Bocconi & Springer Series 7
Mathematics, Statistics, Finance and Economics

Giovanni Peccati
Matthias Reitzner *Editors*

Stochastic Analysis for Poisson Point Processes

Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry

Bocconi University Press

Springer

B&SS – Bocconi & Springer Series

Volume 7

More information about this series at http://www.springer.com/series/8762

Giovanni Peccati • Matthias Reitzner
Editors

Stochastic Analysis for Poisson Point Processes

Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry

Bocconi University Press

Springer

Editors
Giovanni Peccati
Unité de Recherche en Mathématiques
Université du Luxembourg
Luxembourg, Luxembourg

Matthias Reitzner
Osnabrück University
Osnabrück, Germany

ISSN 2039-1471 ISSN 2039-148X (electronic)
B&SS – Bocconi & Springer Series
ISBN 978-3-319-05232-8 ISBN 978-3-319-05233-5 (eBook)
DOI 10.1007/978-3-319-05233-5

Library of Congress Control Number: 2016945014

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing Switzerland

To Stefano Andrejs, Emma Elīza and Ieva
To Ilka, Jan and Lars Jesper

Preface

This book is a collection of original surveys focussing on two very active branches of modern (theoretical and applied) probability, namely the *Malliavin calculus of variations* and *stochastic geometry*. Our aim is to provide (for the first time!) a lively, authoritative and rigorous presentation of the many topics connecting the two fields, in a way that is appealing to researchers from both communities. Each survey has been compiled by leading researchers in the corresponding area. Notation, assumptions and definitions have been harmonized as closely as possible between chapters.

Roughly speaking, stochastic geometry is the branch of mathematics that studies geometric structures associated with random configurations, such as random graphs and networks, random cluster processes, random unions of convex sets, random tilings and mosaics, etc. Due to its strong connections to stereology and spatial statistics, results in this area possess a large number of important applications, e.g. to modelling and statistical analysis of telecommunication networks, geostatistics, image analysis, material science, and many more.

On the other hand, the Malliavin calculus of variations is a collection of probabilistic techniques based on the properties of infinite-dimensional operators, acting on smooth functionals of general point processes and Gaussian fields. The operators of Malliavin calculus typically generalize to an infinite-dimensional setting familiar objects from classical analysis, like, for instance, gradients, difference and divergence operators. When dealing with Malliavin calculus in the context of point processes (as is the case in the present book), one has typically to deal with a number of technical difficulties—related in particular to the intrinsic discrete structure of the underlying objects. As explained in the sections to follow, a crucial tool in partially overcoming these difficulties is given by Wiener–Itô chaotic decompositions, which play a role analogous to that of orthogonal expansions into series of polynomials for square-integrable functions of a real variable.

A fundamental point (which constitutes a strong motivation for the present book) is that, for many prominent models in stochastic geometry, Wiener–Itô chaotic decompositions and associated operators from Malliavin calculus are particularly accessible and amenable to analysis because the involved concepts and expressions have an intrinsic and very natural geometric interpretation.

Of particular interest to us is the application of these techniques to the study of probabilistic approximations in a geometric context, that is, of mathematical statements allowing one to assess the distance between the distribution of a given random geometric object and the law of some target random variable. Probabilistic approximations are naturally associated with variance and covariance estimates, as well as with limit theorems, such as Central Limit Theorems and Laws of Large Numbers, and are one of the leading threads of the whole theory of probability.

The interaction between stochastic geometry and Malliavin calculus is a young and active domain of research that has witnessed an explosion in interest during the past 5 years. By its very nature, such an interaction is a topic that stands at the frontier of many different areas of current research. Investigations gained particular momentum during an Oberwolfach conference in 2013, where many prominent researchers from both fields met for discussions. Since then, an increasing number of collaborations have been initiated or strengthened. Also, although several remarkable results have already been achieved in the field, for instance in the asymptotic study of the Boolean model, random graphs, random polytopes and random k-flats, many questions and problems (e.g. the derivation and use of effective concentration inequalities) remain almost completely open for future investigation.

It is the aim of this book to survey these developments at the boundary between stochastic analysis and stochastic geometry, to present the state of the art in both fields and to point out open questions and unsolved problems with the intention of initiating new research in and between the two areas.

The readership we have in mind includes researchers and graduate students who have a basic knowledge of concepts of probability theory and functional analysis. Most of the fundamental notions that are needed for reading the book are introduced and developed from scratch.

Last but not least, as editors, we would like to thank the numerous colleagues and friends who have been involved in this project: this book would not have been possible without their excellent contributions, as well as their enthusiasm and support.

Luxembourg, Luxembourg — Giovanni Peccati
Osnabrück, Germany — Matthias Reitzner
December 2015

Introduction

This book is composed of ten chapters, each of which contains a detailed state-of-the-art survey of a topic lying at the boundary of stochastic analysis and stochastic geometry.

The first four surveys can be seen as a "crash course" in stochastic analysis on the Poisson space. Starting from the careful construction of Malliavin operators on abstract Poisson spaces via Fock space representations (G. Last), the elegant combinatorial properties of (multiple) Poisson stochastic integrals are explored (N. Privault) and an introduction provided to variational formulae, allowing the reader to deal with the analytical study of expectations of Poisson functionals (I. Molchanov and S. Zuyev). Finally, J.-L. Solé and F. Utzet show how these tools can be extended to the more general framework of random measures with independent increments (sometimes called *completely random measures*) and Lévy processes.

The subsequent survey by D. Hug and M. Reitzner provides a careful introduction to the main objects of interest in modern stochastic geometry. As anticipated, this is a crucial step in our text, since the chapters to follow are all motivated by geometric considerations.

The reader will then enter the realm of the Stein and Chen-Stein methods for probabilistic approximations and be shown how to combine these techniques with Malliavin calculus operators (S. Bourguin and G. Peccati): this powerful interaction represents the very heart of the so-called *Malliavin–Stein method*. The survey by M. Reitzner and R. Lachièze-Rey discusses how one can use Malliavin–Stein techniques in order to deal with the asymptotic study of U-statistics. Further deep results involving U-statistics and extreme values in stochastic geometry are presented in the survey by M. Schulte and Ch. Thäle, while the chapter by S. Bourguin, C. Durastanti, D. Marinucci and G. Peccati focusses on some recent applications of the Malliavin–Stein approach in the context of Poisson processes defined on the sphere.

The book closes with an introduction (by L. Decreusefond, I. Flint, N. Privault and G.L. Torrisi) to recently developed Malliavin calculus techniques—in particular, integration by parts formulae—in the context of determinantal point processes.

We now present a more detailed description of the individual contributions composing the book.

Chapter 1: Stochastic Analysis for Poisson Processes (G. Last). The starting point of this survey is the definition of a Poisson point process on a general σ-finite measure space, and the consequent explanation of the fundamental multivariate *Mecke equation*. The next topic is the Fock space representation of square-integrable functions of a Poisson process in terms of iterated difference operators. The survey continues with the definition and properties of multiple stochastic Wiener–Itô integrals, from which one can deduce the *chaotic representation property* of Poisson random measures. This naturally leads to the definition of the fundamental *Malliavin operators*, which represent one of the backbones of the entire book. The final part presents the Poincaré inequality and related variance inequalities, as well as covariance identities based on the use of Glauber dynamics and Mehler's formula (in a spirit close to [16]). The content of this chapter represents a substantial expansion and refinement of the seminal reference [14] and provides a self-contained account of several fundamental analytical results on the Poisson space (see e.g. [10, 20]). A further connection with the classical logarithmic Sobolev estimates proved in [32] is discussed in the subsequent survey by Bourguin and Peccati.

Chapter 2: Combinatorics of Poisson Stochastic Integrals with Random Integrands (N. Privault). This survey provides a unique self-contained account of recent results on moment identities for Poisson stochastic integrals with random integrands, based on the use of functional transforms on the Poisson space. This presentation relies on elementary combinatorics based on the Fàa di Bruno formula, partitions and polynomials, which are used together with multiple stochastic integrals, finite difference operators and integration by parts. Important references that are discussed in this chapter include [3, 6, 26, 27]. The combinatorial content of many formulae can be regarded as a far-reaching generalization of classical product and diagram formulae on the Poisson space—as presented, for example, in the monograph [22].

Chapter 3: Variational Analysis of Poisson Processes (I. Molchanov and S. Zuyev). The framework of this chapter is that of a family of finite Poisson point process distributions on a general phase space. Given a functional F of point configurations, the expectation $\mathbb{E}(F)$ is regarded as a function of the intensity measure of the corresponding Poisson processes. Thus, the domain of F is the set of finite measures—which is a cone in the Banach space of all signed measures with a finite total variation norm. By explicitly developing the expectation $\mathbb{E}(F)$, one can show that under rather mild assumptions the function $\mathbb{E}(F)$ is analytic. As a byproduct, one establishes Margulis–Russo type formulae for the Poisson process and the Gamma-type result, which have proved extremely useful, for example in percolation theory and in stochastic geometry. The variation formulae obtained are then applied to constrained optimization where first order optimality conditions are established for functionals of a measure. The final part of the survey contains a discussion of applications of the above-described variational calculus, in particular, to numerical integration, statistical experiment design, and quantization of distributions. This chapter expands and refines the content of the seminal paper [17].

Chapter 4: Malliavin Calculus for Stochastic Processes and Random Measures with Independent Increments (J. L. Solé and F. Utzet). The purpose of this survey is twofold: first, to review the extension of Malliavin calculus for Poisson processes based on the *difference operator* or *add one cost operator* to Lévy processes; second, to extend that construction to some classes of random measures, mainly completely random measures. For Lévy processes, the approach is based, on the one hand, on the Itô–Lévy representation (Itô [11]) and on the chaotic expansion property, which provides a direct definition of the Malliavin operators, and, on the other hand, a construction of the canonical space for Lévy processes that allows a pathwise definition of the Malliavin derivative as a quotient operator (Solé et al. [31]). Recent results of Murr [18] concerning extensions of Mecke's formula to that setup will also be discussed.

Chapter 5: Introduction to Stochastic Geometry (D. Hug and M. Reitzner). This chapter introduces some fundamental notions from stochastic geometry and from convex geometry (see e.g. [12, 25, 29] for some comprehensive references on the subject). First, the necessary definitions from convex geometry are given, including Hausdorff distance, Minkowski addition, parallel sets, intrinsic volumes and their local extensions, which are used in the subsequent chapters of the book. Second, some important models of stochastic geometry are introduced: the Boolean model, random geometric graphs, intersection processes of Poisson flat processes and random mosaics. A selection of open problems from stochastic geometry is also presented, together with a description of important new results and directions of research.

Chapter 6: The Malliavin–Stein Method on the Poisson Space (S. Bourguin and G. Peccati). This chapter provides a detailed and unified discussion of a collection of recently introduced techniques (see e.g. [16, 21, 23, 24]), allowing one to establish limit theorems for sequences of Poisson functionals with explicit rates of convergence, by combining Stein's method (see e.g. [5, 19]) and Malliavin calculus. The Gaussian and Poisson asymptotic regimes are discussed in detail. It is also shown how the main estimates of the theory may be applied in order to deduce information about the asymptotic independence of geometric objects (see [1]).

Chapter 7: *U*-Statistics in Stochastic Geometry (R. Lachièze-Rey and M. Reitzner). A U-statistic of order k with kernel $f : X^k \to \mathbb{R}$ over a Poisson process is defined in [28] as

$$\sum_{x_1,\dots,x_k \in \eta^k_{\neq}} f(x_1,\dots,x_k)$$

under appropriate integrability assumptions on f. U-statistics play an important role in stochastic geometry since many interesting functionals can be written as U-statistics, such as intrinsic volumes of intersection processes, characteristics of random geometric graphs, volumes of random simplices, and many others (see for instance [7, 13, 15, 28]). It turns out that the Wiener–Itô chaos expansion of a U-statistic is finite and thus Malliavin calculus is a particularly suitable method. Variance estimates, the approximation of the covariance structure and limit theorems which have been out of reach for many years can be derived. In this chapter the reader will find the fundamental properties of U-statistics as well as an investigation of associated moment formulae. The main object of the chapter is to discuss the available univariate and multivariate limit theorems.

Chapter 8: Poisson Point Process Convergence and Extreme Values in Stochastic Geometry (M. Schulte and Ch. Thäle). Let η_t be a Poisson point process of intensity $t > 0$ over a measurable space X. One then constructs a point process ξ_t on the real line by applying a measurable function f to every k-tuple of distinct points of η_t. It is shown that ξ_t behaves after appropriate rescaling locally like a Poisson point process as $t \to \infty$ under suitable conditions on η_t and f. Via a de-Poissonization argument a similar result is derived for an underlying binomial point process. This method is applied to investigate several problems arising in stochastic geometry, including the Gilbert graph, the Voronoi tessellation, triangular counts with angular constraints, and line tessellations. The core of the survey rests on techniques originally introduced in reference [30].

