storm

The cytokine storm is a synthetic definition of the set of reactions leading to the symptomatic COVID-19 and to death. The core process of the infection is the unbalance between ACE/ACE2. SARS-CoV-2 binds to ACE-2 and sequester it, causing an ACE prevalence and a sharp increase of ROS [17, 18]. All pre-existing processes leading to the prevalence of ACE are pro-inflammatory, those leading to a prevalence of ACE2, are anti-inflammatory. A low level of the active Vitamin D (1,25-dihydroxy-Vitamin D) leads to an increased expression of ACE. Cortisol has the same, but with an independent mechanism, effect on ACE expression and,

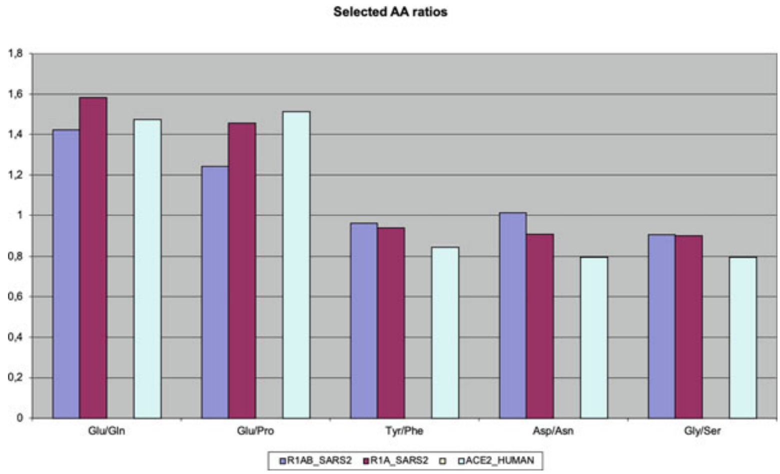


Fig. 2 Comparison of some ratios between selected Amino Acids in the most representative SARS-Cov-2 proteins and the human ACE2, receptor on the surface of host cells. These ratios supply specific information about the local metabolic condition inside the cell [13]

additionally, decreases the expression of ACE2. In all the cases the ROS released by ACE activity are scavenged by GSH and its ancillary enzymes. Downstream of ROS, the inflammatory pathway includes NF-kB, TNF-alpha, IL-6, PLA2, COX1, and COX2.

From the clinical point of view the COVID-19 pandemic is affecting differently the world population: in presence of conditions such as aging, diabetes, obesity, and hypertension the virus triggers a lethal cytokine storm and patients die from acute respiratory distress syndrome, whereas in many cases the disease has a mild or even asymptomatic progression [19]. The identification of the biochemical patterns underlying the severe disease may allow the identification of fragile people in need of more accurate protection.

Combining the biochemical determinants listed in Table 1 within the model described in Fig. 3 is possible to evaluate the risk for every individual, or class of similar individuals, of developing a severe form of the disease.

DHEA is an adrenal hormone, a precursor of testosterone and estrogens, that activates heme synthesis. Heme is required for plenty of reactions, including the respiratory chain (ATP synthesis) and Vitamin D activation. ATP is required for GSH synthesis, BMR reflects the activity of the respiratory chain and therefore depends again on heme. Heme synthesis requires iron, whose availability depends on diet, correct digestion, and absorption. Table 1 lists also some of the environmental factors that can interfere with the molecules involved in the COVID-19 dependent inflammatory response. Environmental pollution and drugs abuse in older people are among the factors that can explain the excess mortality in developed countries. The

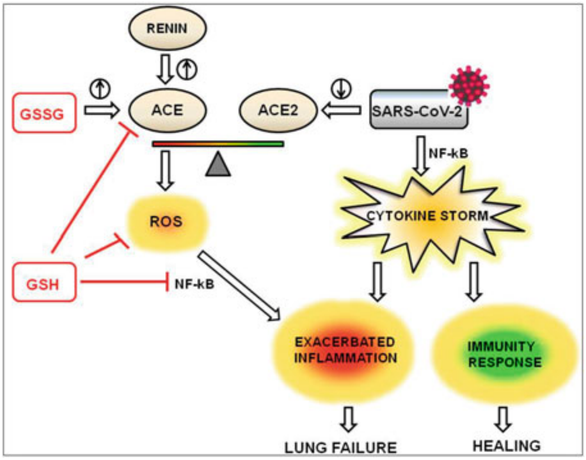


Fig. 3 All the main agents involved in the inflammatory response during COVID-19 are depicted here, with their relationships (reprinted with permission from [16])

Table 1 A synopsis of all the metabolic features associated with the clinical conditions favoring a severe COVID-19 development. DHEA: Dehydroepiandrosterone, GSH: Glutathione, Vit D: 25(OH)-Vitamin D, BMR: Basal Metabolic Rate

Risk factors	DHEA	Cortisol	GSH	Vit D	BMR
Aging	Low	High	Low	Low	Low
Diabetes	Low	High	Low	Low	Low
Hypertension	Low	High	Low	Low	?
Obesity	Low	High	Low	Low	Low
Diuretics	–	High	–	–	?
Drugs	–	–	Low	Low	?
Air pollution	–	–	Low	–	?
Paracetamol	–	–	Low	–	?
Chloroquine	–	–	Low	–	?
Glucocorticoids	–	High	–	–	?
Ibuprofen	–	–	–	–	?
Aspirin					

therapeutic use of paracetamol, chloroquine, and glucocorticoids to prevent severe symptoms looks inappropriate on the basis of their action mechanism.

This set of considerations can be used to tentatively identify and protect fragile people but can be easily modified according to the epidemiological data. Unfortunately, up to now, the prevailing approach has been different, and not so much data about the characteristics of the patients severely ill have been published to allow validation of our criteria for fragility.

1.3 Our Model

With our model, we move from a macro compartmental vision to a meso and micro-analysis capability. Its main characteristics are:

- scalability: we take into account the interactions between virus and molecules inside the host, determining individual susceptibility; the interactions between individuals in more or less restricted contexts; the movement between different environments (home, school, workplace, open spaces, shops); the movements occur in different parts of the daily life, as in [20]; in detail, the scales are:
  - *micro*, with the internal biochemical mechanism involved in reacting to the virus, as in [16], from where we derive the critical importance assigned to an individual attribute of intrinsic susceptibility related to the age and previous morbidity episodes; the model indeed incorporates the medical insights and consistent perspectives of one of its co-authors, former full professor of clinical biochemistry, signing also the quoted article; a comment on Lancet [21] consistently signals the syndemic character of the current event: «Two categories of disease are interacting within specific populations—infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and an array of non-communicable diseases (NCDs)»;
  - *meso*, with the open and closed contexts where the agents behave, as reported above;
  - *macro*, with the emergent effects of the actions of the agents;
- granularity: at any level, the interactions are partially random and therefore the final results will always reflect the sum of the randomness at the different levels; changing the constraints at different levels and running multiple simulations should allow the identification of the most critical points, where to focus the intervention.

Summing up, S.I.s.a.R. (https://terna.to.it/simul/SIsaR.html) is an agent-based model designed to reproduce the diffusion of the COVID-19 using agent-based modeling in NetLogo [22]. We have Susceptible, Infected, symptomatic, asymptomatic, and Recovered people: hence the name S.I.s.a.R. The model works on the structural data of Piedmont, an Italian region, but we can quite easily calibrate it for other areas. It reproduces the events following a realistic calendar (national or local government decisions, as in Sect. 2.2), via its script interpreter. At the above address, it is also possible to run the code online without installation. Into the *Info* sheet of the model, we have more than 20 pages of Supporting Information about both the structure and the calibration of the model.

The micro-based structure of the model allows factual, counterfactual, and conditional simulations. Examples of counterfactual situations are those considering:

(i) different timing in the adoption of the non-pharmaceutical containment measures;
(ii) an alternative strategy, focusing exclusively on the defense of fragile people.

The model generates complex epidemic dynamics, emerging from the consequences of agents' actions and interactions, with high variability in outcomes, and with a stunning realistic reproduction of the contagion waves that occurred in the reference region.

We take charge of the variability of the epidemic paths within the simulation, running batches of executions with 10,000 occurrences for each experiment.

Following [8], the AI and inverse generative side of the model comes from constructing a meta-agent optimizing the vaccine distribution among people groups—characterized by age, fragility, work conditions—to minimize the number of symptomatic people (as deceased persons come from there).

We can characterize the action of the planner both:

(i) introducing ex-ante rules following "plain" or "wise" strategies that we imagine as observers or
(ii) evolving those strategies via the application of a genetic algorithm, where the genome is a matrix of vaccination quotas by people groups, with their time range of adoption.

2 How S.I.s.a.R. Works

We have two initial infected individuals in a population of 4350 individuals, on a scale of 1:1000 with Piedmont. The size of the initial infected group is out of scale: it is the smallest number ensuring the epidemic's activation in a substantial number of cases. Initial infected people bypass the incubation period. For plausibility reasons, we never choose initial infected people among persons in nursing homes or hospitals. The presence of agents in close spaces—such as classrooms, factories, homes, hospitals, nursing homes—is set with realistic numbers, out of scale: e.g., a classroom contains 25 students, a home two persons, large factories up to 150 employees, small ones up to 15, etc.; the movements occur in different parts of the daily life, as in [20].

In Fig. 4 we have a 3D representation of the model world, with one of the possible random maps that the simulation generates. Persons are in gray, houses in cyan, nursing homes in orange, hospitals in pink, schools in yellow, factories (with shops and offices) in brown. Persons have a cylinder as shape, if regular or robust (young); a capital X if fragile; temporary their colors can be: red, if symptomatic; violet, if asymptomatic; turquoise, if symptomatic recovered; green, if asymptomatic recovered.

Doing the batches of repetitions of the simulation, we use random maps to have a neutral effect of the structure of the space.

We can set:

- min and max duration of the individual infection;
- the length of the incubation interval;
- the critical distance, i.e., the radius of the possibility of infection in open air, with a given probability;
- the corrections of that probability, due to the personal characteristics of both active and the passive agents;

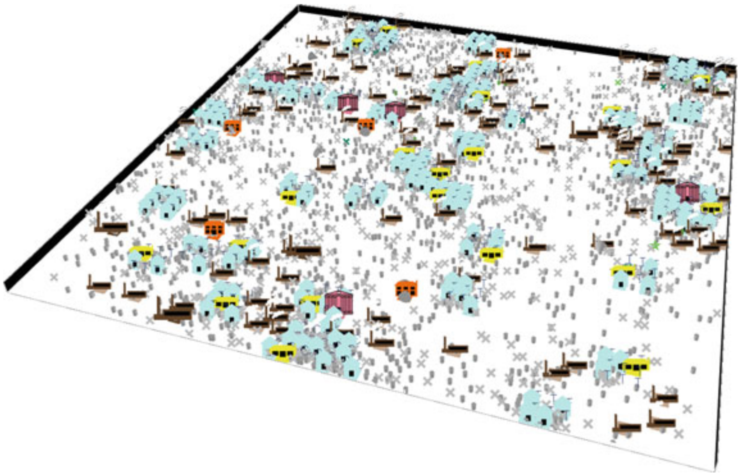


Fig. 4 A live 3D picture of the model world

- active agents can be symptomatic or asymptomatic, with different spreading characteristics (see (ii) in Sect. 2.2);
- passive agents, as receivers, can be robust (young), regular, fragile, and extra fragile.

We have two main types of contagion: (a) within a radius, for people moving around, temporary in a house/factory/nursing home/hospital; (b) in a given space (room or apartment) for people resident in their home or in a hospital or in a nursing home or being in school or in a working environment.

People in hospitals and nursing homes can be infected in ways (a) and (b). Instead, while people are at school, they can only receive the disease from people in the same classroom, where only teachers and students are present, so this is a third infection mechanism (c). In all cases, the personal characteristics of the recipients are decisive.

We remark that workplaces are open to all persons, as clients, vendors, suppliers, external workers can go there. In contrast, schools are reserved for students and school operators.

All agents have their home, inside a city, or a town. The agents also have usual places (UPs) where they act and interact, moving around. These positions can be interpreted as free time elective places. When we activate the schools, students and teachers have both UPs and schools; healthcare operators have both UPs and hospitals or nursing homes; finally, workers have both UPs and working places. In each day (or tick of the model), we simulated full sequences of actions.

Figure 5 describes what happens during every *day* in our simulated world, with the daily sequences of actions.

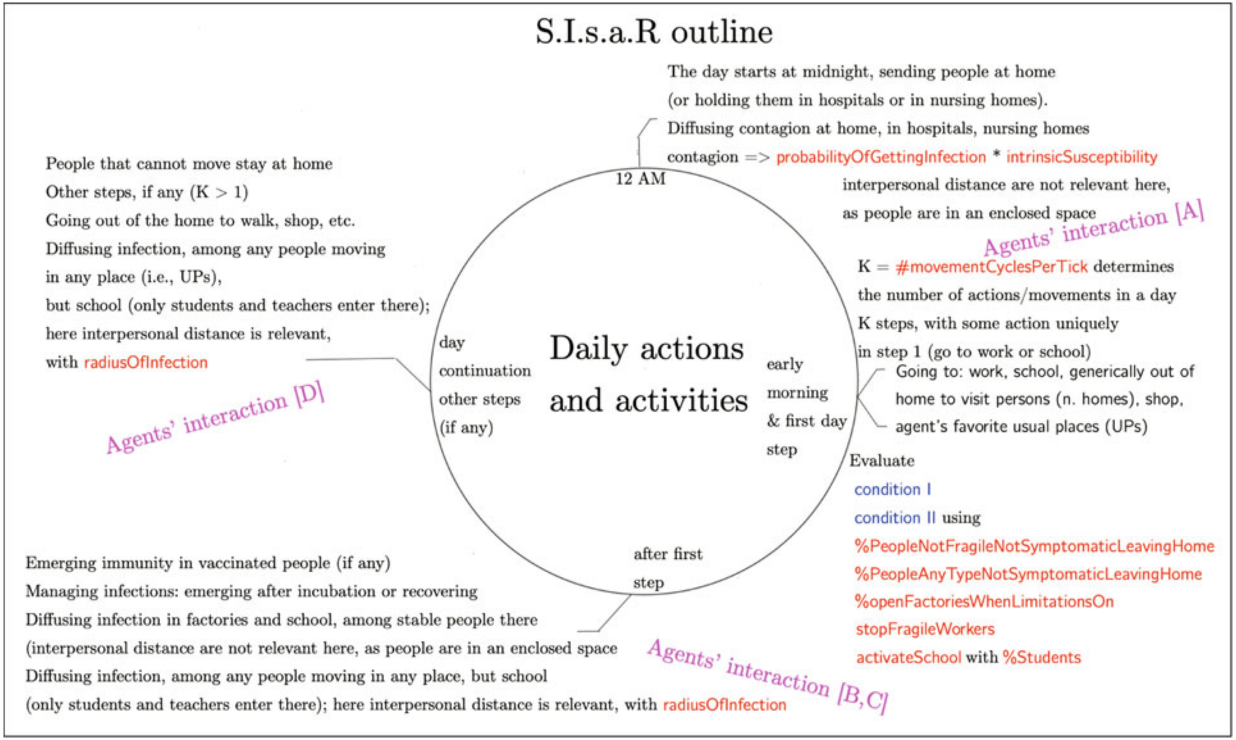


Fig. 5 A day in the simulation, with N repetitions where N is the duration of a given outbreak; look at: Sect. 2.1 for the rules of the conditional actions; Sect. 2.2 for the parameter definitions; and Sect. 2.3 for details on the agent interactions

2.1 Conditional Actions

Agents' movements in space, to go to work, school, and other UPs are subject to two interrelated general conditions.