Chapter 9: *U*-Statistics on the Spherical Poisson Space (S. Bourguin, C. Durastanti, D. Marinucci and G. Peccati). This survey reviews a recent stream of research on normal approximations for linear functionals and more general U-statistics of wavelet and needlet coefficients evaluated on a homogeneous spherical Poisson field (see [2, 9]). It is shown how, by exploiting results from [23] based on Malliavin calculus and Stein's method, it is possible to assess the rate of convergence to Gaussianity for a triangular array of statistics with growing dimensions. These results can be applied in a number of statistical applications, such as spherical density estimations, searching for point sources, estimation of variance and the spherical two-sample problem.

Chapter 10: Determinantal Point Processes (L. Decreusefond, I. Flint, N. Privault and G.L. Torrisi). Determinantal and permanental point processes were introduced in the 1970s in order to incorporate repulsion and attraction properties in particle models. They have recently regained interest due to their close links with random matrix theory. In this paper we survey the main properties of such processes from the point of view of stochastic analysis and Malliavin calculus, including quasi-invariance, integration by parts, Dirichlet forms and the associated Markov diffusion processes (see [4, 8]).

References

1. Bourguin, S., Peccati, G.: Portmanteau inequalities on the Poisson space: mixed limits and multidimensional clustering. Electron. J. Probab. **19**(66) (2014)
2. Bourguin, S., Durastanti, C., Marinucci, D., Peccati, G.: Gaussian approximation of nonlinear statistics on the sphere. J. Math. Anal. Appl. **436**(2), 1121–1148 (2016)
3. Breton, J.-C., Privault, N.: Factorial moments of point processes. Stoch. Process. Appl. **124**, 3412–3428 (2014)
4. Camilier, I., Decreusefond, L.: Quasi-invariance and integration by parts for determinantal and permanental processes. J. Funct. Anal. **259**(1), 268–300 (2011)
5. Chen, L.H.Y., Goldstein, L., Shao, Q.-M.: Normal Approximation by Stein's Method. Springer, Berlin (2011)
6. Decreusefond, L., Flint, I.: Moment formulae for general point processes. J. Funct. Anal. **267**, 452–476 (2014)
7. Decreusefond, L., Ferraz, E., Randriambololona, H.: Simplicial homology of random configurations. Adv. Appl. Probab. **46**(2), 325–347 (2014)
8. Decreusefond, L., Flint, I., Privault, N., Torrisi, G.L.: Stochastic dynamics of determinantal processes by integration by parts. Commun. Stoch. Anal. (2016, to appear)
9. Durastanti, C., Marinucci, D., Peccati, G.: Normal approximation for wavelets coefficients on spherical Poisson fields. J. Math. Anal. Appl. **409**(1), 212–227 (2014)
10. Houdré, C., Perez-Abreu, V.: Covariance identities and inequalities for functionals on Wiener space and Poisson space. Ann. Probab. **23**, 400–419 (1995)
11. Itô, K.: Spectral type of the shift transformation of differential processes with stationary increments. Trans. Am. Math. Soc. **81**, 253–263 (1956)
12. Kendall, W., Molchanov, I. (ed.): New Perspectives in Stochastic Geometry. Oxford University Press, Oxford (2010)
13. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space. I: contractions, cumulants and geometric random graphs. Electron. J. Probab. **18**(32) (2013)
14. Last, G., Penrose, M.: Poisson process Fock space representation, chaos expansion and covariance inequalities. Probab. Theory Relat. Fields **150**, 663–690 (2011)
15. Last, G., Penrose, M.D., Schulte, M., Thäle, C.: Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**, 348–364 (2014)
16. Last, G., Peccati, G., Schulte, M.: Normal approximations on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization. Probab. Theory Relat. Fields (2016+, to appear)
17. Molchanov, I., Zuyev, S.: Variational analysis of functionals of Poisson processes. Math. Oper. Res. **25**(3), 485–508 (2000)
18. Murr, R.: Characterization of infinite divisibility by duality formulas. Application to Lévy processes and random measures. Stoch. Process. Appl. **123**, 1729–1249 (2013)
19. Nourdin, I., Peccati, G.: Normal Approximations Using Malliavin Calculus: From Stein's Method to Universality. Cambridge University Press, Cambridge (2011)
20. Nualart, D., Vives, J.: Anticipative calculus for the Poisson process based on the Fock space. Lect. Notes Math. **1426**, 154–165 (1990)
21. Peccati G.: The Chen-Stein method for Poisson functionals. Manuscript (2011). Available at https://arxiv.org/abs/1112.5051
22. Peccati, G., Taqqu, M.: Wiener Chaos: Moments, Cumulants and Diagrams. Bocconi & Springer Series, vol. 1, Springer, Milan (2011)
23. Peccati, G., Zheng, C.: Multi-dimensional Gaussian fluctuations on the Poisson space. Electron. J. Probab. **15**, 1487–1527 (2010)
24. Peccati, G., Solé, J.L., Taqqu, M.S., Utzet, F.: Stein's method and normal approximation of Poisson functionals. Ann. Probab. **38**, 443–478 (2010)
25. Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)

26. Privault, N.: Generalized Bell polynomials and the combinatorics of Poisson central moments. Electron. J. Combin. **18**(1), Paper 54, 1–10 (2011)
27. Privault, N.: Moments of Poisson stochastic integrals with random integrands. Probab. Math. Stat. **32**(2), 227–239 (2012)
28. Reitzner, M., Schulte, M.: Central limit theorems for U statistics of poisson point processes. Ann. Prob. **41**(6), 3879–3909 (2013)
29. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Springer, New York (2008)
30. Schulte, M., Thäle, C.: The scaling limit of Poisson-driven order statistics with applications in geometric probability. Stoch. Process. Appl. **122**(12), 4096–4120 (2012)
31. Solé, J.-L., Utzet, F., Vives, J.: Canonical Lévy process and Malliavin calculus. Stoch. Process. Appl. **117**(2), 165–187 (2007)
32. Wu, L.: A new modified logarithmic Sobolev inequality for poisson point processes and several applications. Probab. Theory Relat. Fields **118**, 427–438 (2000)

Contents

Stochastic Analysis for Poisson Processes

Günter Last

Abstract This chapter develops some basic theory for the stochastic analysis of Poisson process on a general σ-finite measure space. After giving some fundamental definitions and properties (as the multivariate Mecke equation) the chapter presents the Fock space representation of square-integrable functions of a Poisson process in terms of iterated difference operators. This is followed by the introduction of multivariate stochastic Wiener–Itô integrals and the discussion of their basic properties. The chapter then proceeds with proving the chaos expansion of square-integrable Poisson functionals, and defining and discussing Malliavin operators. Further topics are products of Wiener–Itô integrals and Mehler's formula for the inverse of the Ornstein–Uhlenbeck generator based on a dynamic thinning procedure. The chapter concludes with covariance identities, the Poincaré inequality, and the FKG-inequality.

1 Basic Properties of a Poisson Process

Let $(\mathbb{X}, \mathscr{X})$ be a measurable space. The idea of a *point process* with *state space* $\mathbb{X}$ is that of a random countable subset of $\mathbb{X}$, defined over a fixed probability space $(\Omega, \mathscr{A}, \mathbb{P})$. It is both convenient and mathematically fruitful to define a point process as a random element η in the space $\mathbf{N}_\sigma(\mathbb{X}) \equiv \mathbf{N}_\sigma$ of all σ-finite measures χ on $\mathbb{X}$ such that $\chi(B) \in \mathbb{Z}_+ \cup \{\infty\}$ for all $B \in \mathscr{X}$. To do so, we equip $\mathbf{N}_\sigma$ with the smallest σ-field $\mathscr{N}_\sigma(\mathbb{X}) \equiv \mathscr{N}_\sigma$ of subsets of $\mathbf{N}_\sigma$ such that $\chi \mapsto \chi(B)$ is measurable for all $B \in \mathscr{X}$. Then $\eta : \Omega \to \mathbf{N}_\sigma$ is a point process if and only if

$$\{\eta(B) = k\} \equiv \{\omega \in \Omega : \eta(\omega, B) = k\} \in \mathscr{A}$$

for all $B \in \mathscr{X}$ and all $k \in \mathbb{Z}_+$. Here we write $\eta(\omega, B)$ instead of the more clumsy $\eta(\omega)(B)$. We wish to stress that the results of this chapter do not require special (topological) assumptions on the state space.

G. Last (✉)
Karlsruhe Institute of Technology, Institute of Stochastics, 76128 Karlsruhe, Germany
e-mail: guenter.last@kit.edu

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_1

The *Dirac measure* δ_x at the point $x \in \mathbb{X}$ is the measure on $\mathbb{X}$ defined by $\delta_x(B) = \mathbb{1}_B(x)$, where $\mathbb{1}_B$ is the indicator function of $B \in \mathcal{X}$. If X is a random element of $\mathbb{X}$, then δ_X is a point process on $\mathbb{X}$. Suppose, more generally, that $X_1, \dots, X_m$ are independent random elements in $\mathbb{X}$ with distribution $\mathbb{Q}$. Then

$$\eta := \delta_{X_1} + \cdots + \delta_{X_m} \tag{1}$$

is a point process on $\mathbb{X}$. Because

$$\mathbb{P}(\eta(B) = k) = \binom{m}{k} \mathbb{Q}(B)^k (1 - \mathbb{Q}(B))^{m-k}, \quad k = 0, \dots, m,$$

η is referred to as *binomial process* with *sample size m* and *sampling distribution* $\mathbb{Q}$. Taking an infinite sequence $X_1, X_2, \dots$ of independent random variables with distribution $\mathbb{Q}$ and replacing in (1) the deterministic sample size m by an independent $\mathbb{Z}_+$-valued random variable κ (and interpreting an empty sum as null measure) yields a *mixed binomial process*. Of particular interest is the case where κ has a Poisson distribution with parameter $\lambda \geq 0$, see also (5) below. It is then easy to check that

$$\mathbb{E} \exp\Big[-\int u(x)\eta(\mathrm{d}x) \Big] = \exp\Big[-\int (1 - e^{-u(x)})\mu(\mathrm{d}x) \Big], \tag{2}$$

for any measurable function $u : \mathbb{X} \to [0, \infty)$, where $\mu := \lambda \mathbb{Q}$. It is convenient to write this as

$$\mathbb{E} \exp[-\eta(u)] = \exp\big[-\mu(1 - e^{-u}) \big], \tag{3}$$

where $\nu(u)$ denotes the integral of a measurable function u with respect to a measure ν. Clearly,

$$\mu(B) = \mathbb{E}\eta(B), \quad B \in \mathcal{X}, \tag{4}$$

so that μ is the *intensity measure* of η. The identity (3) or elementary probabilistic arguments show that η has *independent increments*, that is, the random variables $\eta(B_1), \dots, \eta(B_m)$ are stochastically independent whenever $B_1, \dots, B_m \in \mathcal{X}$ are pairwise disjoint. Moreover, $\eta(B)$ has a Poisson distribution with parameter $\mu(B)$, that is

$$\mathbb{P}(\eta(B) = k) = \frac{\mu(B)^k}{k!} \exp[-\mu(B)], \quad k \in \mathbb{Z}_+. \tag{5}$$

Let μ be a σ-finite measure on $\mathbb{X}$. A *Poisson process* with intensity measure μ is a point process η on $\mathbb{X}$ with independent increments such that (5) holds, where an expression of the form $\infty e^{-\infty}$ is interpreted as 0. It is easy to see that these two requirements determine the distribution $\mathbb{P}_\eta := \mathbb{P}(\eta \in \cdot)$ of a Poisson process η. We have seen above that a Poisson process exists for a finite measure μ. In the general case, it can be constructed as a countable sum of independent Poisson processes, see [12, 15, 18] for more details. Equation (3) remains valid. Another consequence of this construction is that η has the same distribution as

$$\eta = \sum_{n=1}^{\eta(\mathbb{X})} \delta_{X_n}, \tag{6}$$

where $X_1, X_2, \ldots$ are random elements in $\mathbb{X}$. A point process that can be (almost surely) represented in this form will be called *proper*. Any locally finite point process on a Borel subset of a complete separable metric space is proper. However, there are examples of Poisson processes which are not proper.

Let η be a Poisson process with intensity measure μ. A classical and extremely useful formula by Mecke [18] says that

$$\mathbb{E} \int h(\eta, x)\eta(\mathrm{d}x) = \mathbb{E} \int h(\eta + \delta_x, x)\mu(\mathrm{d}x) \tag{7}$$

for all measurable $h : \mathbf{N}_\sigma \times \mathbb{X} \to [0, \infty]$. One can use the mixed binomial representation to prove this result for finite Poisson processes. An equivalent formulation for a proper Poisson process is

$$\mathbb{E} \int h(\eta - \delta_x, x)\eta(\mathrm{d}x) = \mathbb{E} \int h(\eta, x)\mu(\mathrm{d}x) \tag{8}$$

for all measurable $h : \mathbf{N}_\sigma \times \mathbb{X} \to [0, \infty]$. Although $\eta - \delta_x$ is in general a signed measure, we can use (6) to see that

$$\int h(\eta - \delta_x, x)\eta(\mathrm{d}x) = \sum_i h\Big(\sum_{j \neq i} \delta_{X_j}, X_i\Big)$$

is almost surely well defined. Both (7) and (8) characterize the distribution of a Poisson process with given intensity measure μ.