I Symptomatic persons are at home or in a hospital or a nursing home and do not move.

II People not constrained by *condition I* can move if (primary rule) there are no general limitations (e.g., lockdown) *OR* if one of the following sub-conditions applies:

(a) agents who are hospital healthcare operators or nursing home healthcare operators;
(b) all people, according to the probability of moving of the whole non-symptomatic agents (Sect. 2.2, (iv));
(c) regular people, according to the probability of moving of the regular non-symptomatic agents (Sect. 2.2, (v));
(d) workers, if all the factories are open or it is open their own workplace (Sect. 2.2, (vi));
(e) teachers, if the schools are open (Sect. 2.2, (vii));
(f) students, if the schools are open, but with a possible quota limitation (Sect. 2.2, (viii)).

2.2 Parameter Definition

We define the parameters of Fig. 5, also with their short names used in program scripts, in round brackets. The values of the parameters are reported in detail in Appendix 1—Parameter values (Sect. 9).

(i) *probabilityOfGettingInfection* (`prob`) is the base probability of getting infected, to be multiplied by the *intrinsicSusceptibility* factor (iii); it is activated if the subject is within a circle of radius (ix) with an infected person; values at (Sect. 9, (i));
(ii) *D%*, without the short name, is the percent increasing or decreasing factor of the contagion spread of an asymptomatic subject, compared to that of a symptomatic one, value at (Sect. 9, (ii));
(iii) the *intrinsicSusceptibility* in defined in Eq. (1)

$$intrinsicSusceptibility = intrinsicSusceptibilityFactor^{groupFragility} \tag{1}$$

with *intrinsicSusceptibilityFactor* set to 5, and $groupFragility$ exponent set to:

1 for extra-fragile persons,
0 for fragile persons,
−1 for regular persons,
−2 young people from 0 to 24 years old;

(iv) *%PeopleAnyTypeNotSymptomaticLeavingHome* (`%PeopleAny`) determines, in a probabilistic way, the number of people of any kind going around in case of limitations/lockdown; the limitations operate only if the lockdown is on (into our simulated world, from day 20); values at (Sect. 9, (iv));
(v) *%PeopleNotFragileNotSymptomaticLeavingHome* (`%PeopleNot`) determines, in a probabilistic way, the number of regular people going around in case of limitations/lockdown; as above, the limitations operate only if the lockdown is on (into our simulated world, from day 20); values at (Sect. 9, (v)); we try to reproduce the uncertainty of the decisions in the real world into the model via frequent changes of the parameters (iv) and (v);

NB, the parameters (iv) and (v) produce independent effects, as in the following examples: (a) the activation of *%PeopleAny at 31, 0* and, simultaneously, of *%PeopleNot at 31, 80*, means that people had to stay home on that day, but people specifically not fragile could go out in 80% of the cases; (b) *%PeopleAny at 339, 80* and, simultaneously, *%PeopleNot at 339, 100* means that fragile and not fragile persons cannot always go around, but only in the 80% of the cases; instead, considering uniquely non-fragile persons they are free to go out; the construction is an attempt to reproduce a fuzzy situation; in future versions of the model, we will define the quotas straightforwardly:

- `%FragilePeopleNotSymptomaticLeavingHome`;
- `%NotFragilePeopleNotSymptomaticLeavingHome`;

(vi) *%openFactoriesWhenLimitationsOn* (`%Fac`) determines, in a probabilistic way, the factories (small and large industries, commercial surfaces, private and government offices) that are open when limitations are on; if the factory of a worker is open, the subject can go to work, not considering the restrictions (but uniquely in the first step of activity of each day); values at (Sect. 9, (vi));
(vii) *stopFragileWorkers* (`sFW`) is *off* (set to 0) by default; if *on* (set to 1), fragile workers (i.e., people fragile due to prior illnesses) can move out of their homes following the (iv) and (v) parameters, but cannot go to work; in the *off* case, workers (fragile or regular) can go to their factory (if open) also when limitations are on; values at (Sect. 9, (vii)); alternatively, we also have the `fragileWorkersAtHome` parameter; if *on* (set to 1) the total of the workers is unchanged, but the workers are all regular; we can activate this counterfactual operation uniquely at the beginning of the simulation;
(viii) when *activateSchools* (`aSch`) is *on* (set to 1), teachers and students go to school avoiding restrictions (but uniquely in the first step of activity of each day); *%Students* (`%St`) sets the quota of the students moving to school; the residual part is following the lessons from home; values at (Sect. 9, (viii));

(ix) following *radiusOfInfection* (`radius`), the effect of the contagion—outside enclosed spaces, or there, but for temporary presences—is possible within that distance; values at (Sect. 9, (ix));

(x) *asymptomaticRegularInfected%* and *asymptomaticFragileInfected%* are the parameters determining the percentage of asymptomatic persons after a contagion for non-fragile (all cases) or fragile people; they are without short names, as they come directly from the model interface; we can see the interface online, activating the model at https://terna.to.it/simul/SIsaR.html; values at (Sect. 9, (x)).

2.3 Agents' Interaction

We underline that our simulation tool is not based on micro-simulation sequences, calculating the contagion agent by agent, on the base of their characteristics and ex-ante probabilities. It implements a true agent-based simulation, with the agents acting and, most of all, interacting. The effect is that of generating continuously contagion situations.

Each run creates a population with expected characteristics, but also with random specifications, to assure the heterogeneity in agents. The daily choices of the agents are partially randomized, to reproduce real-life variability.

Contagions arise from agents' interactions, in four time phases, as specified in Fig. 5:

A in houses (at night), hospitals, nursing homes;
B in schools and workplaces in general, among people stable there;
C in the places above (excluding schools) by people temporary there and in open spaces (UPs above);
D interactions mainly in open spaces (UPs above).

3 Contagion Representation

We introduce a tool analyzing the contagions' sequences in simulated epidemics and identifying the places where they occur.

- We represent each infected agent as a horizontal segment (from the starting date to the final date of the infection) with vertical connections to other agents receiving the disease from it.
- We represent the new infected agents via further segments at an upper level.
- We display multiple information using three elements.

  - Colors in horizontal segments (areas of the infections): black for unknown places, gray for open spaces, cyan for houses, orange for nursing homes, pink for hospitals, yellow for schools, brown for factories, with shops and offices.
  - Vertical connecting segments keep the same color of the horizontal generating one.
  - Line thickness; proportional to fragility.
  - Styles: dotted lines for incubation, dashed lines for asymptomatic subjects, solid lines of symptomatic ones.

- This graphical presentation enables understanding at a glance how an epidemic episode is developing. In this way, it is easier to reason about countermeasures and, thus, to develop intervention policies.

At https://github.com/terna/contagionSequence we have the program *sequential-Records.ipynb*, generating these sequences.

Figure 6 is useful as an example. We start with two agents from the outside, with black as the color code (unknown place). The first one is young—as reported by the thickness of the segment, with the infection starting at day 0 and finishing at day 22—and asymptomatic (dashed line); it infects no one. The second one—regular, as reported by the thickness of the segment, with the infection starting at day 0 and finishing at day 15—is asymptomatic (dashed line) and infects four agents on day 2. All the four infected agents receive the infection at work (brown color) and turn to be asymptomatic after the days of incubation (dotted line); the first and the fourth are regular agents; the second and the third are fragile ones.

Continuing the analysis: on day 3, the second agent infects three other agents (at home, at work, at work) [...]; on day 13, agent number five infects seven regular agents at work and an extra-fragile one in a nursing home (orange color), etc.

If a vertical segment changes its color, we have an agent in an upper layer infecting someone on the same day of the infection transmitted by an agent in a lower row, so we lose some graphical information.

In Fig. 7 we see the example of an epidemic with non-pharmaceutical containment measures in adoption: a first wave shows an interlaced effect of contagions at home, in nursing homes, and at work. After a phase in which contagions develop mainly at home, a skinny bridge connects the first wave to a second one, which restarts from workplaces. The thickness of the *snake* of the contagions measures the stock on infects agents on a given date; the slope reports the speediness of the epidemic development; the upper vertical coordinate reports the cumulative number of infected people.

In Appendix 2—A gallery of contagion sequences (Sect. 10), we have several examples of contagion sequences.

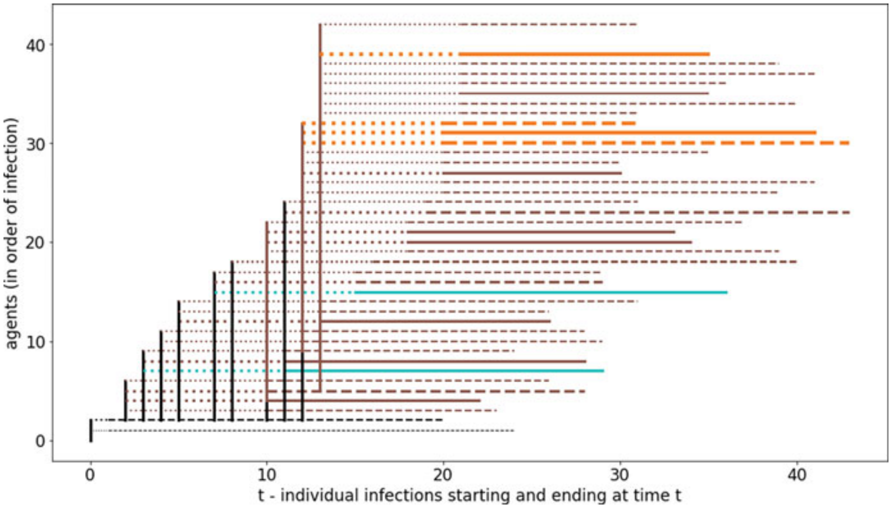


Fig. 6 A case with containment measures, first 40 infections: workplaces (brown) and nursing homes (orange) interweaving

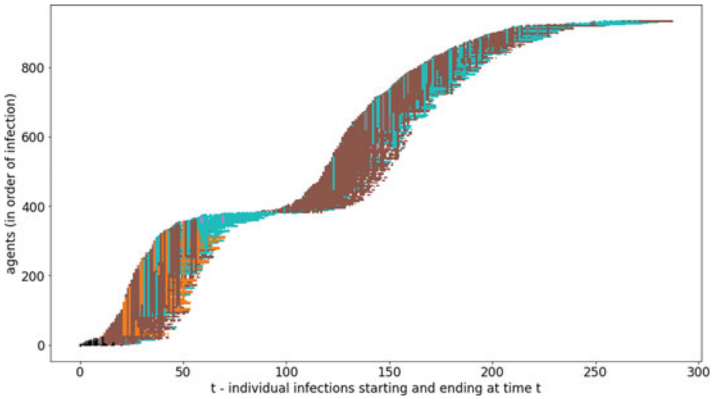
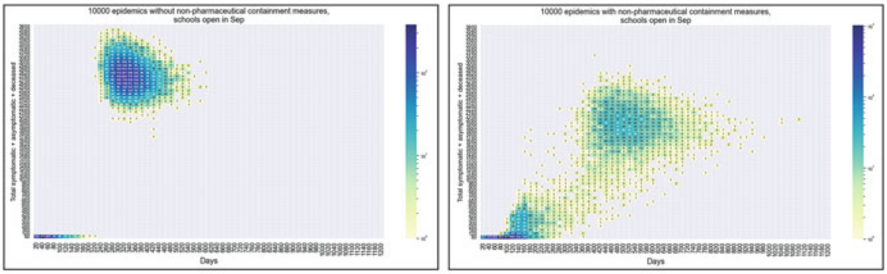


Fig. 7 A case with containment measures, the whole epidemics: workplaces (brown) and nursing homes (orange) and then houses (cyan), with a bridge connecting two waves

4 Exploring Scenarios with Simulation Batches

The sequences described in Sect. 3 suggest possible interventions, but are single cases. To explore systematically the introduction of factual, counterfactual, and prospective actions, we need to analyze batches of simulations. In this perspective, each simulation run—whose length coincides with the disappearance of symptomatic or asymptomatic contagion cases—is a datum in a set of different duration and contagion outcomes. To compare the consequences of each batch's basic



(a) (b)

Fig. 8 Starting our analyses: 10,000 epidemics in Piedmont. (**a**) Outbreaks without non-pharmaceutical containment measures. (**b**) Outbreaks with non-pharmaceutical containment measures

assumptions, we need to represent compactly the results emerging from simulation repetitions.

We use blocks of ten thousand repetitions. Besides summarizing the results with the usual statistical indicators, we adopt the technique of the heat-maps. With [23], our goal is that of making comparative analyzes, not forecasts. This consideration is consistent with the enormous standard deviation values that are intrinsic to the specific reality.

At https://github.com/terna/readSIsaR_BatchResults we have the codes producing the maps of the batches. A heat-map is a double histogram: in our application, it displays each simulated epidemic's duration in the x axis and the total number of the symptomatic, asymptomatic, and deceased agents in the y axis (on a scale of 1:1000). Each cell contains the number of epidemics with x duration and y outcome. Besides the number, a logarithmic color scale improves the readability of the maps.