Equation (7) admits a useful generalization involving multiple integration. To formulate this version we consider, for $m \in \mathbb{N}$, the m-th power $(\mathbb{X}^m, \mathscr{X}^m)$ of $(\mathbb{X}, \mathscr{X})$. Let η be a proper point process given by (6). We define another point process $\eta^{(m)}$ on $\mathbb{X}^m$ by

$$\eta^{(m)}(C) = \sum_{i_1,\ldots,i_m \leq \eta(\mathbb{X})}^{\neq} \mathbb{1}_C(X_{i_1}, \ldots, X_{i_m}), \quad C \in \mathscr{X}^m, \tag{9}$$

where the superscript $\neq$ indicates summation over m-tuples with pairwise different entries. (In the case $\eta(\mathbb{X}) = \infty$ this involves only integer-valued indices.) In the case $C = B^m$ for some $B \in \mathcal{X}$ we have that

$$\eta^{(m)}(B^m) = \eta(B)(\eta(B)-1)\cdots(\eta(B)-m+1).$$

Therefore $\eta^{(m)}$ is called m-th *factorial measure* of η. It can be readily checked that, for any $m \in \mathbb{N}$,

$$\eta^{(m+1)} = \int \Big[\int \mathbb{1}\{(x_1,\dots,x_{m+1}) \in \cdot\}\eta(\mathrm{d}x_{m+1}) - \sum_{j=1}^{m} \mathbb{1}\{(x_1,\dots,x_m,x_j) \in \cdot\}\Big]\eta^{(m)}(\mathrm{d}(x_1,\dots,x_m)), \tag{10}$$

where $\eta^{(1)} := \eta$. This suggests a recursive definition of the factorial measures of a general point process, without using a representation as a sum of Dirac measures. The next proposition confirms this idea.

Proposition 1 *Let η be a point process on $\mathbb{X}$. Then there is a uniquely determined sequence $\eta^{(m)}$, $m \in \mathbb{N}$, of symmetric point processes on $\mathbb{X}^m$ satisfying $\eta^{(1)} := \eta$ and the recursion* (10)*.*

The proof of Proposition 1 is given in the appendix and can be skipped without too much loss. It is enough to remember that $\eta^{(m)}$ can be defined by (9), whenever η is given by (6) and that any Poisson process has a proper version.

The multivariate version of (7) (see e.g. [15]) says that

$$\mathbb{E}\int h(\eta, x_1,\dots,x_m)\eta^{(m)}(\mathrm{d}(x_1,\dots,x_m)) = \mathbb{E}\int h(\eta+\delta_{x_1}+\cdots+\delta_{x_m}, x_1,\dots,x_m)\mu^m(\mathrm{d}(x_1,\dots,x_m)), \tag{11}$$

for all measurable $h : \mathbf{N}_\sigma \times \mathbb{X}^m \to [0,\infty]$. In particular the *factorial moment measures* of η are given by

$$\mathbb{E}\eta^{(m)} = \mu^m, \quad m \in \mathbb{N}. \tag{12}$$

Of course (11) remains true for a measurable $h : \mathbf{N}_\sigma \times \mathbb{X}^m \to \mathbb{R}$ provided that the right-hand side is finite when replacing h with $|h|$.

2 Fock Space Representation

In the remainder of this chapter we consider a Poisson process η on $\mathbb{X}$ with σ-finite intensity measure μ and distribution $\mathbb{P}_\eta$.

In this and later chapters the following *difference operators* (sometimes called *add-one cost operators*) will play a crucial role. For any $f \in \mathbf{F}(\mathbf{N}_\sigma)$ (the set of all measurable functions from $\mathbf{N}_\sigma$ to $\mathbb{R}$) and $x \in \mathbb{X}$ the function $D_x f \in \mathbf{F}(\mathbf{N}_\sigma)$ is defined by

$$D_x f(\chi) := f(\chi + \delta_x) - f(\chi), \quad \chi \in \mathbf{N}_\sigma. \tag{13}$$

Iterating this definition, for $n \geq 2$ and $(x_1, \ldots, x_n) \in \mathbb{X}^n$ we define a function $D^n_{x_1,\ldots,x_n} f \in \mathbf{F}(\mathbf{N}_\sigma)$ inductively by

$$D^n_{x_1,\ldots,x_n} f := D^1_{x_1} D^{n-1}_{x_2,\ldots,x_n} f, \tag{14}$$

where $D^1 := D$ and $D^0 f = f$. Note that

$$D^n_{x_1,\ldots,x_n} f(\chi) = \sum_{J \subset \{1,2,\ldots,n\}} (-1)^{n-|J|} f\Big(\chi + \sum_{j \in J} \delta_{x_j}\Big), \tag{15}$$

where $|J|$ denotes the number of elements of J. This shows that $D^n_{x_1,\ldots,x_n} f$ is symmetric in $x_1, \ldots, x_n$ and that $(x_1, \ldots, x_n, \chi) \mapsto D^n_{x_1,\ldots,x_n} f(\chi)$ is measurable. We define symmetric and measurable functions $T_n f$ on $\mathbb{X}^n$ by

$$T_n f(x_1, \ldots, x_n) := \mathbb{E} D^n_{x_1,\ldots,x_n} f(\eta), \tag{16}$$

and we set $T_0 f := \mathbb{E} f(\eta)$, whenever these expectations are defined. By $\langle \cdot, \cdot \rangle_n$ we denote the scalar product in $L^2(\mu^n)$ and by $\| \cdot \|_n$ the associated norm. Let $L^2_s(\mu^n)$ denote the symmetric functions in $L^2(\mu^n)$. Our aim is to prove that the linear mapping $f \mapsto (T_n(f))_{n \geq 0}$ is an isometry from $L^2(\mathbb{P}_\eta)$ into the *Fock space* given by the direct sum of the spaces $L^2_s(\mu^n)$, $n \geq 0$ (with L^2 norms scaled by $n!^{-1/2}$) and with $L^2_s(\mu^0)$ interpreted as $\mathbb{R}$. In Sect. 4 we will see that this mapping is surjective. The result (and its proof) is from [13] and can be seen as a crucial first step in the stochastic analysis on Poisson spaces.

Theorem 1 *Let $f, g \in L^2(\mathbb{P}_\eta)$. Then*

$$\mathbb{E} f(\eta) g(\eta) = \mathbb{E} f(\eta) \mathbb{E} g(\eta) + \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n, \tag{17}$$

where the series converges absolutely.

We will prepare the proof with some lemmas. Let $\mathscr{X}_0$ be the system of all measurable $B \in \mathscr{X}$ with $\mu(B) < \infty$. Let $\mathbf{F}_0$ be the space of all bounded and measurable functions $v : \mathbb{X} \to [0, \infty)$ vanishing outside some $B \in \mathscr{X}_0$. Let $\mathbf{G}$

denote the space of all (bounded and measurable) functions $g : \mathbf{N}_\sigma \to \mathbb{R}$ of the form

$$g(\chi) = a_1 e^{-\chi(v_1)} + \cdots + a_n e^{-\chi(v_n)}, \tag{18}$$

where $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{R}$ and $v_1, \ldots, v_n \in \mathbf{F}_0$.

Lemma 1 *Relation* (17) *holds for* $f, g \in \mathbf{G}$.

Proof By linearity it suffices to consider functions f and g of the form

$$f(\chi) = \exp[-\chi(v)], \quad g(\chi) = \exp[-\chi(w)]$$

for $v, w \in \mathbf{F}_0$. Then we have for $n \geq 1$ that

$$D^n f(\chi) = \exp[-\chi(v)](e^{-v} - 1)^{\otimes n},$$

where $(e^{-v} - 1)^{\otimes n}(x_1, \ldots, x_n) := \prod_{i=1}^n (e^{-v(x_i)} - 1)$. From (3) we obtain that

$$T_n f = \exp[-\mu(1 - e^{-v})](e^{-v} - 1)^{\otimes n}. \tag{19}$$

Since $v \in \mathbf{F}_0$ it follows that $T_n f \in L_s^2(\mu^n)$, $n \geq 0$. Using (3) again, we obtain that

$$\mathbb{E} f(\eta) g(\eta) = \exp[-\mu(1 - e^{-(v+w)})]. \tag{20}$$

On the other hand we have from (19) (putting $\mu^0(1) := 1$) that

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n \\
&\quad = \exp[-\mu(1 - e^{-v})] \exp[-\mu(1 - e^{-w})] \sum_{n=0}^{\infty} \frac{1}{n!} \mu^n(((e^{-v} - 1)(e^{-w} - 1))^{\otimes n}) \\
&\quad = \exp[-\mu(2 - e^{-v} - e^{-w})] \exp[\mu((e^{-v} - 1)(e^{-w} - 1))].
\end{aligned}$$

This equals the right-hand side of (20). □

To extend (17) to general $f, g \in L^2(\mathbb{P}_\eta)$ we need two further lemmas.

Lemma 2 *The set* $\mathbf{G}$ *is dense in* $L^2(\mathbb{P}_\eta)$.

Proof Let $\mathbf{W}$ be the space of all bounded measurable $g : \mathbf{N}_\sigma \to \mathbb{R}$ that can be approximated in $L^2(\mathbb{P}_\eta)$ by functions in $\mathbf{G}$. This space is closed under monotone and uniformly bounded convergence and also under uniform convergence. Moreover, it contains the constant functions. The space $\mathbf{G}$ is stable under multiplication and we denote by $\mathcal{N}'$ the smallest σ-field on $\mathbf{N}_\sigma$ such that $\chi \mapsto h(\chi)$ is measurable for all $h \in \mathbf{G}$. A functional version of the monotone class theorem (see e.g. Theorem I.21

in [1]) implies that **W** contains any bounded $\mathcal{N}'$-measurable g. On the other hand we have that

$$\chi(C) = \lim_{t \to 0+} t^{-1}(1 - e^{-t\chi(C)}), \quad \chi \in \mathbf{N}_\sigma,$$

for any $C \in \mathcal{X}$. Hence $\chi \mapsto \chi(C)$ is $\mathcal{N}'$-measurable whenever $C \in \mathcal{X}_0$. Since μ is σ-finite, for any $C \in \mathcal{X}$ there is a monotone sequence $C_k \in \mathcal{X}_0$, $k \in \mathbb{N}$, with union C, so that $\chi \mapsto \chi(C)$ is $\mathcal{N}'$-measurable. Hence $\mathcal{N}' = \mathcal{N}_\sigma$ and it follows that **W** contains all bounded measurable functions. But then **W** is clearly dense in $L^2(\mathbb{P}_\eta)$ and the proof of the lemma is complete. □

Lemma 3 *Suppose that $f, f^1, f^2, \ldots \in L^2(\mathbb{P}_\eta)$ satisfy $f^k \to f$ in $L^2(\mathbb{P}_\eta)$ as $k \to \infty$, and that $h : \mathbf{N}_\sigma \to [0, 1]$ is measurable. Let $n \in \mathbb{N}$, let $C \in \mathcal{X}_0$ and set $B := C^n$. Then*

$$\lim_{k\to\infty} \mathbb{E} \int_B |D^n_{x_1,\ldots,x_n} f(\eta) - D^n_{x_1,\ldots,x_n} f^k(\eta)| h(\eta) \mu^n(\mathrm{d}(x_1, \ldots, x_n)) = 0. \tag{21}$$

Proof By (15), the relation (21) is implied by the convergence

$$\lim_{k\to\infty} \mathbb{E} \int_B \Big| f\Big(\eta + \sum_{i=1}^m \delta_{x_i}\Big) - f^k\Big(\eta + \sum_{i=1}^m \delta_{x_i}\Big)\Big| h(\eta) \mu^n(\mathrm{d}(x_1, \ldots, x_n)) = 0 \tag{22}$$

for all $m \in \{0, \ldots, n\}$. For $m = 0$ this is obvious. Assume $m \in \{1, \ldots, n\}$. Then the integral in (22) equals

$$\begin{aligned}
&\mu(C)^{n-m} \mathbb{E} \int_{C^m} \Big| f\Big(\eta + \sum_{i=1}^m \delta_{x_i}\Big) - f^k\Big(\eta + \sum_{i=1}^m \delta_{x_i}\Big)\Big| h(\eta) \mu^m(\mathrm{d}(x_1, \ldots, x_m)) \\
&= \mu(C)^{n-m} \mathbb{E} \int_{C^m} |f(\eta) - f^k(\eta)| h\Big(\eta - \sum_{i=1}^n \delta_{x_i}\Big) \eta^{(m)}(\mathrm{d}(x_1, \ldots, x_m)) \\
&\le \mu(C)^{n-m} \mathbb{E} |f(\eta) - f^k(\eta)| \eta^{(m)}(C^m),
\end{aligned}$$

where we have used (11) to get the equality. By the Cauchy–Schwarz inequality the last expression is bounded above by

$$\mu(C)^{n-m} (\mathbb{E}(f(\eta) - f^k(\eta))^2)^{1/2} (\mathbb{E}(\eta^{(m)}(C^m))^2)^{1/2}.$$