4.1 Epidemics Without and With Control Measures

As a starting point, we compare the situations represented in Fig. 8a, b. In Fig. 8a, the heat-map reports the distribution in duration and infection causation of 10,000 simulated outbreaks left to spread without any control; coherently, with the school always open. The results in Table 2 are scary. The concentration of the cases in the heat-map shows that, except a few instances spontaneously concluding in a short period (left bottom corner), produces a heavy *cloud* of cases lasting one year or one year and a half, hitting (as symptomatic, asymptomatic, and deceased) from 2000 to 3500 persons on a total of 4350 in the region (scale of 1:1000).

In Fig. 8b and the related Table 3, we report a similar simulation batch of 10,000 runs of the model, but with the adoption of the basic non-pharmaceutical containment measures, registered in the values of the parameters in Appendix 1—Parameter values (Sect. 9). A calendar is at https://terna.to.it/simul/calendario092.pdf,

Table 2 Mean values and standard deviations in Fig. 8a cases

(000)	Symptomatic	Totalinfected&Deceased	Duration
Mean	969.46	2500.45	303.10
Std	308.80	802.88	93.50

Table 3 Mean values and standard deviations in Fig. 8b cases

(000)	Symptomatic	Totalinfected&Deceased	Duration
Mean	344.22	851.64	277.93
Std	368.49	916.41	213.48

and the model—version 0.9.6—is updated until April 2021. The results are dramatically different, showing the efficacy of the containment measures.

4.2 Actual Data

The critical points for our simulation experiments in Piedmont are Summer and Fall 2020 in Fig. 9a, where we have the time series of the first part of Piedmont's actual epidemic. The blue line represents the cumulative number of infected persons. Initially, only symptomatic cases were accounted for, but after the 2020 Summer, with more generalized tests, also asymptomatic patients are included:

- from http://www.protezionecivile.it/web/guest/department, the Italian Civil Protection Department web site, we find at https://github.com/pcm-dpc/COVID-1, i.e., the repository of regional data;
- we observe data about symptomatic infected people in the first wave, but from October 2020, data are mixed: in the above *git* repository, in October and November, we had "Positive cases emerged from clinical activity," unfortunately then reported as "No longer populated" (from the end of November, our observation) and "Positive cases emerging from surveys and tests, planned at national or regional level," again "No longer populated" (from the end of November, our observation);
- as a consequence, the subdivision between symptomatic and asymptomatic cases is impossible after that date.

Considering the dynamic of the data in Fig. 9a, we search within the simulation batch for cases with both:

(i) numbers of infected persons quite similar at `cp2` and at `cp3`; besides, numbers not too different from those of the figure; (with *cp*, we indicate the internal check points of the simulation program; in Fig. 9a we also report the number of days from the beginning of the epidemic for each check point);
(ii) the number of infected persons at `cp4` has to be significantly greater than those at the previous check point.

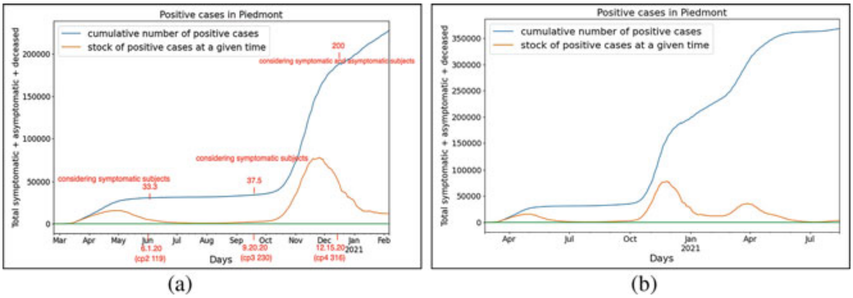


Fig. 9 Actual data. (**a**) Critical points in epidemic dynamic in Summer and Fall 2020 in Piedmont. (**b**) Data in Piedmont until July 2021, showing three waves

In a lot of cases, epidemics satisfying condition (i) fail to match condition (ii); both the situations happen only in less than the 1.5% of the instances in a batch of ten thousand epidemic. We can guess that the second wave registered in Piedmont after the Summer "pause" is due to new infected agents coming from outside and restarting the contagion process.

Other critical points in our analysis are the day on which the vaccination campaign starts, 373 of the simulation (Feb. 12th, 2021), and the day of the effectiveness of the initial vaccinations, 40 days later, day 413 (Mar. 22nd, 2021). At those dates, within the simulations, we can find either the presence of many infected agents or of few ones, as effectively was the situation in Piedmont.

NB, we concluded model calculations in April 2021. In Fig. 9b, the time series covering the whole period.

5 Factual and Counterfactual Analyses

In Fig. 10 we collect the heat-maps of the experiments:

- observing the emergence of spontaneous second waves, in the absence of specific control measures (Sect. 5.1);
- causing the emergence of the second wave through infections from outside, again in the absence of specific control measures (Sect. 5.2);
- causing the emergence of the second wave through infections from outside, in the presence of specific control measures (Sect. 5.3);
- reproducing the case of Sect. 5.3, anticipating by twenty days the start and end of all control measures (Sect. 5.5);
- reproducing the case of Sect. 5.3, limiting the control measures to fragile workers and other fragile people (Sect. 5.6).

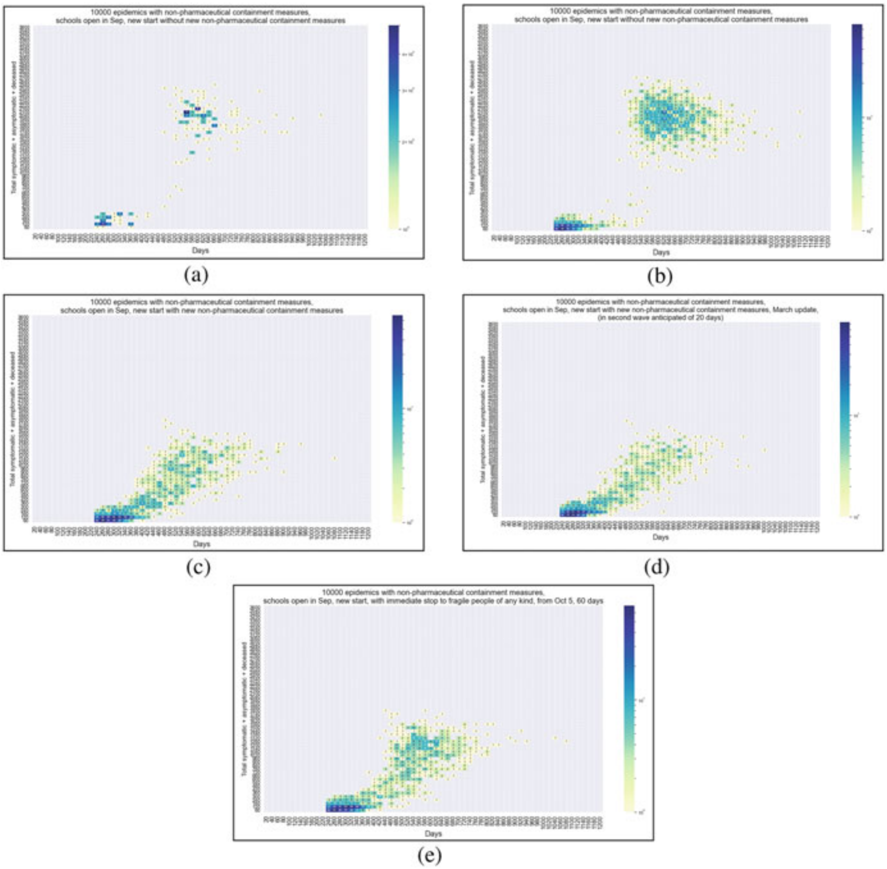


Fig. 10 Heat-maps of the factual and counterfactual analyses. (**a**) First wave with non-pharmaceutical containment measures, spontaneous second wave, without specific measures. (**b**) First wave with non-pharmaceutical containment measures, forcing the second wave, without specific measures. (**c**) First wave with non-pharmaceutical containment measures, forcing the second wave, *with new specific non-ph. containment measures*. (**d**) First wave with non-pharmaceutical cont. meas,, forcing the second w., *with new specific non-ph. cont. meas., acting 20 days in advance*. (**e**) First wave with non-ph. cont. meas., forcing the sec. wave; *in sec. wave, uniquely stopping fragile people, including fragile workers*

5.1 *Spontaneous Second Wave, Without Specific Containment Measures*

In an initial plain batch of runs of the Piedmont model, we count only 140 cases of epidemics with both the absence of new contagions in Summer 2020 and their explosion in Fall, as in Fig. 9a.

The steps are:

- we select, first of all, the 170 cases of epidemics that have, on June 1st, a number of symptomatic agents in the (10, 70] interval (with mean: 37.9) and, on September 20th, a number of symptomatic agents in the (20, 90] interval (with mean: 60.4);
- due to the lack of data described in Sect. 4.2, to compare December 15th and September 20th situations, we use symptomatic plus asymptomatic agents' count;
- we observe the existence of 140 outbreaks with the required characteristics; the December mean of the infected agents is 648.7, sensibly larger than the actual value: ≈200.0.

We overestimate the reality being the long-lasting simulated outbreaks, the larger ones, and, most of all, having no containment measures operating in the simulations.

Figure 10a and Table 4 show the outbreaks with similar cumulative numbers before and after the Summer 2020 "pause" (170 cases), with the second wave (140 cases) in the absence of containment measures.

140 out of 10,000, i.e., 1.4%, is a very light spontaneous ratio for the second wave occurred in the Fall. The transition to the third wave, that we see in Fig. 9b, is easy to explain, as the second wave never completely ended.

5.2 Second Wave, New Infections from Outside, Without Specific Containment Measures

To generate a framework consistent with the presence of a second wave after a period of substantial inactivity of the epidemic, we introduced two cases of infected persons coming back from outside after Summer vacancies, conventionally on September 1st, 2020.

As above, the steps are:

- we select, first of all, the 1407 cases of epidemics that. on June 1st have, a number of symptomatic agents in the (10, 70] interval (with mean: 35.6) and, on September 20th, a number of symptomatic agents in the (20, 90] interval (with mean: 40.0);
- due to the lack of data described in Sect. 4.2, to compare December 15th and September 20th situations, we use symptomatic plus asymptomatic agents' count;
- we observe the existence of 1044 outbreaks with the required characteristics; the December mean of the infected agents is 462.1, again sensibly larger than the actual value: ≈200.0.

We overestimate the reality being the simulations run without the adoption of containment measures.

Both Fig. 10b and Table 5 show the outbreaks with similar cumulative numbers before and after the Summer 2020 "pause" (1407 cases), with the second wave of

Table 4 Spontaneous second wave, without specific measures

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
Cum. v.	Sym.	All	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	days
Count	170.0	170.0	170.0	170.0	140.0	140.0	131.0	131.0	128.0	128.0	140.0	140.0	140.0
Mean	37.9	100.2	60.4	159.3	248.4	648.7	432.2	1109.5	656.3	1655.5	701.1	1757.9	594.2
Std	16.4	61.0	19.6	71.7	167.4	424.3	220.4	538.4	215.4	513.3	246.4	599.7	118.9

Table 5 Second wave, new infections from outside, without specific measures

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
Cum. v.	Sym.	All	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Days
Count	1407.0	1407.0	1407.0	1407.0	1044.0	1044.0	1005.0	1005.0	980.0	980.0	1044.0	1044.0	1044.0
Mean	35.6	72.7	40.0	84.1	180.4	462.1	354.1	900.4	623.8	1563.3	726.6	1810.9	620.9
Std	14.1	42.6	16.7	52.8	134.6	354.6	213.8	535.4	217.9	527.0	221.9	544.0	110.8

1044 cases. In the absence of containment measures, we have a heavy cloud as that of Fig. 8a, with infected people of any kind in a range approximately of 1500 to 2800 realizations, with an equivalence, to the Piedmont scale, to 1.5–2.8 millions of subjects.

The number of cases is now sufficient to evaluate the effects of factual (Sect. 5.3 and counterfactual (Sects. 5.5 and 5.6) simulation experiments.

5.3 Second Wave, New Infections from Outside, with New Specific Containment Measures

Repeating the third step above:

- we observe the existence of 874 outbreaks with the required characteristics; the December mean of the infected agents is 340.6, closer to the actual value (≈200.0) due to the introduction into the simulation of specific control measures for the second wave.

We always overestimate the reality because the surviving epidemics are the larger ones.

In Fig. 10c we see that the heavy cloud of the previous figure dissolved, and in Table 6 the numbers in italic emphasize the positive effects of the containment interventions on the cases of epidemic continuation (which have also dropped in quantity).

5.4 Calculating the Reproduction Number Without Delays

The reproduction number R_t [24, 25]

> is the average number of secondary cases of disease caused by a single infected individual over his or her infectious period

and is defined as follows:

$$R_t = \frac{I_t}{\sum_{s=1}^{t} w_s I_{t-s}}, \tag{2}$$

where:

- I_t is the number of new infected individuals at time t
- $w_s = \Gamma(s; \alpha, \beta)$ is the infectivity profile, usually approximated with the serial interval distribution [25]; it shapes the infectious period of each individual by weighting the infected individuals so that when their period is over, they do not

Table 6 Second wave, new infections from outside, with new specific measures

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
Cum. v.	Sym.	All	Sympt.	TotalInf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	TotalInf.	Sympt.	Totalinf.	Days
Count	1407.0	1407.0	1407.0	1407.0	874.0	874.0	719.0	719.0	523.0	523.0	874.0	874.0	874.0
Mean	35.6	72.7	40.0	84.1	*130.0*	*340.6*	*194.4*	*512.8*	*295.7*	*791.2*	252.7	666.4	494.1
Std	14.1	42.6	16.7	52.8	83.9	232.6	104.1	276.9	119.1	300.6	156.8	416.4	122.7

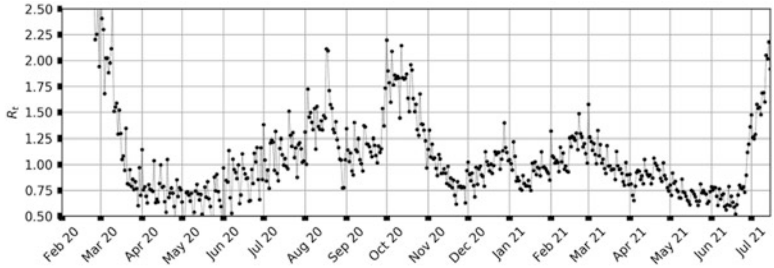


Fig. 11 Naive R_t calculated on raw infected cases by symptoms onset date, data-set by ISS

count any more in the sum; it is usually assumed to be the Gamma distribution [25]

– there is great uncertainty on the parameters of the Gamma distribution, which have been fitted to different values on different national data-sets ([26], table 1 page 25)
– following the Istituto Superiore di Sanità (ISS), italian R_t estimates are based on the parameters fitted in [27], namely $\alpha = 1.87$ and $\beta = 0.28$

While Eq. (2) could in principle be applied naively to any time series of the new infected cases, it usually leads to noisy results caused by the noise contained in the original raw series, as can be seen in Fig. 11. Moreover, despite the noisy content of the original signal, the naive approach does not give any clue on the confidence interval of the result, which is fundamental if the reproduction number has to be used to take decisions about the restrictions.