Since the Poisson distribution has moments of all orders, we obtain (22) and hence the lemma. □

Proof of Theorem 1 By linearity and the polarization identity

$$4\langle u, v\rangle_n = \langle u+v, u+v\rangle_n - \langle u-v, u-v\rangle_n$$

it suffices to prove (17) for $f = g \in L^2(\mathbb{P}_\eta)$. By Lemma 2 there are $f^k \in \mathbf{G}$, $k \in \mathbb{N}$, satisfying $f^k \to f$ in $L^2(\mathbb{P}_\eta)$ as $k \to \infty$. By Lemma 1, Tf^k, $k \in \mathbb{N}$, is a Cauchy sequence in $\mathbf{H} := \mathbb{R} \oplus \oplus_{n=1}^{\infty} L_s^2(\mu^n)$. The direct sum of the scalar products $(n!)^{-1}\langle\cdot,\cdot\rangle_n$ makes $\mathbf{H}$ a Hilbert space. Let $\tilde{f} = (\tilde{f}_n) \in \mathbf{H}$ be the limit, that is

$$\lim_{k\to\infty} \sum_{n=0}^{\infty} \frac{1}{n!} \|T_n f^k - \tilde{f}_n\|_n^2 = 0. \tag{23}$$

Taking the limit in the identity $\mathbb{E}f^k(\eta)^2 = \langle Tf^k, Tf^k\rangle_{\mathbf{H}}$ yields $\mathbb{E}f(\eta)^2 = \langle \tilde{f}, \tilde{f}\rangle_{\mathbf{H}}$. Equation (23) implies that $\tilde{f}_0 = \mathbb{E}f(\eta) = T_0 f$. It remains to show that for any $n \geq 1$,

$$\tilde{f}_n = T_n f, \quad \mu^n\text{-a.e.} \tag{24}$$

Let $C \in \mathscr{X}_0$ and $B := C^n$. Let μ_B^n denote the restriction of the measure μ^n to B. By (23) $T_n f^k$ converges in $L^2(\mu_B^n)$ (and hence in $L^1(\mu_B^n)$) to $\tilde{f}_n$, while by the definition (16) of T_n, and the case $h \equiv 1$ of (22), $T_n f^k$ converges in $L^1(\mu_B^n)$ to $T_n f$. Hence these $L^1(\mathbb{P})$ limits must be the same almost everywhere, so that $\tilde{f}_n = T_n f$ μ^n-a.e. on B. Since μ is assumed σ-finite, this implies (24) and hence the theorem. □

3 Multiple Wiener–Itô Integrals

For $n \geq 1$ and $g \in L^1(\mu^n)$ we define (see [6, 7, 28, 29])

$$I_n(g) := \sum_{J \subset [n]} (-1)^{n-|J|} \iint g(x_1, \dots, x_n) \eta^{(|J|)}(\mathrm{d}x_J) \mu^{n-|J|}(\mathrm{d}x_{J^c}), \tag{25}$$

where $[n] := \{1, \dots, n\}$, $J^c := [n] \setminus J$ and $x_J := (x_j)_{j\in J}$. If $J = \emptyset$, then the inner integral on the right-hand side has to be interpreted as $\mu^n(g)$. (This is to say that $\eta^{(0)}(1) := 1$.) The multivariate Mecke equation (11) implies that all integrals in (25) are finite and that $\mathbb{E}I_n(g) = 0$.

Given functions $g_i : \mathbb{X} \to \mathbb{R}$ for $i = 1, \dots, n$, the *tensor product* $\otimes_{i=1}^n g_i$ is the function from $\mathbb{X}^n$ to $\mathbb{R}$ which maps each $(x_1, \dots, x_n)$ to $\prod_{i=1}^n g_i(x_i)$. When the functions $g_1, \dots, g_n$ are all the same function h, we write $h^{\otimes n}$ for this tensor product function. In this case the definition (25) simplifies to

$$I_n(h^{\otimes n}) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} \eta^{(k)}(h^{\otimes k})(\mu(h))^{n-k}. \tag{26}$$

Let Σ_n denote the set of all permutations of $[n]$, and for $g \in \mathbb{X}^n \to \mathbb{R}$ define the *symmetrization* $\tilde{g}$ of g by

$$\tilde{g}(x_1, \dots, x_n) := \frac{1}{n!} \sum_{\pi \in \Sigma_n} g(x_{\pi(1)}, \dots, x_{\pi(n)}). \tag{27}$$

The following *isometry properties* of the operators I_n are crucial. The proof is similar to the one of [16, Theorem 3.1] and is based on the product form (12) of the factorial moment measures and some combinatorial arguments. For more information on the intimate relationships between moments of Poisson integrals and the combinatorial properties of partitions we refer to [16, 21, 25, 28].

Lemma 4 *Let $g \in L^2(\mu^m)$ and $h \in L^2(\mu^n)$ for $m, n \geq 1$ and assume that $\{g \neq 0\} \subset B^m$ and $\{h \neq 0\} \subset B^n$ for some $B \in \mathcal{X}_0$. Then*

$$\mathbb{E} I_m(g) I_n(h) = \mathbb{1}\{m = n\} m! \langle \tilde{g}, \tilde{h} \rangle_m. \tag{28}$$

Proof We start with a combinatorial identity. Let $n \in \mathbb{N}$. A *subpartition* of $[n]$ is a (possibly empty) family σ of nonempty pairwise disjoint subsets of $[n]$. The cardinality of $\cup_{J \in \sigma} J$ is denoted by $\|\sigma\|$. For $u \in \mathbf{F}(\mathbb{X}^n)$ we define $u_\sigma : \mathbb{X}^{|\sigma|+n-\|\sigma\|} \to \mathbb{R}$ by identifying the arguments belonging to the same $J \in \sigma$. (The arguments $x_1, \dots, x_{|\sigma|+n-\|\sigma\|}$ have to be inserted in the order of occurrence.) Now we take $r, s \in \mathbb{Z}_+$ such that $r + s \geq 1$ and define $\Sigma_{r,s}$ as the set of all partitions of $\{1, \dots, r+s\}$ such that $|J \cap \{1, \dots, r\}| \leq 1$ and $|J \cap \{r+1, \dots, r+s\}| \leq 1$ for all $J \in \sigma$. Let $u \in \mathbf{F}(\mathbb{X}^{r+s})$. It is easy to see that

$$\begin{aligned} &\iint u(x_1, \dots, x_{r+s}) \eta^{(r)}(\mathrm{d}(x_1, \dots, x_r)) \eta^{(s)}(\mathrm{d}(x_{r+1}, \dots, x_{r+s})) \\ &\quad = \sum_{\sigma \in \Sigma_{r,s}} \int u_\sigma \, \mathrm{d}\eta^{(|\sigma|)}, \end{aligned} \tag{29}$$

provided that $\eta(\{u \neq 0\}) < \infty$. (In the case $r = 0$ the inner integral on the left-hand side is interpreted as 1.)

We next note that $g \in L^1(\mu^m)$ and $h \in L^1(\mu^n)$ and abbreviate $f := g \otimes h$. Let $k := m + n$, $J_1 := [m]$ and $J_2 := \{m+1, \dots, m+n\}$. The definition (25) and Fubini's theorem imply that

$$\begin{aligned} I_m(g) I_n(h) = \sum_{I \subset [k]} (-1)^{n-|I|} \iiint f(x_1, \dots, x_k) \\ \eta^{(|I \cap J_1|)}(\mathrm{d}x_{I \cap J_1}) \eta^{(|I \cap J_2|)}(\mathrm{d}x_{I \cap J_2}) \mu^{n-|I|}(\mathrm{d}x_{I^c}), \end{aligned} \tag{30}$$

where $I^c := [k] \setminus I$ and $x_J := (x_j)_{j \in J}$ for any $J \subset [k]$. We now take the expectation of (30) and use Fubini's theorem (justified by our integrability assumptions on g

and h). Thanks to (29) and (12) we can compute the expectation of the inner two integrals to obtain that

$$\mathbb{E} I_m(g) I_n(h) = \sum_{\sigma \in \Sigma^*_{m,n}} (-1)^{k-\|\sigma\|} \int f_\sigma \, \mathrm{d}\mu^{k-\|\sigma\|+|\sigma|}, \tag{31}$$

where $\Sigma^*_{m,n}$ is the set of all subpartitions σ of $[k]$ such that $|J \cap J_1| \leq 1$ and $|J \cap J_2| \leq 1$ for all $J \in \sigma$. Let $\Sigma^{*,2}_{m,n} \subset \Sigma^*_{m,n}$ be the set of all subpartitions of $[k]$ such that $|J| = 2$ for all $J \in \sigma$. For any $\pi \in \Sigma^{*,2}_{m,n}$ we let $\Sigma^*_{m,n}(\pi)$ denote the set of all $\sigma \in \Sigma^*_{m,n}$ satisfying $\pi \subset \sigma$. Note that $\pi \in \Sigma^*_{m,n}(\pi)$ and that for any $\sigma \in \Sigma^*_{m,n}$ there is a unique $\pi \in \Sigma^{*,2}_{m,n}$ such that $\sigma \in \Sigma^*_{m,n}(\pi)$. In this case

$$\int f_\sigma \mathrm{d}\mu^{k-\|\sigma\|+|\sigma|} = \int f_\pi \mathrm{d}\mu^{k-\|\pi\|},$$

so that (31) implies

$$\mathbb{E} I_m(g) I_n(h) = \sum_{\pi \in \Sigma^{*,2}_{m,n}} \int f_\pi \mathrm{d}\mu^{k-\|\pi\|} \sum_{\sigma \in \Sigma^*_{m,n}(\pi)} (-1)^{k-\|\sigma\|}. \tag{32}$$

The inner sum comes to zero, except in the case where $\|\pi\| = k$. Hence (32) vanishes unless $m = n$. In the latter case we have

$$\emptyset \mathbb{E} I_m(g) I_n(h) = \sum_{\pi \in \Sigma^{*,2}_{m,m} : |\pi| = m} \int f_\pi \, \mathrm{d}\mu^m = m! \langle \tilde{g}, \tilde{h} \rangle_m,$$

as asserted. □

Any $g \in L^2(\mu^m)$ is the L^2-limit of a sequence $g_k \in L^2(\mu^m)$ satisfying the assumptions of Lemma 4. For instance we may take $g_k := \mathbb{1}_{(B_k)^m} g$, where $\mu(B_k) < \infty$ and $B_k \uparrow \mathbb{X}$ as $k \to \infty$. Therefore the isometry (28) allows us to extend the linear operator I_m in a unique way to $L^2(\mu^m)$. It follows from the isometry that $I_m(g) = I_m(\tilde{g})$ for all $g \in L^2(\mu^m)$. Moreover, (28) remains true for arbitrary $g \in L^2(\mu^m)$ and $h \in L^2(\mu^n)$. It is convenient to set $I_0(c) := c$ for $c \in \mathbb{R}$. When $m \geq 1$, the random variable $I_m(g)$ is the (m-th order) *Wiener–Itô integral* of $g \in L^2(\mu^m)$ with respect to the *compensated Poisson process* $\hat{\eta} := \eta - \mu$. The reference to $\hat{\eta}$ comes from the explicit definition (25). We note that $\hat{\eta}(B)$ is only defined for $B \in \mathcal{X}_0$. In fact, $\{\hat{\eta}(B) : B \in \mathcal{X}_0\}$ is an *independent random measure* in the sense of [7].

4 The Wiener–Itô Chaos Expansion

A fundamental result of Itô [7] and Wiener [29] says that every square integrable function of the Poisson process η can be written as an infinite series of orthogonal stochastic integrals. Our aim is to prove the following explicit version of this *Wiener–Itô chaos expansion*. Recall definition (16).

Theorem 2 *Let $f \in L^2(\mathbb{P}_\eta)$. Then $T_n f \in L^2_s(\mu^n)$, $n \in \mathbb{N}$, and*

$$f(\eta) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(T_n f), \tag{33}$$

where the series converges in $L^2(\mathbb{P})$. Moreover, if $g_n \in L^2_s(\mu^n)$ for $n \in \mathbb{Z}_+$ satisfy $f(\eta) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(g_n)$ with convergence in $L^2(\mathbb{P})$, then $g_0 = \mathbb{E} f(\eta)$ and $g_n = T_n f$, μ^n-a.e. on $\mathbb{X}^n$, for all $n \in \mathbb{N}$.

For a homogeneous Poisson process on the real line, the explicit chaos expansion (33) was proved in [8]. The general case was formulated and proved in [13]. Stroock [27] has proved the counterpart of (33) for Brownian motion. Stroock's formula involves iterated Malliavin derivatives and requires stronger integrability assumptions on $f(\eta)$.

Theorem 2 and the isometry properties (28) of stochastic integrals show that the isometry $f \mapsto (T_n(f))_{n \geq 0}$ is in fact a bijection from $L^2(\mathbb{P}_\eta)$ onto the Fock space. The following lemma is the key for the proof.