The most widely adopted approach to extract statistics about the R_t estimate, and hence its confidence interval, is to apply Bayesian statistical inference, assuming a prior distribution for the serial interval and a posterior for the reproduction number [25].

While Bayesian inference allows us to compute any kind of statistics on the estimate, it still fails dealing with the noise in the original signal, leading again to spiky estimates of R_t.

The standard solution to smooth out the noise is to assume that the transmissibility is constant over a time window (e.g., a week): we can then estimate the average R_t over the time window [25], by computing the total number of new infected cases over a window τ instead of those of each single day: $\hat{I}_{t,\tau} = \sum_{s=t-\tau}^{t} I_s$ and replacing it to I_t in Eq. (2); note that this is equivalent to compute R_t on the average of I_t over the window, as Eq. (2) is invariant under constant scaling of I_t.

The result is smoothed, but it turns out that it is delayed by the size of the windows. Figure 12 shows R_t calculated over a 14-day rolling window (14 days is the window size officially adopted in Italy); it is clearly visible that the average R_t is systematically delayed: maximizing the cross correlation of the signals confirms a measure of the delay of 14 days.

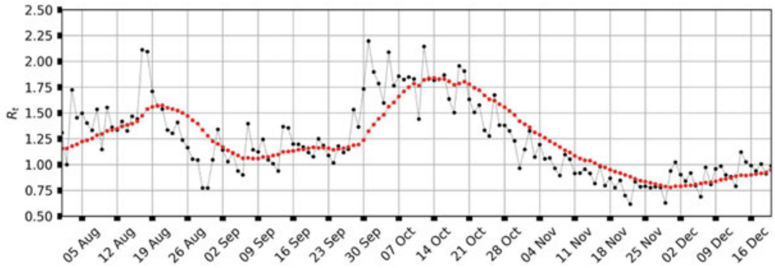


Fig. 12 In black R_t calculated on raw infected cases, in red the average R_t calculated over a window of 14 days; both series are by symptoms onset date, data-set by ISS

Official Data-Sets
Data used in all the computations refers to the following sources:

- data-set by ISS: count of new infected individuals by symptoms onset date, at https://github.com/tomorrowdata/COVID-19/blob/main/data/sources/ISS/covid_19-iss_2021-07-30T22:34:44%2B00:00.inizio_sintomi.csv downloaded on Jul 30
- data-set by Protezione Civile: count of new infected individuals by notification date, at https://github.com/pcm-dpc/COVID-19 downloaded on Jul 31

5.4.1 Tikhonov Regularization to Smooth the Original Signal

As an alternative solution to averages, we adopt Tikhonov regularization to the original signal, which does not introduce delays.

I_t is smoothed by fitting a series to represent the derivative of I_t and then integrating it back to the original signal, which then results in a smoothed one.

We search for the differential signal $\boldsymbol{\omega}$ such that:

$$\mathbf{I} = \mathbf{X} \cdot \boldsymbol{\omega},$$

where $\mathbf{I}$ denotes the array of elements I_t and $\mathbf{X}$ is the matrix representing the integration operator:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

$\boldsymbol{\omega}$ is obtained by minimizing the following cost function:

$$F(\boldsymbol{\omega}) = \|\mathbf{I} - \mathbf{X} \cdot \boldsymbol{\omega}\|^2 + \alpha^2 \|\boldsymbol{\Gamma} \cdot \boldsymbol{\omega}\|^2 . \qquad (3)$$

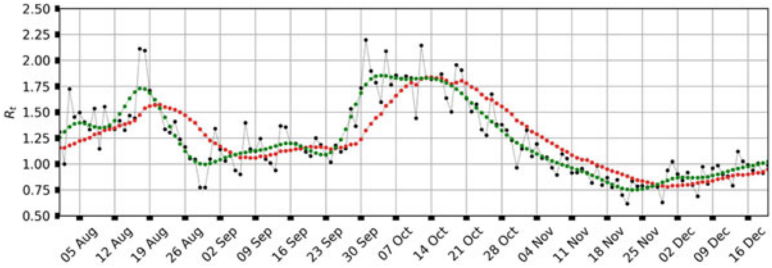


Fig. 13 In black R_t calculated on raw infected cases, in red the average R_t calculated over a window of 14 days; in green $\bar{R}_t$ calculated on the signal smoothed with Tikhonov regularization; each series is by symptoms onset date, data-set by ISS

Hence the derivative $\boldsymbol{\omega}$ is fitted using a Ridge regression with a generalized Tikhonov regularization factor:

- $\boldsymbol{\Gamma}$: the Tikhonov regularization matrix, chosen to be the second derivative operator;
- α: the regularization factor.

The regularization factor penalizes the spikes in the second derivative, forcing the derivative to be a smoothed signal. Once the derivative is fitted, the original signal is reconstructed by applying again the integral matrix to the differentiated smoothed signal; denoting the smoothed signal by $\bar{\mathbf{I}}$:

$$\bar{\mathbf{I}} = \mathbf{X} \cdot \boldsymbol{\omega}.$$

The parameter α can be obtained by searching the maximal smoothness constrained to the desired degree of information still available in the signal. It can be shown empirically that $\alpha = 100$ represents a reasonable trade-off.

Once we have $\bar{I}_t$ it can be fed into Eq. (2) to obtain the reproduction number computed on the smoothed signal, which we denote by $\bar{R}_t$.

Figure 13 shows in green the result of calculating $\bar{R}_t$ on the signal smoothed by minimizing Eq. (3). It is clearly visible that the green line anticipates the red one: maximizing the cross correlation of the original noisy R_t wrt $\bar{R}_t$ confirms a measure of the delay of 0 days.

5.4.2 Do Not Wait for the Symptoms Onset Date

Delays are not as important in literature, where we usually look at historical data, as they are in policy making, where we do need near real-time data.

Figure 14 shows the zoom of the series in Fig. 13 to the “present” days (which is Jul 30 at the time of writing). The last available “consolidated” count of new infected cases dates back to Jul 15, as the full process of data collection must be

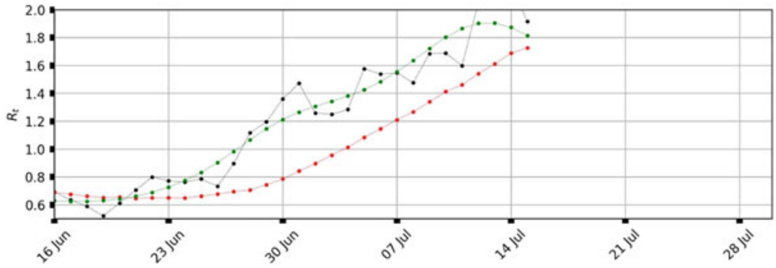


Fig. 14 Zoom of Fig. 13 to the most recent data available and consolidated, data-set by ISS

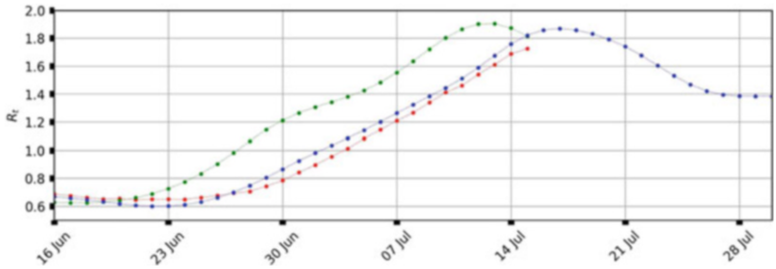


Fig. 15 In red the average R_t calculated over a window of 14 days on the symptoms onset distribution, data-set by ISS; in green $\bar{R}_t$ calculated on the symptoms onset distribution smoothed with Tikhonov regularization, data-set by ISS; in blue $\hat{\bar{R}}_t$ calculated on the notification date distribution smoothed, data-set by Protezione Civile

completed if we want to know the symptoms onset date. This problem, known as right censoring, is true for every country, with delays which vary depending on the particular data collection process. Moreover, it is well known in Italy that the collection process greatly depends on the pressure that the epidemic is producing on the Health System.

Instead of using the distribution of new infected cases by symptoms onset date, we propose to adopt the smoothed distribution by notification date as the input for Eq. (2), to obtain $\hat{\bar{R}}_t$. The difference is that as soon as a case is detected, it is notified. The advantage that the series is consolidated by definition, without the need of past revisions, comes with the following drawbacks:

1. there is a certain amount of delay from the symptoms onset date to the notification date;
2. the series accounts for more noise, as it makes no distinctions between symptomatic cases and asymptomatic cases.

Figure 15 shows the comparison of three R_t calculations: average R_t calculated on a 14-day window (in red), $\bar{R}_t$ calculated on smoothed cases by symptoms onset date (in green) and $\hat{\bar{R}}_t$ calculated on smoothed cases by notification date (in blue). The blue line exhibits a delay wrt the green one, but it is still anticipating the red line.

Maximizing the cross correlations provides the following measures of the relative delays:

- $\bar{R}_t$ anticipates $\hat{\hat{R}}_t$ by 8 days, but the last available value of $\bar{R}_t$ dates back 15 days prior to the present;
- $\hat{\hat{R}}_t$ anticipates R_t (calculated on a 14-day window) by 6 days.

Hence, we can conclude that, thanks to the smoothing procedure without delays (via Tikhonov regularization), we can replace the distribution of new cases by symptoms onset date with the distribution by notification date, obtaining the following advantages:

1. earn 6 days of anticipation with respect to the averaged R_t;
2. being able to compute the reproduction number up to the present, without having to wait for varying consolidation times in the data collection processes.

5.4.3 Residuals

As the original raw series are noisy and uncertain, we want a method to extract the noise and use it to calculate confidence intervals on the estimated R_t, in a way such that confidence intervals can directly reflect the uncertainty in the effective measuring process. This is much more relevant as we plan to estimate the reproduction number on the series of new infected cases by notification date, which includes both symptomatic and asymptomatic cases, with the latter exhibiting high noise.

The noise can be measured by the relative residuals of the signal with respect to its smoothed version, $\epsilon_t = (I_t - \bar{I}_t)/\bar{I}_t$.

Figure 16 shows that the distribution of ϵ_t calculated on the series of new infected cases by notification date **is unbalanced**.

It turns out that the unbalancing is directly related to the weekly seasonality which affects the series (the seasonality can be seen in Fig. 11 or Fig. 12). The reason is that the smoothing obtained by Eq. (3) is not able to capture the seasonality.

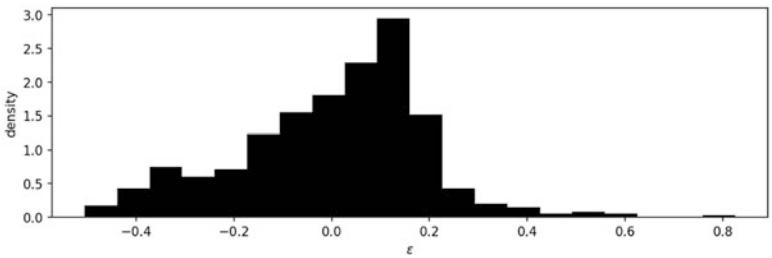


Fig. 16 Distribution of $\epsilon_t = (I_t - \bar{I}_t)/\bar{I}_t$ calculated on the series of new infected cases by notification date, data-set by Protezione Civile

5.4.4 Deseasoning via Singular Value Decomposition

Standard techniques to deal with seasonality, like SARIMA (Seasonal Autoregressive Integrated Moving Average), rely on moving averages.

To avoid the delays introduced by moving averages, we instead adopt Regularized Singular Value Decomposition (RSVD) proposed by Lin, Huang and Mcelroy in [28]. RSVD allows to detect the seasonal component of the signal by casting the signal vector into a matrix whose columns are the seasons and the rows are the repetitive periods of a complete series of seasons. Singular Value Decomposition is then applied to the matrix so that singular values represent the seasonal component of the signal. Each seasonal component is regularized via Tikhonov regularization, following the hypothesis that each seasonal component must change smoothly, period after period. The Tikhonov regularization parameter is fitted via "leave one out cross validation."

The advantage of this method with respect to the SARIMA approach is that *we do not need to take moving averages*, and we do not need to tune any meta-parameter of the model.

The python porting of the original R code is available in the supplementary material at https://github.com/tomorrowdata/COVID-19, within the library `covid19_pytoolbox`. The following features have been added to the original work:

- take the logarithm of the seasonal series, to remove exponential trends;
- differentiate the signal to a desired degree, to remove non-stationary trends in the original data, with an augmented Dickey–Fuller (ADF) test to check if any non-stationary component is present;
- apply Tikhonov regularization to the deseasoned signal to obtain the trend.

Denoting by $\tilde{I}_t$ the trend of the raw signal I_t after removing the seasonality, we obtain the following decomposition of the original series:

$$I_t = \tilde{I}_t + S_t + \tilde{E}_t \tag{4}$$

where S_t is the seasonal component and $\tilde{E}_t$ is the residual after deseasoning.

Figure 17 shows the result of applying RSVD to the series of new infected cases by notification date. RSVD has been applied to the logarithm of the second difference of the original series, with the ADF test confirming the removal of any non-stationary component. The smoothness of the seasonal components can be noted clearly.

5.4.5 Residuals of the Deseasoned Series

Now that we have removed the seasonality, we can look at the distribution of the residuals again. Figure 18 shows the distribution of the relative residuals after

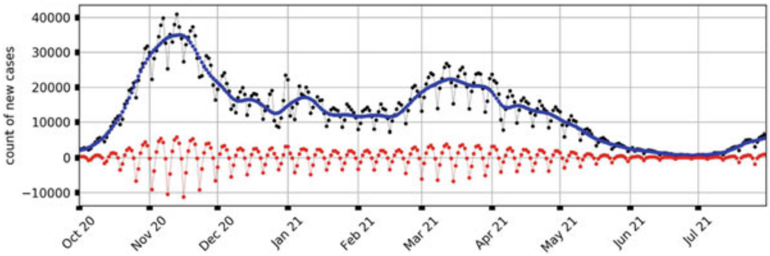


Fig. 17 Raw series I_t of new infected cases by notification date (in black), its trend $\tilde{I}_t$ after removing the seasonal component (in blue), the seasonal components S_t (in red); data-set Protezione Civile

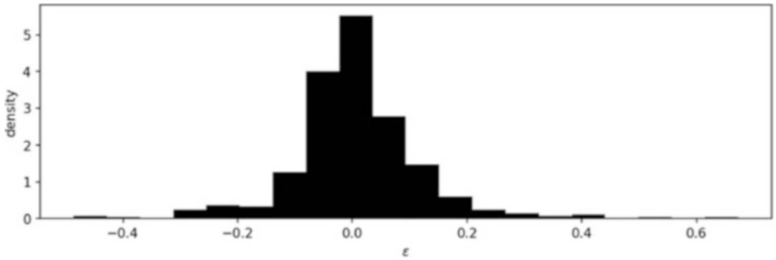


Fig. 18 Distribution of $\tilde{\epsilon}_t = \tilde{E}_t/\tilde{I}_t$ calculated on the series of new infected cases by notification date, data-set by Protezione Civile

removing the seasonal component: removing the seasonality produced a much more balanced, almost gaussian, distribution, if compared to Fig. 16.

5.4.6 Putting it All Together with Markov Chain Monte Carlo

We start from the series I_t of new infected cases by notification date, as explained in Sect. 5.4.2. We then apply RSVD to obtain the deseasoned smoothed trend $\tilde{I}_t$ of the series and the respective relative residuals $\tilde{\epsilon}_t = \tilde{E}_t/\tilde{I}_t$, as explained in Sect. 5.4.4.

With those ingredients, we can setup Markov chain Monte Carlo simulations to sample multiple chains of R_t values, as follows, denoting by $C(\cdot)$ the chains obtained via sampling:

1. $C(R_t)$ chains are sampled from a prior normal distribution, with $\mu = 1.3$ and $\sigma = 10$; a Gaussian process could be used instead, but it is less computationally efficient; the length of the chains is the same as the length of I_t;
2. $C(\tilde{\epsilon}_t)$ chains are sampled from a prior normal distribution, with $\mu = \sum_{s=t-7}^{t} \tilde{\epsilon}_s/7$ and $\sigma = \sqrt{\sum_{s=t-7}^{t}(\tilde{\epsilon}_s - \mu)^2/7}$; the length of the chains is the same as the length of I_t;

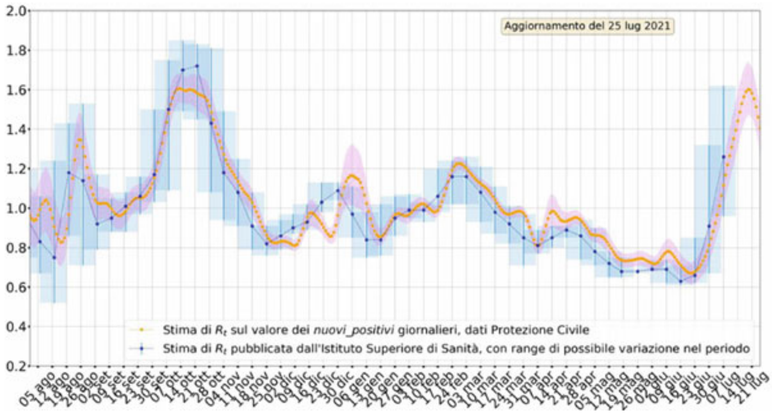


Fig. 19 In blue, the R_t values as reported by the Istituto Superiore di Sanità and in red the anticipated calculation published regularly, from the end of November 2020, at https://mondoeconomico.eu by Stefano Terna

3. $C(\tilde{I}_t)$, the chains of new cases with random noise, are obtained as $\tilde{I}_t + \tilde{I}_t \cdot C(\tilde{\epsilon}_t)$;
   - **note**: this is where the original noise of the series is transferred to the simulation, so that the confidence interval will account for uncertainties in the original series;
4. the estimated count T_t of new cases in each day of the chain is calculated from Eq. (2) as $C(T_t) = C(R_t) \cdot \sum_{s=1}^{t} w_s C(\tilde{I}_{t-s})$;
5. finally, a posterior Poisson distribution is tested via Monte Carlo, between the estimated cases, T_t, and the expected ones, $\tilde{I}_t$.

We sample 4 chains with 1000 iterations each discarded for tuning, and 500 iterations each kept for sampling. The final data-set contains 2000 samples from which day by day statistics, like the confidence interval, can be calculated.

Figure 19 shows the result, where the confidence interval (in violet) succeeds in representing periods of higher uncertainty in the data.

5.5 Second Wave, New Infections from Outside, Introducing 20 Days in Advance the New Specific Containment Measures

The counterfactual situation described in this section—inspired by Sect. 5.4—is related to the start and end dates of the actions of containment, both occurring 20 days in advance, with a natural barrier set on October 5th, 2020. Before that date, no one could plan to start new control measures.

As in the last two sections, we have 1407 cases of epidemics alive at the critical dates of June 1st and September 20th, after a Summer interval characterized by a quiet phase. Considering December 15th and September 20th situations, the second wave epidemics are 769, again decreasing because the anticipated actions have eliminated some other cases. The December mean of the infected agents is 294.2, still higher than the actual value (≈200.0). We always overestimate the mean of the epidemic effects, being the surviving epidemics the larger ones.

Comparing Fig. 10d and c the difference is not evident; instead, the italic figures, and most of all, the red bold ones—in Table 7—report clearly the comparative advantage of this counterfactual experiment with respect to the values of Table 6.

5.6 Second Wave, New Infections from Outside, with a Unique Intervention Measure: Stopping Fragile People for 60 Days

The second counterfactual experiment is based on an immediate stop to the circulation of fragile persons and specifically of fragile workers, plus isolating nursing homes and hospitals. Schools are always open in this experiment. The decision is activated on October 5th, 2020, when the second wave was becoming evident. In [29] we have important consideration suggesting the importance of taking into account fragility in a long-term fighting perspective against this kind of epidemics.

As in the last three sections, we have 1407 cases of epidemics alive at the critical dates of June 1st and September 20th, after a Summer interval characterized by a quiet phase. Considering December 15th and September 20th situations, the second wave epidemics are 886, lightly above the values of Sects. 5.3 and 5.5, but without locking the economy and the society as a whole. The December mean of the infected agents is 326.3, higher than the actual value (≈200.0) for the explained overestimation bias.

Comparing Fig. 10e and c the difference is not evident; instead, the italic figures, and most of all, the violet bold ones—in Table 8—signal the close proximity of the effects of this counterfactual experiment with those of Table 6.

5.7 To Recap

Table 9 reports the different cases synthetically and, most of all allows an easy comparative interpretation of the actual and counterfactual situations.

6 Economic Analysis of the of Interventions

The pandemic has an impact on the general economy. First, we take into account the additional health expenditure, which in Piedmont has risen from € 8880 million

Table 7 Second wave, new infections from outside, with new specific measure anticipation of −20 days

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
Cum. v.	Sym.	All	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Days
Count	1407.0	1407.0	1407.0	1407.0	769.0	769.0	637.0	637.0	471.0	471.0	769.0	769.0	769.0
Mean	35.6	72.7	40.0	84.1	**112.2**	**294.2**	*172.0*	*467.9*	*276.5*	*748.6*	248.9	663.4	499.3
Std	14.1	42.6	16.7	52.8	66.8	188.4	91.5	251.3	112.9	286.9	158.0	417.5	124.1

Table 8 Second wave, new infections from outs., stop fragile people. 60 days from Oct. 5

(1000)	Jun 1, 20		Sep 20, 20		Dec 15, 20		Feb 1, 21		May 1, 21		Dec 15, 20 to end		
Cum. v.	Sym.	All	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Sympt.	Totalinf.	Days
Count	1407.0	1407.0	1407.0	1407.0	886.0	886.0	761.0	761.0	637.0	637.0	886.0	886.0	886.0
Mean	35.6	72.7	40.0	84.1	**128.1**	**326.3**	*211.0*	*555.1*	*323.3*	*862.1*	301.1	792.3	515.5
Std	14.1	42.6	16.7	52.8	89.6	234.2	118.1	306.7	126.4	315.9	170.7	450.2	116.9

Table 9 Report of the key results, with count, mean, and std

Scenarios		Dec 15, 20		Dec 15, 20 to end		
		Sympt.	Totalinf.	Sympt.	Totalinf.	Days
No containment	Count	140.0	140.0	140.0	140.0	140.0
In spontaneous	Mean	248.4	648.7	701.1	1757.9	594.2
Second wave	Std	167.4	424.3	246.4	599.7	118.9
No containment	Count	1044.0	1044.0	1044.0	1044.0	1044.0
In forced	Mean	180.4	462.1	726.6	1810.9	620.9
Second wave	Std	134.6	354.6	221.9	544.0	110.8
Basic containment	Count	874.0	874.0	874.0	874.0	874.0
In forced	Mean	*130.0*	*340.*6	252.7	666.4	494.1
Second wave	Std	83.9	232.6	156.8	416.4	122.7
−20 days cont.	Count	769.0	769.0	769.0	769.0	769.0
In forced	Mean	**112.2**	**294.2**	248.9	663.4	499.3
Second wave	Std	66.8	188.4	158.0	417.5	124.1
Frag. subj. & workers control	Count	886.0	886.0	886.0	886.0	886.0
In forced	Mean	**128.1**	**326.3**	301.1	792.3	515.5
Second wave	Std	89.6	234.2	170.7	450.2	116.9

to €9200 million, with an increase in pressure on GDP of 0.2%. It is an increment that cannot be generalized. In other regions and States, health expenditure has even decreased, due to the lower demand for diagnostic and treatment services, precisely because of the pandemic and the precautionary reduced access to health services. Apart from the additional health expenditure, the main impact to be considered is the loss in production induced by the contagion containment measures, i.e., the so-called lockdown of the economy and the associated mobility bans.

The impact assessment of production stoppages and mobility bans can be measured by applying an Input-Output model. The main quality of Input-Output models is the possibility of determining the total effect of changes in output in all sectors of the economy due to a unit change in final demand in a given sector. This is achieved by applying a matrix of multipliers, i.e., Leontief's inverse matrix, to a sectoral vector of demand changes. The inverse matrix makes it possible to calculate the sum of the direct impact of the stopped productions, sector by sector, and the indirect impact, due to the infinite feedback on the purchases of the affected sectors from the first drop in demand received. However, the standard representation is not complete. The literature tends to extend these effects to consider the feedback not only by the purchases of the impacted sectors but also by the drop in the final demand of households affected by the unexpected change in income through their marginal propensity to consume. This third effect is the so-called induced impact.

The matrix of direct, indirect, and induced impacts of the Piedmont economy in Table 10 has been originally estimated by one of the authors.

As we can see, the economic effects of lockdowns can be very different depending on whether they selectively affect one sector (normally the sectors most

Table 10 Multipliers of direct, indirect, and induced impact, and overall impact as well added value (GDP) multipliers per 1 euro of final demand change, related to each of the five sectors on the rows

	Final/total demand	Direct impact	Indirect impact	Total Induced impact	Production multiplier	Added value multiplier
Agriculture	0.53	1.40	0.50	1.30	3.20	1.40
Manufacturing	0.33	1.80	1.20	1.60	4.50	1.60
Construction	0.38	1.70	0.90	1.60	3.20	1.60
Distribution	0.53	1.40	0.50	1.30	3.10	1.60
Services	0.50	1.50	0.50	1.40	3.40	1.50

affected are the last two, distribution and services), or whether all sectors are affected. The manufacturing sector, which is strongly linked with other sectors, has a total, direct, indirect, and induced multiplication coefficient of 4.5 times the initial reduction in final demand. Therefore, to calculate the impacts, we started from three different assumptions, or scenarios, which we have called A, B, and C.

[A] The restrictions affected all economic activities that could be stopped, safeguarding only those businesses that were essential. This meant stopping approximately half of the regional production system. Schooling was only permitted with distance learning. This case occurred in the period from 9 March 2020 to the end of April 2020.

[B] Only businesses in sectors whose activities were rated with a high risk of contagion were stopped: these activities included non-food retail trade, the tourism restaurant and hotel sector, the sport, recreation and entertainment sector, the cultural sector, and, of course, the whole education system, that was served by distance learning. The transportation sector was legally active but still impacted by an almost obligatory drop in demand. This case actually occurred at different times during 2020 and 2021 and significantly from October 2020 until spring 2021, with a break of a few weeks during the winter.

[C] Purely theoretical and not put in place, it was considered to stop only the fragile workers, leaving intact the education and all the activities stopped in case B. In this case, fragile workers are estimated to be 14% of the total, based on a national projection of the total number of 5.6 million fragile people under 65 in Italy. To make the calculation of the impact realistic, we assume that all fragile workers received sickness compensation equal to the lost wage that impacted on the overall tax loss, increasing it; we also assume that half of the production of fragile workers could still be produced with overtime or temporary work by other workers.

The results of the simulations are reported in Table 11 and are expressed in points per thousand of Piedmont's GDP.