Lemma 5 *Let $f(\chi) := e^{-\chi(v)}$, $\chi \in \mathbf{N}_\sigma(\mathbb{X})$, where $v : \mathbb{X} \to [0, \infty)$ is a measurable function vanishing outside a set $B \in \mathcal{X}$ with $\mu(B) < \infty$. Then (33) holds $\mathbb{P}$-a.s. and in $L^2(\mathbb{P})$.*

Proof By (3) and (19) the right-hand side of (33) equals the formal sum

$$I := \exp[-\mu(1 - e^{-v})] + \exp[-\mu(1 - e^{-v})] \sum_{n=1}^{\infty} \frac{1}{n!} I_n((e^{-v} - 1)^{\otimes n}). \tag{34}$$

Using the pathwise definition (25) we obtain that almost surely

$$\begin{aligned} I &= \exp[-\mu(1 - e^{-v})] \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \eta^{(k)}((e^{-v} - 1)^{\otimes k})(\mu(1 - e^{-v}))^{n-k} \\ &= \exp[-\mu(1 - e^{-v})] \sum_{k=0}^{\infty} \frac{1}{k!} \eta^{(k)}((e^{-v} - 1)^{\otimes k}) \sum_{n=k}^{\infty} \frac{1}{(n-k)!} (\mu(1 - e^{-v}))^{n-k} \\ &= \sum_{k=0}^{N} \frac{1}{k!} \eta^{(k)}((e^{-v} - 1)^{\otimes k}), \end{aligned} \tag{35}$$

where $N := \eta(B)$. Assume for the moment that η is proper and write $\delta_{X_1} + \cdots + \delta_{X_N}$ for the restriction of η to B. Then we have almost surely that

$$I = \sum_{J \subset \{1,\ldots,N\}} \prod_{i \in J} (e^{-v(X_i)} - 1) = \prod_{i=1}^{N} e^{-v(X_i)} = e^{-\eta(v)},$$

and hence (33) holds with almost sure convergence of the series. To demonstrate that convergence also holds in $L^2(\mathbb{P})$, let the partial sum $I(m)$ be given by the right-hand side of (34) with the series terminated at $n = m$. Then since $\mu(1 - e^{-v})$ is nonnegative and $|1 - e^{-v(y)}| \leq 1$ for all y, a similar argument to (35) yields

$$\begin{aligned}|I(m)| &\leq \sum_{k=0}^{\min(N,m)} \frac{1}{k!} |\eta^{(k)}((e^{-v} - 1)^{\otimes k})| \\ &\leq \sum_{k=0}^{N} \frac{N(N-1)\cdots(N-k+1)}{k!} = 2^N.\end{aligned}$$

Since 2^N has finite moments of all orders, by dominated convergence the series (34) (and hence (33)) converges in $L^2(\mathbb{P})$.

Since the convergence of the right-hand side of (34) as well as the almost sure identity $I = e^{-\eta(v)}$ remain true for any point process with the same distribution as η (that is, for any Poisson process with intensity measure μ), it was no restriction of generality to assume that η is proper. □

Proof of Theorem 2 Let $f \in L^2(\mathbb{P}_\eta)$ and define $T_n f$ for $n \in \mathbb{Z}_+$ by (16). By (28) and Theorem 1,

$$\sum_{n=0}^{\infty} \mathbb{E}\Big(\frac{1}{n!} I_n(T_n f)\Big)^2 = \sum_{n=0}^{\infty} \frac{1}{n!} \|T_n f\|_n^2 = \mathbb{E} f(\eta)^2 < \infty.$$

Hence the infinite series of orthogonal terms

$$S := \sum_{n=0}^{\infty} \frac{1}{n!} I_n(T_n f)$$

converges in $L^2(\mathbb{P})$. Let $h \in \mathbf{G}$, where $\mathbf{G}$ was defined at (18). By Lemma 5 and linearity of $I_n(\cdot)$ the sum $\sum_{n=0}^{\infty} \frac{1}{n!} I_n(T_n h)$ converges in $L^2(\mathbb{P})$ to $h(\eta)$. Using (28) followed by Theorem 1 yields

$$\mathbb{E}(h(\eta) - S)^2 = \sum_{n=0}^{\infty} \frac{1}{n!} \|T_n h - T_n f\|_n = \mathbb{E}(f(\eta) - h(\eta))^2.$$

Hence if $\mathbb{E}(f(\eta)-h(\eta))^2$ is small, then so is $\mathbb{E}(f(\eta)-S)^2$. Since **G** dense in $L^2(\mathbb{P}_\eta)$ by Lemma 2, it follows that $f(\eta)=S$ almost surely.

To prove the uniqueness, suppose that also $g_n \in L^2_s(\mu^n)$ for $n \in \mathbb{Z}_+$ are such that $\sum_{n=0}^\infty \frac{1}{n!} I_n(g_n)$ converges in $L^2(\mathbb{P})$ to $f(\eta)$. By taking expectations we must have $g_0 = \mathbb{E}f(\eta) = T_0 f$. For $n \geq 1$ and $h \in L^2_s(\mu^n)$, by (28) and (33) we have

$$\mathbb{E}f(\eta)I_n(h) = \mathbb{E}I_n(T_n f)I_n(h) = n!\langle T_n f, h\rangle_n$$

and similarly with $T_n f$ replaced by g_n, so that $\langle T_n f - g_n, h\rangle_n = 0$. Putting $h = T_n f - g_n$ gives $\|T_n f - g_n\|_n = 0$ for each n, completing the proof of the theorem.

□

5 Malliavin Operators

For any $p \geq 0$ we denote by L^p_η the space of all random variables $F \in L^p(\mathbb{P})$ such that $F = f(\eta)$ $\mathbb{P}$-almost surely, for some $f \in \mathbf{F}(\mathbf{N}_\sigma)$. Note that the space L^p_η is a subset of $L^p(\mathbb{P})$ while $L^p(\mathbb{P}_\eta)$ is the space of all measurable functions $f \in \mathbf{F}(\mathbf{N}_\sigma)$ satisfying $\int |f|^p \, d\mathbb{P}_\eta = \mathbb{E}|f(\eta)|^p < \infty$. The *representative* f of $F \in L^p(\mathbb{P})$ is is $\mathbb{P}_\eta$-a.e. uniquely defined element of $L^p(\mathbb{P}_\eta)$. For $x \in \mathbb{X}$ we can then define the random variable $D_x F := D_x f(\eta)$. More generally, we define $D^n_{x_1,\dots,x_n} F := D^n_{x_1,\dots,x_n} f(\eta)$ for any $n \in \mathbb{N}$ and $x_1,\dots,x_n \in \mathbb{X}$. The mapping $(\omega, x_1,\dots,x_n) \mapsto D^n_{x_1,\dots,x_n} F(\omega)$ is denoted by $D^n F$ (or by DF in the case $n=1$). The multivariate Mecke equation (11) easily implies that these definitions are $\mathbb{P} \otimes \mu$-a.e. independent of the choice of the representative.

By (33) any $F \in L^2_\eta$ can be written as

$$F = \mathbb{E}F + \sum_{n=1}^\infty I_n(f_n), \tag{36}$$

where $f_n := \frac{1}{n!}\mathbb{E}D^n F$. In particular we obtain from (28) (or directly from Theorem 1) that

$$\mathbb{E}F^2 = (\mathbb{E}F)^2 + \sum_{n=1}^\infty n!\|f_n\|_n^2. \tag{37}$$

We denote by $\operatorname{dom} D$ the set of all $F \in L^2_\eta$ satisfying

$$\sum_{n=1}^\infty nn!\|f_n\|_n^2 < \infty. \tag{38}$$

The following result is taken from [13] and generalizes Theorem 6.5 in [8] (see also Theorem 6.2 in [20]). It shows that under the assumption (38) the pathwise defined difference operator DF coincides with the *Malliavin derivative* of F. The space $\operatorname{dom} D$ is the *domain* of this operator.

Theorem 3 *Let $F \in L^2_\eta$ be given by (36). Then $DF \in L^2(\mathbb{P} \otimes \mu)$ iff $F \in \operatorname{dom} D$. In this case we have $\mathbb{P}$-a.s. and for μ-a.e. $x \in \mathbb{X}$ that*

$$D_x F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(x,\cdot)). \tag{39}$$

The proof of Theorem 3 requires some preparations. Since

$$\int \Big(\sum_{n=1}^{\infty} n n! \|f_n(x,\cdot)\|_{n-1}^2 \Big) \mu(\mathrm{d}x) = \sum_{n=1}^{\infty} n n! \int \|f_n\|_n^2,$$

(28) implies that the infinite series

$$D'_x F := \sum_{n=1}^{\infty} n I_{n-1} f_n(x,\cdot) \tag{40}$$

converges in $L^2(\mathbb{P})$ for μ-a.e. $x \in \mathbb{X}$ provided that $F \in \operatorname{dom} D$. By construction of the stochastic integrals we can assume that $(\omega, x) \mapsto (I_{n-1} f_n(x,\cdot))(\omega)$ is measurable for all $n \geq 1$. Therefore we can also assume that the mapping $D'F$ given by $(\omega, x) \mapsto D'_x F(\omega)$ is measurable. We have just seen that

$$\mathbb{E} \int (D'_x F)^2 \mu(\mathrm{d}x) = \sum_{n=1}^{\infty} n n! \int \|f_n\|_n^2, \quad F \in \operatorname{dom} D. \tag{41}$$

Next we introduce an operator acting on random functions that will turn out to be the *adjoint* of the difference operator D, see Theorem 4. For $p \geq 0$ let $L^p_\eta(\mathbb{P} \otimes \mu)$ denote the set of all $H \in L^p(\mathbb{P} \otimes \mu)$ satisfying $H(\omega, x) = h(\eta(\omega), x)$ for $\mathbb{P} \otimes \mu$-a.e. (ω, x) for some *representative* $h \in \mathbf{F}(\mathbf{N}_\sigma \times \mathbb{X})$. For such a H we have for μ-a.e. x that $H(x) := H(\cdot, x) \in L^2(\mathbb{P})$ and (by Theorem 2)

$$H(x) = \sum_{n=0}^{\infty} I_n(h_n(x,\cdot)), \quad \mathbb{P}\text{-a.s.}, \tag{42}$$

where $h_0(x) := \mathbb{E} H(x)$ and $h_n(x, x_1, \ldots, x_n) := \frac{1}{n!} \mathbb{E} D^n_{x_1,\ldots,x_n} H(x)$. We can then define the *Kabanov–Skorohod integral* [3, 10, 11, 26] of H, denoted $\delta(H)$, by

$$\delta(H) := \sum_{n=0}^{\infty} I_{n+1}(h_n), \tag{43}$$

which converges in $L^2(\mathbb{P})$ provided that

$$\sum_{n=0}^{\infty}(n+1)!\int \tilde{h}_n^2 \mathrm{d}\mu^{n+1} < \infty. \tag{44}$$

Here

$$\tilde{h}_n(x_1,\dots,x_{n+1}) := \frac{1}{(n+1)!}\sum_{i=1}^{n+1}\mathbb{E}D^n_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_{n+1}}H(x_i) \tag{45}$$

is the symmetrization of h_n. The set of all $H \in L^2_\eta(\mathbb{P}\otimes\mu)$ satisfying the latter assumption is the domain $\operatorname{dom}\delta$ of the operator δ.

We continue with a preliminary version of Theorem 4.

Proposition 2 *Let* $F \in \operatorname{dom} D$. *Let* $H \in L^2_\eta(\mathbb{P}\otimes\mu)$ *be given by* (42) *and assume that*

$$\sum_{n=0}^{\infty}(n+1)!\int h_n^2 \mathrm{d}\mu^{n+1} < \infty. \tag{46}$$

Then

$$\mathbb{E}\int (D'_x F)H(x)\mu(\mathrm{d}x) = \mathbb{E}F\delta(H). \tag{47}$$

Proof Minkowski inequality implies (44) and hence $H \in \operatorname{dom}\delta$. Using (40) and (42) together with (28), we obtain that

$$\mathbb{E}\int (D'_x F)H(x)\mu(\mathrm{d}x) = \int\left(\sum_{n=1}^{\infty} n!\langle f_n(x,\cdot), h_{n-1}(x,\cdot)\rangle_{n-1}\right)\mu(\mathrm{d}x),$$

where the use of Fubini's theorem is justified by (41), the assumption on H and the Cauchy–Schwarz inequality. Swapping the order of summation and integration (to be justified soon) we see that the last integral equals

$$\sum_{n=1}^{\infty} n!\langle f_n, h_{n-1}\rangle_n = \sum_{n=1}^{\infty} n!\langle f_n, \tilde{h}_{n-1}\rangle_n,$$

where we have used the fact that f_n is a symmetric function. By definition (43) and (28), the last series coincides with $\mathbb{E}F\delta(H)$. The above change of order is permitted since

$$\sum_{n=1}^{\infty} n! \int |\langle f_n(x,\cdot), h_{n-1}(x,\cdot)\rangle_{n-1}|\mu(\mathrm{d}x)$$
$$\leq \sum_{n=1}^{\infty} n! \int \|f_n(x,\cdot)\|_{n-1}\|h_{n-1}(x,\cdot)\|_{n-1}\mu(\mathrm{d}x)$$

and the latter series is finite in view of the Cauchy–Schwarz inequality, the finiteness of (36) and assumption (46). □

Proof of Theorem 3 We need to show that

$$DF = D'F, \quad \mathbb{P} \otimes \mu\text{-a.e.} \tag{48}$$

First consider the case with $f(\chi) = e^{-\chi(v)}$ with a measurable $v : \mathbb{X} \to [0,\infty)$ vanishing outside a set with finite μ-measure. Then $n!f_n = T_n f$ is given by (19). Given $n \in \mathbb{N}$,

$$n \cdot n! \int f_n^2 d\mu^n = \frac{1}{(n-1)!} \exp[2\mu(e^{-v}-1)](\mu((e^{-v}-1)^2))^n$$

which is summable in n, so (38) holds in this case. Also, in this case, $D_x f(\eta) = (e^{v(x)} - 1)f(\eta)$ by (13), while $f_n(\cdot,x) = (e^{-v(x)} - 1)n^{-1}f_{n-1}$ so that by (40),

$$D'_x f(\eta) = \sum_{n=1}^{\infty}(e^{-v(x)} - 1)I_{n-1}(f_{n-1}) = (e^{-v(x)} - 1)f(\eta)$$

where the last inequality is from Lemma 5 again. Thus (48) holds for f of this form. By linearity this extends to all elements of **G**.