In Table 11, each day of closure of productive activities (leaving only the essential ones and schools closed, i.e., open for distance learning) to counter the

Table 11 Economic impacts of the pandemic with three hypotheses of non-pharmaceutical containment measures applied; values are expressed in GDP points/1000

	Scenario A	Scenario B	Scenario C
Daily impacts			
Total production	−4.86	−1.23	−0.69
Added value	−2.12	−0.55	−0.30
Taxes	−0.91	−0.24	−0.35
Monthly impacts			
Total production	−145.7	−36.9	−20.7
Added value	−63.7	−16.6	−9.1
Taxes	−27.4	−7.1	−10.6

contagion and allow access to hospitals produces a loss of income (added value) equal to 2.1 per thousand of GDP and a worsening of the fiscal budget by 0.9 per thousand of GDP. One month of total closures, therefore, would cost an income loss of 6.4% and a worsening of the fiscal balance of 2.7% (Scenario A, actually implemented in Italy from 9 March 2020 to April).

Conversely, limiting closures to only distribution activities (non-food), as well as to school (open as distance learning), sports, culture and leisure, tourism, and restaurant services (as in the light lockdown established in October 2020 and subsequent months, Scenario B) would have produced a daily income loss of 0.55 per thousand of GDP (equivalent to 1.6% per month) and a fiscal loss of 0.24 per thousand per day and 0.7% of GDP per month.

The solution of protecting at home (and paying) only fragile workers, leaving all schools and productive activities open, would reduce the loss of income to 0.3 per thousand per day (0.91% per month). Although this solution (Scenario C, never actually implemented) is more convenient concerning the overall income loss, even 1/7 of that of scenario A and 1/2 of that of scenario B, it costs slightly more in fiscal terms than scenario B (−0.35 per thousand per day, instead of −0.24). However, it would seem preferable because it is the only option of the three that would allow the regular operation of the schools. According to the reliable Invalsi tests performed in 2021, the percentage of pupils in Italian schools who have not reached the minimum learning standards has increased by 10 percentage points based on the total number of pupils. If we were to put this loss of human capital on an economic balance sheet, we would have to consider the full cost of an additional year of schooling for 10% of the school population, both in terms of the cost of additional education plus the income lost for a 1-year delay in subsequent employment. A raw estimate of this cost would appear to be 1.3% of GDP, of which 0.58% for the additional cost of education and 0.75% for the income lost by postponing entry into employment by 1-year.

Following Table 12, in a C scenario, the cost of pandemic restrictions, for 3 months (hypothesis), would be 0.2% of annual GDP for increased health expenditures +2.7% of direct, indirect, and induced value-added (GDP) losses, plus 3.19% of GDP of public budget deterioration, while there would be no human capital losses. The total losses in scenario C would be 6.1% of annual GDP for three full

Table 12 Total losses simulating the three scenarios A, B, C, from activity and mobility restrictions, in GDP points/1000

	Scenario A three months	Scenario B three months	Scenario C three months
More health expenditure	−2.0	−2.0	−2.0
Added value or GDP loss	−191.0	−49.8	−27.2
Tax loss	−82.1	−21.4	−31.9
Human capital loss	−13.4	−13.4	0.0
Total loss (GDP/1000)	−288.6	−86.6	−61.1

months of restrictions. In scenarios B and A the total loss would have been much higher and specifically 8.6% and 28.8% of the pre-Covid GDP, respectively. It is also worth noting the distribution of losses by row. In scenario C, the losses in value-added, and thus the recession damage to the economy to be recovered, would be minimal, and the losses due to insufficient human capital formation would be zero. Nevertheless, the policies adopted have preferred the adoption of scenarios A and B.

7 Planning Vaccination Campaigns

7.1 Some Notes on Vaccines

Vaccines are biological products made from killed or attenuated microorganisms, from viruses or from some of their components (antigens), or from substances they produce made safe by chemical (e.g., formaldehyde) or heat treatment, while maintaining their immunogenic properties (https://www.who.int/vaccines); today, vaccines can be composed of proteins obtained by recombinant DNA techniques using genetic engineering approaches.

They usually contain, in addition to the antigenic fraction, sterile water (or a saline-based physiological solution), adjuvants, preservatives, and stabilizers. Adjuvants are included in the vaccine in order to enhance the immune system response; preservatives are added to prevent contamination of the preparate by bacteria; stabilizers are introduced to increase the shelf life of the product and to maintain the properties of the vaccine during storage.

How Vaccines Work: A Step Back in the Eighteenth Century

Although early forms of empirical immunization appear to have been present in different cultures (India and China; [30]), the creation of the first vaccine (for smallpox immunization) dates back to 1798 by Edward Jenner, an English physician. Jenner had noticed that milkmaids who became infected with cowpox (Vaccinia Virus), a virus that causes similar symptoms to human smallpox (Variola Virus or Smallpox virus) but not fatal, did not subsequently develop the disease [31].

This suggested that Cowpox inoculation could protect against Smallpox. Jenner decided to test his theory by inoculating an eight-year-old boy, the son of his gardener (sic!), James Phipps, with material taken from the cowpox lesions of a local milkmaid. As expected, James developed few local lesions and a modest fever. Two months later Jenner inoculated James with variolous matter from a case of human smallpox, without a sensible effect was produced: Jenner had proved that the boy had been immunized. By definition, all subsequent immunizations would be called vaccinations as in 1881 Louis Pasteur proposed it as a general term for the new protective inoculations, in honor of Jenner.

Once inoculated, vaccines (all of them), mimicking the first contact between man and pathogen, are able to stimulate an immunological response (humoral and cellular) as if this occurred through a natural contagion, although not leading to disease and without giving the associated complications. The rationale behind this phenomenon is immunological memory: the body/immune system that has already experienced a pathogenic microorganism, treasures the experience by responding rapidly to the same microorganism (the absence of immunological memory is the reason why Covid-19 emerged as a problem for humans). For some vaccines it is necessary to make recalls at a distance of time. Normally our body reacts to an unwanted host, but it can take up to two weeks to produce a sufficient amount of antibodies versus the pathogen. In the absence of vaccination, in this interval of time a pathogen can create damage to the body and even lead to death.

Types of Vaccines

A long way has been covered since 1798, and technologies have steadily improved to arrive at hi-tech vaccines such as those we are using today to fight Covid. The types of vaccine that exist today are:

- live attenuated vaccines (e.g., measles and tuberculosis): these are produced from infectious agents that have been rendered non-pathogenic;
- inactivated vaccines (e.g., poxvirus): these are produced using infectious agents killed by heat or chemicals;
- purified antigen vaccines (e.g., anti-meningococcal): these are produced by purifying specific components (bacterial or viral);
- anatoxin vaccines (e.g., tetanus): these are produced using molecules from the infectious agent, which are not capable of causing the disease on their own, but which can stimulate/activate the immune defenses of the vaccinee;
- recombinant protein vaccines (e.g., hepatitis B): these are produced using recombinant DNA technology, which involves inserting genetic material coding for the antigen (a protein/peptide) into microorganisms capable of producing the antigen specifically, allowing it to be purified;
- recombinant mRNA vaccines (e.g., Pfizer/BioNTec and Moderna): these are produced using an mRNA coding for a target gene encapsulated nanoparticle made with lipid bilayers; this information is able to drive the synthesis of an antigenic protein, through the cell machinery, and the triggering of the immunitary system;

- recombinant viral vector vaccines (e.g., Astrazeneca/Oxford): these are produced using an DNA coding for a target gene carried within a defective adenovirus (from human or from chimpanzee), able to vehicle the gene within the cell nucleus. This information is able to drive the synthesis of an antigenic protein, through the cell machinery, and the triggering of the immunitary system.

The latter types of vaccine (mRNA and Adenoviral vector-based) were today adopted mainly because they can be manufactured very quickly, and being their production process highly standardized. As a matter of fact, they are the fastest way to create a vaccine in the middle of a pandemic.

mRNA and Adenovirus-Based Vaccines
Let us take a step back. The CoV-SARS-2019 virus has on its surface a protein called Spike (S-protein), that it uses to enter a human cell via binding to the ACE2 receptor (Fig. 3). The S-protein has therefore been chosen as the specific target to produce a vaccine since it is exposed in large quantities on the surface of the virus.

mRNA-based and adenovirus-based vaccine for Covid target the S-protein through the production of a RNA messengers (mRNAs), the classical molecule that routinely instructs all the cells what to build. Once the S-protein is produced within the body and presented the immune system, it is considered an antigen, and the body starts producing antibodies against it. The same thing can be done by using a pre-made protein and injecting it, but its production, testing and approval is longer (years to decades) and more expensive.

In a mRNA-based vaccine (e.g., Pfizer/Moderna) the mRNA coding for the Cov-Sars-2019 spike (S-protein) is encapsulated in lipid nanoparticles. This preparate is then injected (usually in the deltoid muscle). After that, the nanoparticles fuse with the cell membranes and mRNA is released into the cell cytoplasm, without entering in the nucleus nor getting incorporated into the genomic DNA. In an adenovirus-based vaccine (e.g., AstraZeneca) the gene coding for the spike protein is inglobated as DNA in a defective Chimpanzee adenovirus which is not able to proliferate, alone. This virus once injected latch on the host cell and released DNA (carrying the spike protein gene) in the cell cytoplasm. DNA then migrates to the nucleus where it is transcribed into mRNA, which will migrate to the cytoplasm.

In both the vaccines, at this particular stage the mRNA coding for the Spike uses the cellular machinery (e.g., ribosomes) for being translated into protein, imitating virus-infection-like humoral immunity and cellular immunity [32]. Both mRNA-based and adenovirus-based vaccines are able to increase the host's anti-virus effects by increasing T cells' antigen reactiveness [33]. Normally, are these white blood cells, as the first defenses of the body, to "detect" the presence of the pathogen and to organize a protection by generating specific antibodies to combat it through B-lymphocytes, as particular blood cells deputed to antibody production [34] These antibodies cover the virus and prevent it from attacking our body.

The immune system memory can be compared to a human's memory. Once it encounters an unwanted visitor, it will remember it and will be able to recognize it in the future. This process typically takes a few weeks for the body to produce the antibodies, but these cells will be there to guard the body for a long time.

7.2 Planning a Vaccination Campaign Using Genetic Algorithms, with Non-pharmaceutical Containment Measures in Action

We compare the effect of choosing the vaccination quotas via genetic algorithms (GAs) with two predetermined strategies. Our model considers three hypotheses: vaccinated people still spread the contagion; they do not spread the contagion; they do it in the 50% of the case. We show here only the results of the first case, the worst (as we write, the Delta variant is spreading, with vaccinated people transmitting the infection).

The parameters of the GAs side of the model are contained in a special file, as described in the Info sheet of the model; at https://terna.to.it/simul/SIsaR.html start the model and look at the Info paragraph named *Using Genetic Algorithms*.

Important dates:

- in the internal calendar of the model, day 373 is February 12th, 2021; it is the starting point of the vaccinations in Piedmont;
- the effectiveness of the initial vaccinations, 40 days later, starts on day 413 (March 22nd, 2021).

A technical detail: we simulate the vaccination campaigns with the GAs using the BehaviorSearch program, https://www.behaviorsearch.org, strictly related to NetLogo.

7.2.1 Vaccination Groups

We take into consideration seven groups, in order of decreasing fragility, also considering the exposure to contagions:

g1 Extra-fragile people with three components;

- due to intrinsic characteristics: people in living in nursing homes;
- due to risk exposure:
  - nursing homes operators;
  - healthcare operators;

g2 teachers;
g3 workers with medical fragility;
g4 regular workers;
g5 fragile people without special characteristics;
g6 regular people, not young, not worker, and not teacher;
g7 young people excluding special activity cases (a limited number in *g1*).

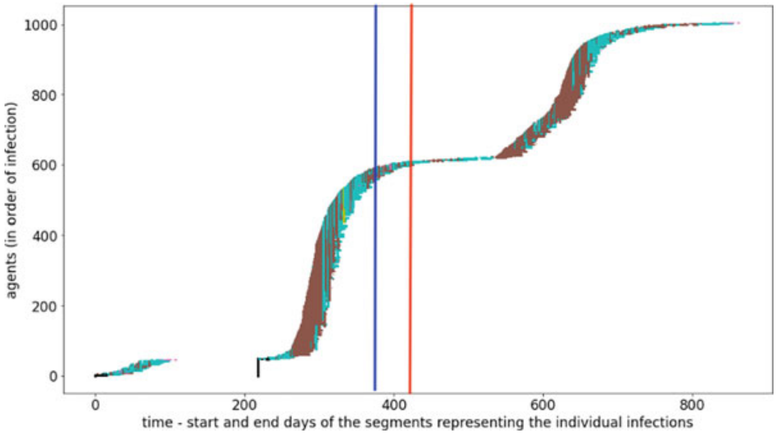


Fig. 20 Crucial dates: blue line for the starting point of the vaccination campaign and red line for the start of the effectiveness of the initial vaccinations; all the situations without vaccination

7.3 A Specific Realistic Case

The description of the vaccination effects on an outbreak is quite lengthy. Considering the collection at https://terna.to.it/simul/GAresultPresentation.pdf, we report here a unique case: the experiment I reported there, maintaining the reference to I in the titles of the figures. Considering the adoption of the government non-pharmaceutical measures, we search—in the batch of the 10,000 outbreaks of Sect. 5.3—for realizations of sequences similar to the actual events that occurred in Piedmont. As we see in Fig. 20 and the related Fig. 21, the artificial case that we adopt for the GAs exploration has the following critical characteristics:

(i) numbers of infected persons quite similar at `cp2` and at `cp3` in Fig. 9a; besides, numbers not too different from those of the same figure;
(ii) number of infected persons at `cp4` significantly greater than those at the previous checkpoint.

In Fig. 21, without vaccinations, we have the first wave in Spring 2020, a larger one in Fall 2020, a limited one between the end of 2020 and the beginning of 2021; then, a relatively quiet interval and successively, just while we write these notes, some restarting signals; finally, a fourth wave. Currently, it is in the future, relative to both the time of writing and the time when the calculations were completed (see NB at the end of Sect. 4.2). Very realistic with Piedmont's actual situation, the limited thickness of the *snake* of Fig. 20, when vaccinations start and when their effectiveness develops. The hole in the series identifies a period of quasi-extinct epidemics. Then it restarts with the arrival of infected persons from outside.