Let us now consider the general case. Choose $g_k \in \mathbf{G}$, $k \in \mathbb{N}$, such that $G_k := g_k(\eta) \to F$ in $L^2(\mathbb{P})$ as $k \to \infty$, see Lemma 2. Let $H \in L^2_\eta(\mathbb{P}_\eta \otimes \mu)$ have the representative $h(\chi,x) := h'(\chi)\mathbb{1}_B(x)$, where h' is as in Lemma 5 and $B \in \mathscr{X}_0$. From Lemma 5 it is easy to see that (46) holds. Therefore we obtain from Proposition 2 and the linearity of the operator D' that

$$\mathbb{E}\int (D'_x F - D'_x G_k)H(x)\mu(\mathrm{d}x) = \mathbb{E}(f - G_k)\delta(H) \to 0 \quad \text{as } k \to \infty. \tag{49}$$

On the other hand,

$$\mathbb{E}\int (D_x F - D_x G_k)H(x)\mu(\mathrm{d}x) = \mathbb{E}\int_B (D_x f(\eta) - D_x g_k(\eta))h'(\eta)\mu(\mathrm{d}x),$$

and by the case $n = 1$ of Lemma 3, this tends to zero as $k \to \infty$. Since $D'_x g_k = D_x g_k$ a.s. for μ-a.e. x we obtain from (49) that

$$\mathbb{E}\int (D'_x f)h(\eta, x)\mu(\mathrm{d}x) = \mathbb{E}\int (D_x f(\eta))h(\eta, x)\mu(\mathrm{d}x). \tag{50}$$

By Lemma 2, the linear combinations of the functions h considered above are dense in $L^2(\mathbb{P}_\eta \otimes \mu)$, and by linearity (50) carries through to h in this dense class of functions too, so we may conclude that the assertion (48) holds.

It follows from (41) and (48) that $F \in \operatorname{dom} D$ implies $DF \in L^2_\eta(\mathbb{P} \otimes \mu)$. The other implication was noticed in [22, Lemma 3.1]. To prove it, we assume $DF \in L^2_\eta(\mathbb{P} \otimes \mu)$ and apply the Fock space representation (17) to $\mathbb{E}(D_x F)^2$ for μ-a.e. x. This gives

$$\begin{aligned}\int \mathbb{E}(D_x F)^2 \mu(\mathrm{d}x) &= \sum_{n=0}^{\infty} \frac{1}{n!} \iint (\mathbb{E}D^{n+1}_{x_1,\dots,x_n,x})^2 \mu^n(\mathrm{d}(x_1,\dots,x_n))\mu(\mathrm{d}x) \\ &= \sum_{n=0}^{\infty}(n+1)(n+1)!\|f_{n+1}\|^2_{n+1}\end{aligned}$$

and hence $F \in \operatorname{dom} D$. □

The following duality relation (also referred to as *partial integration*, or *integration by parts formula*) shows that the operator δ is the adjoint of the difference operator D. It is a special case of Proposition 4.2 in [20] applying to general Fock spaces.

Theorem 4 *Let* $F \in \operatorname{dom} D$ *and* $H \in \operatorname{dom} \delta$. *Then,*

$$\mathbb{E}\int (D_x F)H(x)\mu(\mathrm{d}x) = \mathbb{E}F\delta(H). \tag{51}$$

Proof We fix $F \in \operatorname{dom} D$. Theorem 3 and Proposition 2 imply that (51) holds if $H \in L^2_\eta(\mathbb{P} \otimes \mu)$ satisfies the stronger assumption (46). For any $m \in \mathbb{N}$ we define

$$H^{(m)}(x) := \sum_{n=0}^{m} I_n(h_n(x,\cdot)), \quad x \in \mathbb{X}. \tag{52}$$

Since $H^{(m)}$ satisfies (46) we obtain that

$$\mathbb{E}\int (D_x F)H^{(m)}(x)\mu(\mathrm{d}x) = \mathbb{E}F\delta(H^{(m)}). \tag{53}$$

From (28) we have

$$\int \mathbb{E}(H(x) - H^{(m)}(x))^2 \mu(\mathrm{d}x) = \int \Big(\sum_{n=m+1}^{\infty} n! \|h_n(x,\cdot)\|_n^2 \Big) \mu(\mathrm{d}x)$$
$$= \sum_{n=m+1}^{\infty} n! \|h_n\|_{n+1}^2 .$$

As $m \to \infty$ this tends to zero, since

$$\mathbb{E} \int H(x)^2 \mu(\mathrm{d}x) = \int \mathbb{E}(H(x))^2 \mu(\mathrm{d}x) = \sum_{n=0}^{\infty} n! \|h_n\|_{n+1}^2$$

is finite. It follows that the left-hand side of (53) tends to the left-hand side of (51).

To treat the right-hand side of (53) we note that

$$\mathbb{E}\delta(H - H^{(m)})^2 = \sum_{n=m+1}^{\infty} \mathbb{E}(I_{n+1}(h_n))^2 = \sum_{n=m+1}^{\infty} (n+1)! \|\tilde{h}_n\|_{n+1}^2 . \tag{54}$$

Since $H \in \operatorname{dom} \delta$ this tends to 0 as $m \to \infty$. Therefore $\mathbb{E}(\delta(H) - \delta(H^{(m)}))^2 \to 0$ and the right-hand side of (53) tends to the right-hand side of (51). □

We continue with a basic isometry property of the Kabanov–Skorohod integral. In the present generality the result is in [17]. A less general version is [24, Proposition 6.5.4].

Theorem 5 *Let $H \in L^2_\eta(\mathbb{P} \otimes \mu)$ be such that*

$$\mathbb{E} \iint (D_y H(x))^2 \mu(\mathrm{d}x) \mu(\mathrm{d}y) < \infty. \tag{55}$$

Then, $H \in \operatorname{dom} \delta$ and moreover

$$\mathbb{E}\delta(H)^2 = \mathbb{E} \int H(x)^2 \mu(\mathrm{d}x) + \mathbb{E} \iint D_y H(x) D_x H(y) \mu(\mathrm{d}x) \mu(\mathrm{d}y). \tag{56}$$

Proof Suppose that H is given as in (42). Assumption (55) implies that $H(x) \in \operatorname{dom} D$ for μ-a.e. $x \in \mathbb{X}$. We therefore deduce from Theorem 3 that

$$g(x,y) := D_y H(x) = \sum_{n=1}^{\infty} n I_{n-1}(h_n(x,y,\cdot))$$

$\mathbb{P}$-a.s. and for μ^2-a.e. $(x, y) \in \mathbb{X}^2$. Using assumption (55) together with the isometry properties (28), we infer that

$$\sum_{n=1}^{\infty} nn!\|\tilde{h}_n\|_{n+1}^2 \le \sum_{n=1}^{\infty} nn!\|h_n\|_{n+1}^2 = \mathbb{E}\iint (D_y H(x))^2 \mu(\mathrm{d}x)\mu(\mathrm{d}y) < \infty,$$

yielding that $H \in \operatorname{dom}\delta$.

Now we define $H^{(m)} \in \operatorname{dom}\delta$, $m \in \mathbb{N}$, by (52) and note that

$$\mathbb{E}\delta(H^{(m)})^2 = \sum_{n=0}^{m} \mathbb{E}I_{n+1}(\tilde{h}_n)^2 = \sum_{n=0}^{m} (n+1)!\|\tilde{h}_n\|_{n+1}^2.$$

Using the symmetry properties of the functions h_n it is easy to see that the latter sum equals

$$\sum_{n=0}^{m} n! \int h_n^2 d\mu^{n+1} + \sum_{n=1}^{m} nn! \iint h_n(x, y, z)h_n(y, x, z)\mu^2(\mathrm{d}(x, y))\mu^{n-1}(\mathrm{d}z). \tag{57}$$

On the other hand, we have from Theorem 3 that

$$D_y H^{(m)}(x) = \sum_{n=1}^{m} nI_{n-1}(h_n(x, y, \cdot)),$$

so that

$$\mathbb{E}\int H^{(m)}(x)^2\mu(\mathrm{d}x) + \mathbb{E}\iint D_y H^{(m)}(x) D_x H^{(m)}(y)\mu(\mathrm{d}x)\mu(\mathrm{d}y)$$

coincides with (57). Hence

$$\mathbb{E}\delta(H^{(m)})^2 = \mathbb{E}\int H^{(m)}(x)^2\mu(\mathrm{d}x) + \mathbb{E}\iint D_y H^{(m)}(x) D_x H^{(m)}(y)\mu(\mathrm{d}x)\mu(\mathrm{d}y). \tag{58}$$

These computations imply that $g_m(x, y) := D_y H^{(m)}(x)$ converges in $L^2(\mathbb{P} \otimes \mu^2)$ towards g. Similarly, $g'_m(x, y) := D_x H^{(m)}(y)$ converges towards $g'(x, y) := D_x g(y)$. Since we have seen in the proof of Theorem 4 that $H^{(m)} \to H$ in $L^2(\mathbb{P} \otimes \mu)$ as $m \to \infty$, we can now conclude that the right-hand side of (58) tends to the right-hand side of the asserted identity (56). On the other hand we know by (54) that $\mathbb{E}\delta(H^{(m)})^2 \to \mathbb{E}\delta(H)^2$ as $m \to \infty$. This concludes the proof. □

To explain the connection of (55) with classical stochastic analysis we assume for a moment that $\mathbb{X}$ is equipped with a transitive binary relation $<$ such that $\{(x, y) :$

$x < y\}$ is a measurable subset of $\mathbb{X}^2$ and such that $x < x$ fails for all $x \in \mathbb{X}$. We also assume that $<$ totally orders the points of $\mathbb{X}$ μ-a.e., that is

$$\mu([x]) = 0, \quad x \in \mathbb{X}, \tag{59}$$

where $[x] := \mathbb{X} \setminus \{y \in \mathbb{X} : y < x \text{ or } x < y\}$. For any $\chi \in \mathbf{N}_\sigma$ let χ_x denote the restriction of χ to $\{y \in \mathbb{X} : y < x\}$. Our final assumption on $<$ is that $(\chi, y) \mapsto \chi_y$ is measurable. A measurable function $h : \mathbf{N}_\sigma \times \mathbb{X} \to \mathbb{R}$ is called *predictable* if

$$h(\chi, x) = h(\chi_x, x), \quad (\chi, x) \in \mathbf{N}_\sigma \times \mathbb{X}. \tag{60}$$

A process $H \in L^0_\eta(\mathbb{P} \otimes \mu)$ is predictable if it has a predictable representative. In this case we have $\mathbb{P} \otimes \mu$-a.e. that $D_x H(y) = 0$ for $y < x$ and $D_y H(x) = 0$ for $x < y$. In view of (59) we obtain from (56) the classical Itô isometry

$$\mathbb{E}\delta(H)^2 = \mathbb{E} \int H(x)^2 \mu(\mathrm{d}x). \tag{61}$$

In fact, a combinatorial argument shows that any predictable $H \in L^2_\eta(\mathbb{P} \otimes \mu)$ is in the domain of δ. We refer to [14] for more detail and references to the literature.

We return to the general setting and derive a pathwise interpretation of the Kabanov–Skorohod integral. For $H \in L^1_\eta(\mathbb{P} \otimes \mu)$ with representative h we define

$$\delta'(H) := \int h(\eta - \delta_x, x)\eta(\mathrm{d}x) - \int h(\eta, x)\mu(\mathrm{d}x). \tag{62}$$

The Mecke equation (7) implies that this definition does $\mathbb{P}$-a.s. not depend on the choice of the representative. The next result (see [13]) shows that the Kabanov–Skorohod integral and the operator δ' coincide on the intersection of their domains. In the case of a diffuse intensity measure μ (and requiring some topological assumptions on $(\mathbb{X}, \mathscr{X})$) the result is implicit in [23].

Theorem 6 *Let $H \in L^1_\eta(\mathbb{P} \otimes \mu) \cap \operatorname{dom} \delta$. Then $\delta(H) = \delta'(H)$ $\mathbb{P}$-a.s.*

Proof Let H have representative h. The Mecke equation (7) shows the integrability $\mathbb{E} \int |h(\eta - \delta_x, x)| \eta(\mathrm{d}x) < \infty$ as well as

$$\mathbb{E} \int D_x f(\eta) h(\eta, x) \mu(\mathrm{d}x) = \mathbb{E} f(\eta) \delta'(H), \tag{63}$$

whenever $f : \mathbf{N}_\sigma \to \mathbb{R}$ is measurable and bounded. Therefore we obtain from (51) that $\mathbb{E}F\delta'(H) = \mathbb{E}F\delta(H)$ provided that $F := f(\eta) \in \operatorname{dom} D$. By Lemma 2 the space of such bounded random variables is dense in $L^2_\eta(\mathbb{P})$, so we may conclude that the assertion holds. □

Finally in this section we discuss the *Ornstein–Uhlenbeck generator* L whose domain $\operatorname{dom} L$ is given by the class of all $F \in L^2_\eta$ satisfying

$$\sum_{n=1}^{\infty} n^2 n! \|f_n\|_n^2 < \infty.$$

In this case one defines

$$LF := -\sum_{n=1}^{\infty} n I_n(f_n).$$

The (pseudo) *inverse* L^{-1} of L is given by

$$L^{-1}F := -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n). \tag{64}$$

The random variable $L^{-1}F$ is well defined for any $F \in L^2_\eta$. Moreover, (37) implies that $L^{-1}F \in \operatorname{dom} L$. The identity $LL^{-1}F = F$, however, holds only if $\mathbb{E}F = 0$.

The three Malliavin operators D, δ, and L are connected by a simple formula:

Proposition 3 *Let $F \in \operatorname{dom} L$. Then $F \in \operatorname{dom} D$, $DF \in \operatorname{dom} \delta$ and $\delta(DF) = -LF$.*

Proof The relationship $F \in \operatorname{dom} D$ is a direct consequence of (37). Let $H := DF$. By Theorem 3 we can apply (43) with $h_n := (n+1)f_{n+1}$. We have

$$\sum_{n=0}^{\infty} (n+1)! \|h_n\|_{n+1}^2 = \sum_{n=0}^{\infty} (n+1)!(n+1)^2 \|f_{n+1}\|_{n+1}^2$$

showing that $H \in \operatorname{dom} \delta$. Moreover, since $I_{n+1}(\tilde{h}_n) = I_{n+1}(h_n)$ it follows that

$$\delta(DF) = \sum_{n=0}^{\infty} I_{n+1}(h_n) = \sum_{n=0}^{\infty} (n+1) I_{n+1}(f_{n+1}) = -LF,$$

finishing the proof. □

The following pathwise representation shows that the Ornstein–Uhlenbeck generator can be interpreted as the generator of a free *birth and death process* on $\mathbb{X}$.

Proposition 4 *Let $F \in \operatorname{dom} L$ with representative f and assume $DF \in L^1_\eta(\mathbb{P} \otimes \mu)$. Then*

$$LF = \int (f(\eta - \delta_x) - f(\eta))\eta(\mathrm{d}x) + \int (f(\eta + \delta_x) - f(\eta))\mu(\mathrm{d}x). \tag{65}$$

Proof We use Proposition 3. Since $DF \in L^1_\eta(\mathbb{P} \otimes \mu)$ we can apply Theorem 6 and the result follows by a straightforward calculation. □

6 Products of Wiener–Itô Integrals

In this section we study the chaos expansion of $I_p(f)I_q(g)$, where $f \in L^2_s(\mu^p)$ and $g \in L^2_s(\mu^q)$ for $p, q \in \mathbb{N}$. We define for any $r \in \{0, \dots, p \wedge q\}$ (where $p \wedge q := \min\{p, q\}$) and $l \in [r]$ the *contraction* $f \star^l_r g : \mathbb{X}^{p+q-r-l} \to \mathbb{R}$ by

$$\begin{aligned} f \star^l_r g(x_1, \dots, x_{p+q-r-l}) \\ := \int f(y_1, \dots, y_l, x_1, \dots, x_{p-l}) \\ \times g(y_1, \dots, y_l, x_1, \dots, x_{r-l}, x_{p-l+1}, \dots, x_{p+q-r-l})\mu^l(\mathrm{d}(y_1, \dots, y_l)), \end{aligned} \tag{66}$$

whenever these integrals are well defined. In particular $f \star^0_0 g = f \otimes g$.

In the case $q = 1$ the next result was proved in [10]. The general case is treated in [28], though under less explicit integrability assumptions and for diffuse intensity measure. Our proof is quite different.

Proposition 5 *Let $f \in L^2_s(\mu^p)$ and $f \in L^2_s(\mu^q)$ and assume $f \star^l_r g \in L^2(\mu^{p+q-r-l})$ for all $r \in \{0, \dots, p \wedge q\}$ and $l \in \{0, \dots, r-1\}$. Then*

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q} r! \binom{p}{r}\binom{q}{r} \sum_{l=0}^{r} \binom{r}{l} I_{p+q-r-l}(f \star^l_r g), \quad \mathbb{P}\text{-a.s.} \tag{67}$$

Proof First note that the Cauchy–Schwarz inequality implies $f \star^r_r g \in L^2(\mu^{p+q-2r})$ for all $r \in \{0, \dots, p \wedge q\}$.

We prove (67) by induction on $p + q$. For $p \wedge q = 0$ the assertion is trivial. For the induction step we assume that $p \wedge q \geq 1$. If $F, G \in L^0_\eta$, then an easy calculation shows that

$$D_x(fG) = (D_xF)G + F(D_xG) + (D_xF)(D_xG) \tag{68}$$

holds $\mathbb{P}$-a.s. and for μ-a.e. $x \in \mathbb{X}$. Using this together with Theorem 3 we obtain that

$$D_x(I_p(f)I_q(g)) = pI_{p-1}(f_x)I_q(g) + qI_p(f)I_{q-1}(g_x) + pqI_{p-1}(f_x)I_{q-1}(g_x),$$

where $f_x := f(x,\cdot)$ and $g_x := g(x,\cdot)$. We aim at applying the induction hypothesis to each of the summands on the above right-hand side. To do so, we note that

$$(f_x \star_r^l g)(x_1,\dots,x_{p-1+q-r-l}) = f \star_r^l g(x_1,\dots,x_{p-1-l},x,x_{p-1-l+1}\dots,x_{p-1+q-r-l})$$

for all $r \in \{0,\dots,(p-1)\wedge q\}$ and $l \in \{0,\dots,r\}$ and

$$(f_x \star_r^l g_x)(x_1,\dots,x_{p-1+q-1-r-l}) = f \star_{r+1}^l g(x,x_1,\dots,x_{p-1+q-1-r-l})$$

for all $r \in \{0,\dots,(p-1)\wedge(q-1)\}$ and $l \in \{0,\dots,r\}$. Therefore the pairs (f_x, g), (f, g_x) and (f_x, g_x) satisfy for μ-a.e. $x \in \mathbb{X}$ the assumptions of the proposition. The induction hypothesis implies that

$$\begin{aligned} D_x(I_p(f)I_q(g)) = &\sum_{r=0}^{(p-1)\wedge q} r!p\binom{p-1}{r}\binom{q}{r}\sum_{l=0}^{r}\binom{r}{l} I_{p+q-1-r-l}(f_x \star_r^l g) \\ &+ \sum_{r=0}^{p\wedge(q-1)} r!q\binom{p}{r}\binom{q-1}{r}\sum_{l=0}^{r}\binom{r}{l} I_{p+q-1-r-l}(f \star_r^l g_x) \\ &+ \sum_{r=0}^{(p-1)\wedge(q-1)} r!pq\binom{p-1}{r}\binom{q-1}{r}\sum_{l=0}^{r}\binom{r}{l} I_{p+q-2-r-l}(f_x \star_r^l g_x). \end{aligned}$$

A straightforward but tedious calculation (left to the reader) implies that the above right-hand side equals

$$\sum_{r=0}^{p\wedge q} r!\binom{p}{r}\binom{q}{r}\sum_{l=0}^{r}\binom{r}{l}(p+q-r-l) I_{p+q-r-l-1}((\widetilde{f \star_r^l g})_x),$$

where the summand for $p+q-r-l=0$ has to be interpreted as 0. It follows that

$$D_x(I_p(f)I_q(g)) = D_x G, \quad \mathbb{P}\text{-a.s., } \mu\text{-a.e. } x \in \mathbb{X},$$

where G denotes the right-hand side of (67). On the other hand, the isometry properties (28) show that $\mathbb{E} I_p(f)I_q(g) = \mathbb{E} G$. Since $I_p(f)I_q(g) \in L^1_\eta(\mathbb{P})$ we can use the *Poincaré inequality* of Corollary 1 in Sect. 8 to conclude that

$$\mathbb{E}(I_p(f)I_q(g) - G)^2 = 0.$$

This finishes the induction and the result is proved. □

If $\{f \neq 0\} \subset B^p$ and $\{g \neq 0\} \subset B^q$ for some $B \in \mathscr{X}_0$ (as in Lemma 4), then (67) can be established by a direct computation, starting from (30). The argument is similar to the proof of Theorem 3.1 in [16]. The required integrability follows from the Cauchy–Schwarz inequality; see [16, Remark 3.1]. In the case $q \geq 2$ we do not see, however, how to get from this special to the general case via approximation.

Equation (67) can be further generalized so as to cover the case of a finite product of Wiener–Itô integrals. We again refer the reader to [28] as well as to [16, 21].

7 Mehler's Formula

In this section we assume that η is a proper Poisson process. We shall derive a pathwise representation of the inverse (64) of the Ornstein–Uhlenbeck generator.

To give the idea we define for $F \in L^2_\eta$ with representation (36)

$$T_s F := \mathbb{E}F + \sum_{n=1}^{\infty} e^{-ns} I_n(f_n), \quad s \geq 0. \tag{69}$$

The family $\{T_s : s \geq 0\}$ is the *Ornstein–Uhlenbeck semigroup*, see e.g. [24] and also [19] for the Gaussian case. If $F \in \mathrm{dom} L$ then it is easy to see that

$$\lim_{s \to 0} \frac{T_s F - F}{s} = L$$

in $L^2(\mathbb{P})$, see [19, Proposition 1.4.2] for the Gaussian case. Hence L can indeed be interpreted as the generator of the semigroup. But in the theory of Markov processes it is well known (see, e.g., the resolvent identities in [12, Theorem 19.4]) that

$$L^{-1}F = -\int_0^\infty T_s F ds, \tag{70}$$

at least under certain assumptions. What we therefore need is a pathwise representation of the operators T_s. Our guiding star is the birth and death representation in Proposition 4.

For $F \in L^1_\eta$ with representative f we define,

$$P_s F := \int \mathbb{E}[f(\eta^{(s)} + \chi) \mid \eta] \Pi_{(1-s)\mu}(\mathrm{d}\chi), \quad s \in [0, 1], \tag{71}$$

where $\eta^{(s)}$ is a *s-thinning* of η and where $\Pi_{\mu'}$ denotes the distribution of a Poisson process with intensity measure μ'. The thinning $\eta^{(s)}$ can be defined by removing the points in (6) independently of each other with probability $1 - s$; see [12, p. 226]. Since

$$\Pi_\mu = \mathbb{E}\Big[\int \mathbb{1}\{\eta^{(s)} + \chi \in \cdot\} \Pi_{(1-s)\mu}(\mathrm{d}\chi)\Big], \tag{72}$$

this definition does almost surely not depend on the representative of F. Equation (72) implies in particular that

$$\mathbb{E}P_sF = \mathbb{E}F, \quad F \in L^1_\eta, \tag{73}$$

while Jensen's inequality implies for any $p \geq 1$ the contractivity property

$$\mathbb{E}(P_sF)^p \leq \mathbb{E}|F|^p, \quad s \in [0,1],\ F \in L^2_\eta. \tag{74}$$

We prepare the main result of this section with the following crucial lemma from [17].