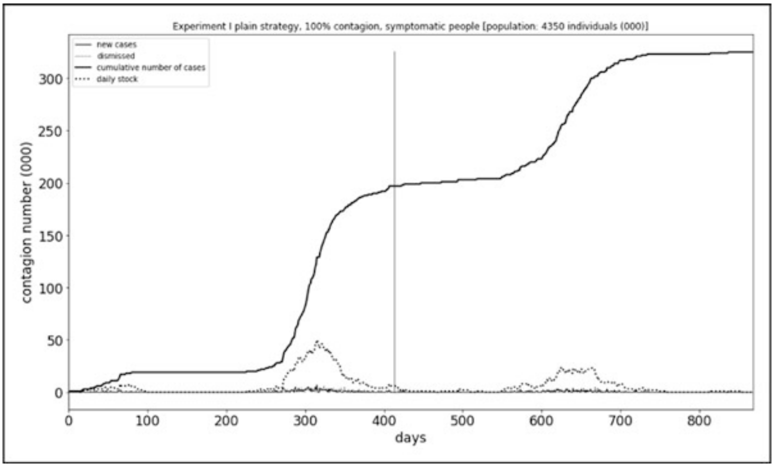


Fig. 21 Base symptomatic series; the vertical line at day 413 is not relevant here

Here and in the following sections, we analyze the count of symptomatic persons, being the goal of our simulated vaccination campaign exactly that of decreasing the number of symptomatic people, as deceased persons come from there.

7.4 Vaccination Quotas, Plain *Strategy*

The vaccination plans are related to the first dose; the second dose is supposed to be automatically scheduled, with an independent supply. The vaccinated person starts to benefit from immunity 40 days after the first dose.

Considering a *plain* option as that adopted in Table 13 with, in each day, the quantities of doses of the first column, we will primarily vaccinate the left column groups to move gradually to people of the other columns, as those on the left have already received the vaccine. The order is (*g1*) extra-fragile people, (*g2*) teachers, (*g3*) fragile workers, (*g4*) regular workers, (*g5*) fragile people, (*g6*) regular people, (*g7*) young people. In Table 14 we have numbers both of persons in each category at the beginning of this experiment (and in the following ones) and when the vaccination campaign starts.

Some of the coefficients in Table 13, and all the successive similar ones, are not used in two situations:

(i) when the persons of a group are fully vaccinated, the quotas in the rows below that day are not relevant;
(ii) when the people in the columns to the left of a given column completely absorb the available doses of vaccine on that day (the quotas in that column have unimportant values).

Table 13 From the day of the first column, considering the quantity of the second column (000), the vaccination of each group follows the quotas of the related columns

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.1	0.1	0.1	0.1	0.1	0.1	0.1
433	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
493	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
553	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
613	20	0.1	0.1	0.1	0.1	0.1	0.1	0.1
738	End							

Table 14 Susceptible persons at the beginning of the simulation and when the vaccination campaign starts, day 373, Feb. 12th, 2021

(000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
Susc. at $t = 0$	133	84	240	1560	1179	254	900
Susc. when vacc. starts	124	81	162	1234	1032	245	891

We anticipate that the GAs procedure does not optimize the coefficients of cases (i) and (ii).

The series that we introduce hereafter are significant from day 413, March 22nd, when the initial vaccinations' effectiveness begins, after 40 days from initial vaccinations.

In Fig. 22a we have the effects of the vaccination plan as numbers of vaccinated persons by groups. In Fig. 22b we have the most important outcome: the no vaccination test-bed is that of Fig. 21. We note the waves after the vertical line—when vaccinations start to operate—are lower than in the test plot, but anyway, those further waves are there.

7.5 *Vaccination Quotas,* Wise *Strategy*

Considering now a *wise* option, as an attempt to mimic the actual (and complex) vaccine distribution in the region, we use the quotas of Table 15, with the exact mechanism of the previous section. We primarily vaccinate the left column groups to move gradually to other columns, but postponing group *g4* (regular workers), *g6* (regular people), and *g7* (young people). In Table 14 we have numbers both of persons in each category at the beginning of this experiment (and in the following ones) and when the vaccination campaign starts. The considerations sub (i) and (ii) in Sect. 7.4 apply also here.

In Fig. 22c we have the effects of the vaccination plan as numbers of vaccinated persons by groups. In Fig. 22d we have the experiment outcome: the no vaccination test-bed is that of Fig. 21. We note the waves after the vertical line—when

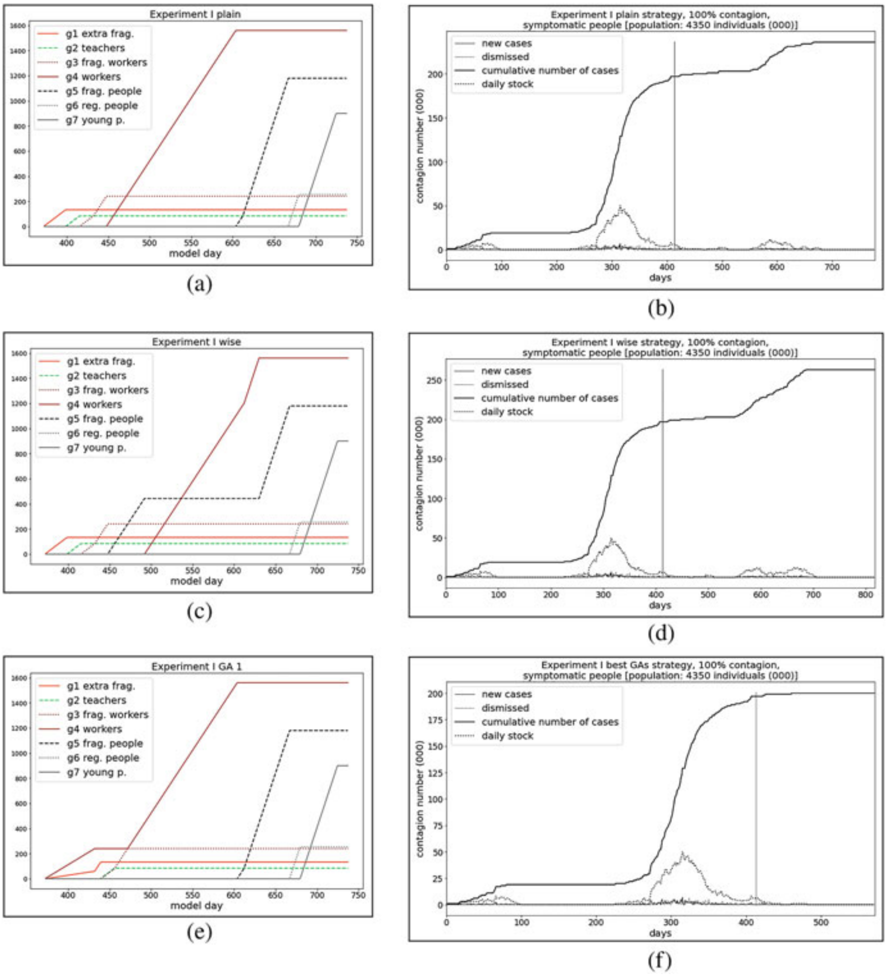


Fig. 22 Vaccination sequences and time series. (**a**) *Plain* vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal). (**b**) *Plain* vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts. (**c**) *Wise* vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal). (**d**) *Wise* vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts. (**e**) GA vaccination sequence; on the y axis the number of vaccinated subjects of each group (if vaccination is complete, the line is horizontal). (**f**) GAs vaccination symptomatic series; the vertical line is at day 413, when the effectiveness of first vaccination starts

vaccinations start to operate—are lower than in the test plot, but we have significant further waves in this case too.

Table 15 From the day of the first column, considering the quantity of the second column (000), the vaccination of each group follows the quotas of the related columns

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.1	0.1	0.1	0.0	0.1	0.0	0.0
433	10	0.1	0.1	0.1	0.0	0.1	0.0	0.0
493	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
553	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
613	20	0.1	0.1	0.1	0.1	0.1	0.1	0.1
738	End							

7.6 GAs Quotas in the Experiment, with Vaccinated People Spreading the Infection

Finally, this whole section's objective is to use GAs to evolve populations of models by choosing "genetically" the parameters to decide daily vaccination. Initially, on a random basis and successively considering them as a genetic chromosome of each model, re-productively crossed with those of other models. The search is for the best fitness related to the goal of reducing the number of symptomatic persons. [35], also quoted at https://www.behaviorsearch.org, is a helpful introduction to the methodology; the sources of the GAs used here are at https://github.com/terna/GAs. The GAs action, determining the vaccination quotas, optimizes the behavior of a *deciding* meta-agent, in a sort of *inverse generative social science perspective* [36].

With the GAs option, we use the quotas of Table 16, with the exact mechanism of the previous section. The considerations sub (i) and (ii) in Sect. 7.4 also apply here. We underline that the GAs procedure does not optimize the coefficients of those two cases, because they do not affect the fitness related to the goal of minimizing the number of symptomatic subjects.

In Table 14 we have numbers both of persons in each category at the beginning of the experiment and when the vaccination campaign starts.

In Fig. 22e we have the effects of the vaccination plan as numbers of vaccinated persons by groups. The main attention of the GAs is initially related to the groups: *g4* (workers), *g1* (extra-fragile persons), *g3* (fragile workers), *g2* (teachers). Then *g5* (fragile people), finally *g6* (regular people), and *g7* (young people). The priority is for highly circulating persons (workers and teachers), then for fragile persons.

In Fig. 22f we have the crucial result of this experiment: the no vaccination test-bed is always that of Fig. 21. With GAs' choices, the waves after the vertical line—when vaccinations start to operate—disappear, and the whole outbreak is a lot shorter.

Table 16 GAs best strategy with *vaccinated people still spreading the infection*: from the day of the first column, considering the quantity of the second column, the vaccination of each group follows the quotas of the related columns

From day	Q. of vaccines (000)	*g1*	*g2*	*g3*	*g4*	*g5*	*g6*	*g7*
373	5	0.01	0	0	0.79	0.18	0.38	0.19
433	10	0.94	0.06	0.32	0.54	0.19	0.83	0.5
493	10	0.97	0.97	0.74	0.79	0.2	0.14	0.52
553	10	0.98	0.83	0.02	0.39	0.99	0.04	0.48
613	20	0.52	0.01	0.83	0.6	1	0.27	0.9
738	End							

8 A New Model and Future Developments

Using SLAPP, https://terna.github.io/SLAPP/ a second model is under development, with a ratio of 1:100 to the Piedmont population, so 43,500 agents. It will contain the same items as the current one, plus transportation and aggregation places: happy hours, nightlife, sports, stadiums, discotheques, etc. We will also consider networks as family networks, professional networks, high-contact individual networks [37]. Finally, we will take into consideration the socioeconomic conditions of the individuals.

As seen, the S.I.s.a.R. model is a tool for comparative analyses, not for forecasting, mainly due to the enormous standard deviation values intrinsic to the problem.

The model is highly parametric, and more it will be, precisely in the comparative perspective. It also represents a small step in using artificial intelligence tools and the inverse generative perspective [8] in agent-based models.

9 Appendix 1—Parameter Values

We report here the values of parameters of Fig. 5, with their short names used in program scripts, in round brackets. Look at Sect. 2.2 for the definition. Day numbering is related to actual dates via Table 17. Day 1 is February 4th, 2020.

The values adopted in the experiments reported in this work are the following.

Table 17 The days of the simulation and their equivalent dates in the calendar

Day	Date	Day	Date	Day	Date	Day	Date
25	28-2-2020	200	21-8-2020	375	12-2-2021	550	6-8-2021
50	24-3-2020	225	15-9-2020	400	9-3-2021	575	31-8-2021
75	18-4-2020	250	10-10-2020	425	3-4-2021	600	25-9-2021
100	13-5-2020	275	4-11-2020	450	28-4-2021	625	20-10-2021
125	7-6-2020	300	29-11-2020	475	23-5-2021	650	14-11-2021
150	2-7-2020	325	24-12-2020	500	17-6-2021	675	9-12-2021
175	27-7-2020	350	18-1-2021	525	12-7-2021	700	3-1-2022

(i) The values of *probabilityOfGettingInfection* (`prob`) are: 0.05 (starting phase); 0.02 at day 49 (adoption of non-pharmaceutical measures); 0.035 at day 149 (some relaxation in compliance); 0.02 at day 266 (again, compliance to rules).

(ii) The value of *D%* is −50 in all the runs.

(iii) *intrinsicSusceptibility* is set discussing Eq. (1) in Sect. 2.2.

(iv) The values of *%PeopleAnyTypeNotSymptomaticLeavingHome* (`%PeopleAny`) are: at (day) 20, 90; at 28, 80; at 31, 0; at 106, 80; at 110, 95; at112, 85; at 117, 95; at 121, 90; at 259, 90; at 266, 80; at 277, 50; at 302, 70; at 320, 90; at 325, 50; at 329, 80; at 332, 50; at 336, 80; at 337, 50; at 339, 80.

(v) The values of *%PeopleNotFragileNotSymptomaticLeavingHome* (`%PeopleNot`) are: at (day) 31, 80; at 35, 70; at 36, 65; at 38, 15; at 42, 25; at 84, 30; at 106, 0; at 302, 90; at 325, 50; at 332, 50; at 337, 50; at 339, 100; at 349, 90.

(vi) The values of *%openFactoriesWhenLimitationsOn* (`%Fac`) are: at (day) 38, value4 0; at 49, 20; at 84, 70; at 106, 100; at 266, 90; at 277, 70; at 302, 80; at 320, 90; at 325, 30; at 329, 90; at 332, 30; 336, 90; at 337, 30; at 339, 100.

(vii) *stopFragileWorkers* (`sFW`): by default, 0; in one of the experiments we used `sFW` with set to 1 (on) at day 245 and to 0 (off) at day 275.

(viii) The values of *activateSchools* (`aSch`) are: at (day) 1, on; at 17, off; at 225, on; at 325, off; at 339, on; the values of *%Students* (`%St`) are: at (day) 0, 100; at 277, 50; at 339, 50; at 350, 50 (repeated values are not relevant for the model, but for the use of the programmer-author).

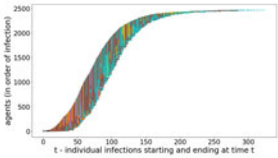
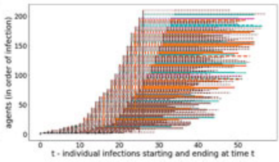
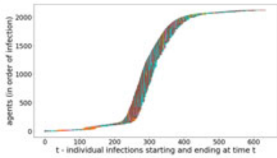
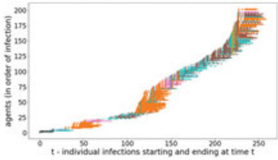
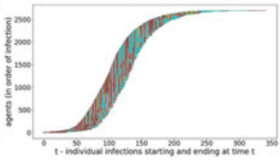
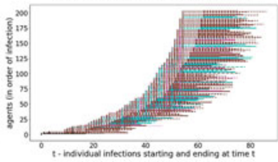
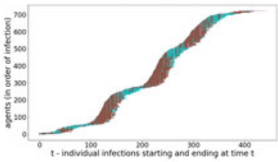
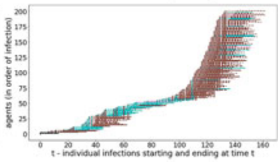
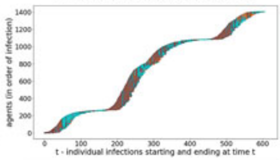
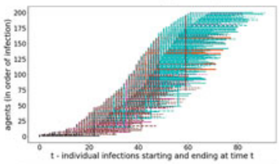
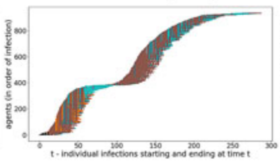
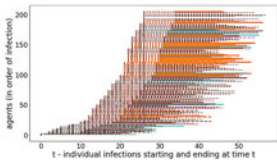
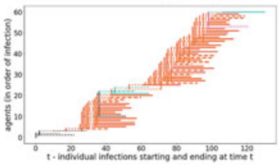
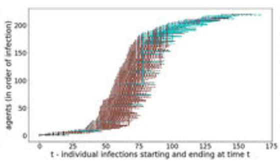
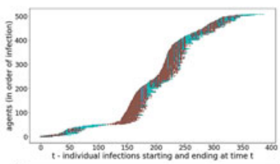
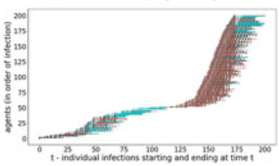
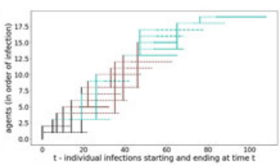
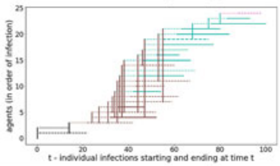
(ix) The value of *radiusOfInfection* (`radius`) is 0.2; in the model, space is missing of a scale, but forcing the area to be in the scale of a region as Piedmont, 0.2 is equivalent to 20 m; we have to better calibrate this measure with movements and probabilities; this is a critical step in future developments of the model.

(x) The values of *asymptomaticRegularInfected%* and *asymptomaticFragileInfected%* are 95 and 20.

10 Appendix 2—A Gallery of Contagion Sequences

The gallery of contagion sequences, reported in Table 18, shows the vast variety of situations generated by our agent-based simulations. What is significant is the variety of the situations.

Table 18 Gallery of sequences, symptomatic, and asymptomatic agents

Nr.	a	b	c
1			
2			
3			
4			
5			
6			

(1a) An outbreak without containment measures, with a unique wave, but very heavy: contagions are in nursing homes (orange), workplaces (brown), homes (cyan), hospitals (pink).
(1b) This is the previous epidemic without containment measures, considering the first 200 infections, with the main contribution of nursing homes (orange) and workplaces (brown).
(1c) Another outbreak, always without containment measures: nursing homes (orange) as a starter.
(2a) The (1c) epidemic, without containment measures, first 200 infections: nursing homes (orange) as a starter; around day 70, a unique contagion at home makes the epidemic continue.
(2b) Another case without containment measures showing the initial action of contagions in workplaces (brown) and homes (cyan).
(2c) Here we see the first 200 infections showing that the initial profound effects of contagions in workplaces (brown) and homes are due, in the beginning, to fragile persons, also asymptomatic,
(3a) An outbreak with containment measures, where we see another influential contribution of workplaces (brown) and homes (cyan) to the epidemic diffusion.
(3b) Here the first 200 infections: after day 100, we observe many significant cases of fragile workers diffusing the infection.
(3c) In this outbreak, with containment measures, the infections arise from workplaces (brown), nursing homes (orange), and homes (cyan), but also hospitals (pink).
(4a) Here we explore the first 200 infections of (3c): in the beginning, workplaces (brown), hospitals (pink), nursing homes (orange), and homes (cyan) are interweaving.
(4b) An outbreak with containment measures where the effect of the contagions in workplaces (brown), nursing homes (orange), and homes (cyan) is evident.
(4c) In the first 200 infections of (4b), workplaces (brown) and nursing homes (orange) are strictly interweaving.
(5a) An outbreak with containment measures where the effect of nursing homes (orange) is prevalent.
(5b) An outbreak with containment measures with a highly significant effect from workplaces (brown).
(5c) Stopping fragile workers at day 20 in the previous case, we obtain a beneficial effect, but home contagions (cyan) keep alive the pandemic, which explodes again in workplaces (brown).
(6a) Exploring the first 200 infections of the case (5c), we have evidence of the event around day 110 with the new phase due to a unique asymptomatic worker.
(6b) Finally, the same epidemic stopping fragile workers and any fragility at day 15 case and isolating nursing homes.
(6c) An outbreak with containment measures spontaneously stopping in a short period.

Acknowledgments Many thanks to Simone Landini, Nizar Mansour, Matteo Morini, Fabio Pammolli, Enrico Scalas, and Federico Tedeschi for their highly valuable discussions, insights, and critics. The usual disclaimer applies.

To run our model, we tremendously benefit from the use of HPC facilities provided by: Sarah de Nigris and Matteo Morini; http://tomorrowdata.io; the HP4AI (https://hpc4ai.unito.it/) and C3S centers at University of Torino.

References

1. A. Scala, A. Flori, A. Spelta, E. Brugnoli, M. Cinelli, W. Quattrociocchi, F. Pammolli, Scientific Reports **10**(1), 13764 (2020). https://doi.org/10.1038/s41598-020-70631-9
2. N. Bellomo, R. Bingham, M.A.J. Chaplain, G. Dosi, G. Forni, D.A. Knopoff, J. Lowengrub, R. Twarock, M.E. Virgillito, Mathematical Models and Methods in Applied Sciences **30**(08), 1591 (2020). https://doi.org/10.1142/S0218202520500323
3. H. Rahmandad, J. Sterman, Management Science **54**(5), 998 (2008)
4. F. Squazzoni, J.G. Polhill, B. Edmonds, P. Ahrweiler, P. Antosz, G. Scholz, E. Chappin, M. Borit, H. Verhagen, F. Giardini, N. Gilbert, Journal of Artificial Societies and Social Simulation **23**(2), 10 (2020). https://doi.org/10.18564/jasss.4298. http://jasss.soc.surrey.ac.uk/23/2/10.html
5. J.B. Bak-Coleman, M. Alfano, W. Barfuss, C.T. Bergstrom, M.A. Centeno, I.D. Couzin, J.F. Donges, M. Galesic, A.S. Gersick, J. Jacquet, A.B. Kao, R.E. Moran, P. Romanczuk, D.I. Rubenstein, K.J. Tombak, J.J. Van Bavel, E.U. Weber, Proceedings of the National Academy of Sciences **118**(27) (2021). https://doi.org/10.1073/pnas.2025764118. https://www.pnas.org/content/118/27/e2025764118
6. J. Epstein, Journal of Artificial Societies and Social Simulation **11**(4), 12 (2008). http://jasss.soc.surrey.ac.uk/11/4/12.html
7. R. Axtell, Why agents? On the varied motivations for agent computing in the social sciences. Tech. rep., Center on Social and Economic Dynamics Brookings Institution (2000)
8. T.M. Vu, C. Probst, J.M. Epstein, A. Brennan, M. Strong, R.C. Purshouse, Genetic and Evolutionary Computation Conference : [proceedings]. Genetic and Evolutionary Computation Conference **2019**, 1356 (2019). https://doi.org/10.1145/3321707.3321840. https://pubmed.ncbi.nlm.nih.gov/33083795
9. J.M. Epstein, Complexity **4**(5), 41 (1999). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.118.546&rep=rep1&type=pdf
10. G. Pescarmona, in *Recent research developments in biophysical chemistry*, ed. by C.A. Condat, A. Baruzzi (Trivandrum, 2002), pp. 1–22. http://flipper.diff.org/static/files/1517/The_life_context_2002.pdf
11. L. Gennero, M.A. Roos, P. D'Amelio, T. Denysenko, E. Morra, K. Sperber, V. Ceroni, M. Panzone, F. Lesca, E. De Vivo, A. Grimaldi, M.L. Gabetti, A. Ponzetto, G.P. Pescarmona, A. Pugliese, Cell Biochemistry and Function **28**(2), 142 (2010)
12. A. Vernone, P. Berchialla, G. Pescarmona, PloS one **8**(4), e60220 (2013). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0060220
13. A. Vernone, C. Ricca, D. Merlo, G. Pescarmona, F. Silvagno, Royal Society open science **6**(4), 181891 (2019). https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.181891
14. F. Scialo, A. Daniele, F. Amato, L. Pastore, M.G. Matera, M. Cazzola, G. Castaldo, A. Bianco, Lung pp. 1–11 (2020). https://link.springer.com/article/10.1007/s00408-020-00408-4
15. S. Shenoy, Inflammation Research **69**(11), 1077 (2020). https://doi.org/10.1007/s00011-020-01389-z. https://doi.org/10.1007%2Fs00011-020-01389-z
16. F. Silvagno, A. Vernone, G.P. Pescarmona, Antioxidants **9**(7), 624 (2020). https://doi.org/10.3390/antiox9070624

17. M. Soy, G. Keser, P. Atagündüz, F. Tabak, I. Atagündüz, S. Kayhan, Clinical rheumatology **39**, 2085 (2020). https://link.springer.com/content/pdf/10.1007/s10067-020-05190-5.pdf
18. B. Hu, S. Huang, L. Yin, Journal of medical virology **93**(1), 250 (2021). https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.26232
19. Y.d. Gao, M. Ding, X. Dong, J.j. Zhang, A. Kursat Azkur, D. Azkur, H. Gan, Y.l. Sun, W. Fu, W. Li, et al., Allergy **76**(2), 428 (2021). https://onlinelibrary.wiley.com/doi/full/10.1111/all.14657
20. A. Ghorbani, F. Lorig, B. de Bruin, P. Davidsson, F. Dignum, V. Dignum, M. van der Hurk, M. Jensen, C. Kammler, K. Kreulen, L.G. Ludescher, A. Melchior, R. Mellema, C. Păstrăv, L. Vanhée, H. Verhagen, Review of Artificial Societies and Social Simulation (2020). https://rofasss.org/2020/04/25/the-assocc-simulation-model/
21. R. Horton, Lancet (London, England) **396**(10255), 874 (2020). https://www.thelancet.com/action/showPdf?pii=S0140-6736%2820%2932000-6
22. U. Wilensky. Netlogo (1999). http://ccl.northwestern.edu/netlogo/
23. P. Steinmann, J.R. Wang, G.A. van Voorn, J.H. Kwakkel, Review of Artificial Societies and Social Simulation **17** (2020). URL https://rofasss.org/2020/04/17/deep-uncertainty/
24. L.M. Bettencourt, R.M. Ribeiro, PloS one **3**(5), e2185 (2008). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002185
25. A. Cori, N.M. Ferguson, C. Fraser, S. Cauchemez, American Journal of Epidemiology **178**(9), 1505 (2013). https://doi.org/10.1093/aje/kwt133
26. R. Anderson, C. Donnelly, D. Hollingsworth, M. Keeling, C. Vegvari, R. Baggaley, R. Maddren, The Royal Society **2020** (2020). https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf
27. D. Cereda, M. Tirani, F. Rovida, V. Demicheli, M. Ajelli, P. Poletti, F. Trentini, G. Guzzetta, V. Marziano, A. Barone, M. Magoni, S. Deandrea, G. Diurno, M. Lombardo, M. Faccini, A. Pan, R. Bruno, E. Pariani, G. Grasselli, A. Piatti, M. Gramegna, F. Baldanti, A. Melegaro, S. Merler. The early phase of the covid-19 outbreak in Lombardy, Italy (2020). https://arxiv.org/abs/2003.09320
28. W. Lin, J.Z. Huang, T. McElroy, Journal of Business & Economic Statistics **38**(3), 487 (2020). https://doi.org/10.1080/07350015.2018.1515081
29. N. Phillips, Nature **590**(7846), 382 (2021). https://www.nature.com/articles/d41586-021-00396-2
30. A. Boylston, Journal of the Royal Society of Medicine **105**(7), 309 (2012). https://doi.org/10.1258/jrsm.2012.12k044
31. E. Jenner, *An inquiry into the causes and effects of the variolae vaccinae, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox* (Springfield [Mass.] : Re-printed for Dr. Samuel Cooley, by Ashley & Brewer, 1802, 1800). http://resource.nlm.nih.gov/2559001R
32. M.A. Monslow, S. Elbashir, N.L. Sullivan, D.S. Thiriot, P. Ahl, J. Smith, E. Miller, J. Cook, S. Cosmi, E. Thoryk, et al., Vaccine **38**(36), 5793 (2020). https://www.sciencedirect.com/science/article/pii/S0264410X20308483
33. Y. Wang, Z. Zhang, J. Luo, X. Han, Y. Wei, X. Wei, Molecular Cancer **20**(1), 33 (2021). https://doi.org/10.1186/s12943-021-01311-z
34. W. Ratajczak, P. Niedźwiedzka-Rystwej, B. Tokarz-Deptuła, W. Deptuła, Central-European journal of immunology **43**(2), 194 (2018). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6102609/
35. J.H. Miller, Management Science **44**(6), 820 (1998). https://www.jstor.org/stable/pdf/2634650.pdf
36. T.M. Vu, C. Probst, J.M. Epstein, A. Brennan, M. Strong, R.C. Purshouse, in *Proceedings of the Genetic and Evolutionary Computation Conference* (2019), pp. 1356–1363. https://dl.acm.org/doi/abs/10.1145/3321707.3321840
37. G. Manzo, A. van de Rijt, Journal of Artificial Societies and Social Simulation **23**(4), 10 (2020). http://jasss.soc.surrey.ac.uk/23/4/10.html