Lemma 6 *Let $F \in L^2_\eta$. Then, for all $n \in \mathbb{N}$ and $s \in [0,1]$,*

$$D^n_{x_1,\dots,x_n}(P_sF) = s^n P_s D^n_{x_1,\dots,x_n}F, \quad \mu^n\text{-a.e. } (x_1,\dots,x_n) \in \mathbb{X}^n,\ \mathbb{P}\text{-a.s.} \tag{75}$$

In particular

$$\mathbb{E}D^n_{x_1,\dots,x_n}P_sF = s^n \mathbb{E}D^n_{x_1,\dots,x_n}F, \quad \mu^n\text{-a.e.} (x_1,\dots,x_n) \in \mathbb{X}^n. \tag{76}$$

Proof To begin with, we assume that the representative of F is given by $f(\chi) = e^{-\chi(v)}$ for some $v : \mathbb{X} \to [0,\infty)$ such that $\mu(\{v > 0\}) < \infty$. By the definition of a s-thinning,

$$\mathbb{E}\big[e^{-\eta^{(s)}(v)} \mid \eta\big] = \exp\bigg[\int \log\big((1-s) + se^{-v(y)}\big)\eta(\mathrm{d}y)\bigg], \tag{77}$$

and it follows from Lemma 12.2 in [12] that

$$\int \exp(-\chi(v))\Pi_{(1-s)\mu}(\mathrm{d}\chi) = \exp\bigg[-(1-s)\int(1-e^{-v})\mathrm{d}\mu\bigg].$$

Hence, the definition (71) of the operator P_s implies that the following function f_s is a representative of P_sF:

$$f_s(\chi) := \exp\bigg[-(1-s)\int\big(1-e^{-v}\big)\mathrm{d}\mu\bigg]\exp\bigg[\int \log\big((1-s)+se^{-v(y)}\big)\chi(\mathrm{d}y)\bigg].$$

Therefore we obtain for any $x \in \mathbb{X}$ that

$$D_xP_sF = f_s(\eta+\delta_x) - f_s(\eta) = s\big(e^{-v(x)}-1\big)f_s(\eta) = s\big(e^{-v(x)}-1\big)P_sF.$$

This identity can be iterated to yield for all $n \in \mathbb{N}$ and all $(x_1,\dots,x_n) \in \mathbb{X}^n$ that

$$D^n_{x_1,\dots,x_n}P_sF = s^n \prod_{i=1}^n \big(e^{-v(x_i)}-1\big)P_sF.$$

On the other hand we have $\mathbb{P}$-a.s. that

$$P_s D^n_{x_1,\dots,x_n} F = P_s \prod_{i=1}^{n} \big(e^{-v(x_i)} - 1\big) F = \prod_{i=1}^{n} \big(e^{-v(x_i)} - 1\big) P_s F,$$

so that (75) holds for Poisson functionals of the given form.

By linearity, (75) extends to all F with a representative in the set $\mathbf{G}$ of all linear combinations of functions f as above. There are $f_k \in \mathbf{G}$, $k \in \mathbb{N}$, satisfying $F_k := f_k(\eta) \to F = f(\eta)$ in $L^2(\mathbb{P})$ as $k \to \infty$, where f is a representative of F (see [13, Lemma 2.1]). Therefore we obtain from the contractivity property (74) that

$$\mathbb{E}[(P_s F_k - P_s F)^2] = \mathbb{E}[(P_s(f_k - F))^2] \le \mathbb{E}[(f_k - F)^2] \to 0,$$

as $k \to \infty$. Taking $B \in \mathcal{X}$ with $\mu(B) < \infty$, it therefore follows from [13, Lemma 2.3] that

$$\mathbb{E} \int_{B^n} |D^n_{x_1,\dots,x_n} P_s F_k - D^n_{x_1,\dots,x_n} P_s F| \mu(\mathrm{d}(x_1,\dots,x_n)) \to 0,$$

as $k \to \infty$. On the other hand we obtain from the Fock space representation (17) that $\mathbb{E}|D^n_{x_1,\dots,x_n} F| < \infty$ for μ^n-a.e. $(x_1,\dots,x_n) \in \mathbb{X}^n$, so that linearity of P_s and (74) imply

$$\mathbb{E} \int_{B^n} |P_s D^n_{x_1,\dots,x_n} F_k - P_s D^n_{x_1,\dots,x_n} F| \mu(\mathrm{d}(x_1,\dots,x_n))$$
$$\le \int_{B^n} \mathbb{E}|D^n_{x_1,\dots,x_n} (f_k - F)| \mu(\mathrm{d}(x_1,\dots,x_n)).$$

Again, this latter integral tends to 0 as $k \to \infty$. Since (75) holds for any F_k we obtain that (75) holds $\mathbb{P} \otimes (\mu_B)^n$-a.e., and hence also $\mathbb{P} \otimes \mu^n$-a.e.

Taking the expectation in (75) and using (73) proves (76). □

The following theorem from [17] achieves the desired pathwise representation of the inverse Ornstein–Uhlenbeck operator.

Theorem 7 *Let $F \in L^2_\eta$. If $\mathbb{E}F = 0$ then we have $\mathbb{P}$-a.s. that*

$$L^{-1} F = -\int_0^1 s^{-1} P_s F \mathrm{d}s. \tag{78}$$

Proof Assume that F is given as in (36). Applying (36) to P_sF and using (76) yields

$$P_sF = \mathbb{E}F + \sum_{n=1}^{\infty} s^n I_n(f_n), \quad \mathbb{P}\text{-a.s., } s \in [0,1]. \tag{79}$$

Furthermore,

$$-\sum_{n=1}^{m} \frac{1}{n} I_n(f_n) = -\int_0^1 s^{-1} \sum_{n=1}^{m} s^n I_n(f_n) \mathrm{d}s, \quad m \geq 1.$$

Assume now that $\mathbb{E}F = 0$. In view of (64) we need to show that the above right-hand side converges in $L^2(\mathbb{P})$, as $m \to \infty$, to the right-hand side of (78). Taking into account (79) we hence have to show that

$$R_m := \int_0^1 s^{-1}\Big(P_sF - \sum_{n=1}^{m} s^n I_n(f_n)\Big)\mathrm{d}s = \int_0^1 s^{-1}\Big(\sum_{n=m+1}^{\infty} s^n I_n(f_n)\Big)\mathrm{d}s$$

converges in $L^2(\mathbb{P})$ to zero. Using that $\mathbb{E}I_n(f_n)I_m(f_m) = \mathbb{1}\{m = n\}n!\|f_n\|_n^2$ we obtain

$$\mathbb{E}R_m^2 \leq \int_0^1 s^{-2}\mathbb{E}\Big(\sum_{n=m+1}^{\infty} s^n I_n(f_n)\Big)^2 \mathrm{d}s = \sum_{n=m+1}^{\infty} n!\|f_n\|_n^2 \int_0^1 s^{2n-2}\mathrm{d}s$$

which tends to zero as $m \to \infty$. □

Equation (79) implies *Mehler's formula*

$$P_{e^{-s}}F = \mathbb{E}F + \sum_{n=1}^{\infty} e^{-ns} I_n(f_n), \quad \mathbb{P}\text{-a.s., } s \geq 0, \tag{80}$$

which was proved in [24] for the special case of a finite Poisson process with a diffuse intensity measure. Originally this formula was first established in a Gaussian setting, see, e.g., [19]. The family $\{P_{e^{-s}} : s \geq 0\}$ of operators describes a special example of *Glauber dynamics*. Using (80) in (78) gives the identity (69).

8 Covariance Identities

The fundamental Fock space isometry (17) can be rewritten in several other disguises. We give here two examples, starting with a covariance identity from [5] involving the operators P_s.

Theorem 8 *Assume that η is a proper Poisson process. Then, for any $F, G \in \operatorname{dom} D$,*

$$\mathbb{E}FG = \mathbb{E}F\,\mathbb{E}G + \mathbb{E}\int\int_0^1 (D_xF)(P_tD_xG)\mathrm{d}t\mu(\mathrm{d}x). \tag{81}$$

Proof The Cauchy–Schwarz inequality and the contractivity property (74) imply that

$$\left(\mathbb{E}\int\int_0^1 |D_xF||P_sD_xG|\mathrm{d}s\mu(\mathrm{d}x)\right)^2 \le \mathbb{E}\int(D_xF)^2\mu(\mathrm{d}x)\mathbb{E}\int(D_xG)^2\mu(\mathrm{d}x)$$

which is finite due to Theorem 3. Therefore we can use Fubini's theorem and (75) to obtain that the right-hand side of (81) equals

$$\mathbb{E}F\,\mathbb{E}G + \int\int_0^1 s^{-1}\mathbb{E}(D_xF)(D_xP_sG)\mathrm{d}s\mu(\mathrm{d}x). \tag{82}$$

For $s \in [0,1]$ and μ-a.e. $x \in \mathbb{X}$ we can apply the Fock space isometry Theorem 1 to D_xF and D_xP_sG. Taking into account Lemma 6, (73) and applying Fubini again (to be justified below) yields that the second summand in (82) equals

$$\begin{aligned}
&\int\int_0^1 s^{-1}\mathbb{E}D_xF\,\mathbb{E}D_xP_sG\,\mathrm{d}s\mu(\mathrm{d}x)\\
&\quad+\sum_{n=1}^{\infty}\frac{1}{n!}\iint\int_0^1 s^{-1}\mathbb{E}D^{n+1}_{x_1,\dots,x_n,x}F\,\mathbb{E}D^{n+1}_{x_1,\dots,x_n,x}P_sG\,\mathrm{d}s\mu^n(\mathrm{d}(x_1,\dots,x_n))\mu(\mathrm{d}x)\\
&=\int\mathbb{E}D_xF\,\mathbb{E}D_xG\mu(\mathrm{d}x)\\
&\quad+\sum_{n=1}^{\infty}\frac{1}{n!}\iint\int_0^1 s^n\mathbb{E}D^{n+1}_{x_1,\dots,x_n,x}F\,\mathbb{E}D^{n+1}_{x_1,\dots,x_n,x}G\,\mathrm{d}s\mu^n(\mathrm{d}(x_1,\dots,x_n))\mu(\mathrm{d}x)\\
&=\sum_{m=1}^{\infty}\frac{1}{m!}\int\mathbb{E}D^m_{x_1,\dots,x_m}F\,\mathbb{E}D^m_{x_1,\dots,x_m}G\,\mu^m(\mathrm{d}(x_1,\dots,x_m)).
\end{aligned}$$

Inserting this into (82) and applying Theorem 1 yield the asserted formula (81). The use of Fubini's theorem is justified by Theorem 1 for $f = g$ and the Cauchy–Schwarz inequality. □

The integrability assumptions of Theorem 8 can be reduced to mere square integrability when using a symmetric formulation. Under the assumptions of Theorem 8 the following result was proved in [4, 5]. An even more general version is [13, Theorem 1.5].

Theorem 9 *Assume that η is a proper Poisson process. Then, for any $F \in L^2_\eta$,*

$$\mathbb{E}\int\int_0^1 (\mathbb{E}[D_xF \mid \eta^{(t)}])^2 \mathrm{d}t\mu(\mathrm{d}x) < \infty, \tag{83}$$

and for any $F, G \in L^2_\eta$,

$$\mathbb{E}FG = \mathbb{E}F\,\mathbb{E}G + \mathbb{E}\int\int_0^1 \mathbb{E}[D_xF \mid \eta^{(t)}]\mathbb{E}[D_xG \mid \eta^{(t)}]\mathrm{d}t\mu(\mathrm{d}x). \tag{84}$$

Proof It is well known (and not hard to prove) that $\eta^{(t)}$ and $\eta - \eta^{(t)}$ are independent Poisson processes with intensity measures $t\mu$ and $(1-t)\mu$, respectively. Therefore we have for $F \in L^2_\eta$ with representative f that

$$\mathbb{E}[D_xF|\eta_t] = \int D_xf(\eta^{(t)} + \chi)\Pi_{(1-t)\mu}(\mathrm{d}\chi) \tag{85}$$

holds almost surely. It is easy to see that the right-hand side of (85) is a measurable function of (the suppressed) $\omega \in \Omega$, $x \in \mathbb{X}$, and $t \in [0,1]$.

Now we take $F, G \in L^2_\eta$ with representatives f and g. Let us first assume that $DF, DG \in L^2(\mathbb{P} \otimes \mu)$. Then (83) follows from the (conditional) Jensen inequality while (85) implies for all $t \in [0,1]$ and $x \in \mathbb{X}$, that

$$\begin{aligned}\mathbb{E}(D_xF)(P_tD_xG) &= \mathbb{E}D_xF\int D_xg(\eta^{(t)} + \mu)\Pi_{(1-t)\mu}(\mathrm{d}\mu)\\ &= \mathbb{E}\mathbb{E}[D_xF\,\mathbb{E}[D_xG \mid \eta^{(t)}]] = \mathbb{E}\mathbb{E}[D_xF \mid \eta^{(t)}]\mathbb{E}[D_xG \mid \eta^{(t)}].\end{aligned}$$

Therefore (84) is just another version of (81).

In this second step of the proof we consider general $F, G \in L^2_\eta$. Let $F_k \in L^2_\eta$, $k \in \mathbb{N}$, be a sequence such that $DF_k \in L^2(\mathbb{P} \otimes \mu)$ and $\mathbb{E}(f - F_k)^2 \to 0$ as $k \to \infty$. We have just proved that

$$\mathrm{Var}[F_k - F^l] = \mathbb{E}\int(\mathbb{E}[D_xF_k \mid \eta^{(t)}] - \mathbb{E}[D_xF^l \mid \eta^{(t)}])^2\mu^*(\mathrm{d}(x,t)), \quad k, l \in \mathbb{N},$$

where μ^* is the product of μ and Lebesgue measure on $[0,1]$. Since $L^2(\mathbb{P}\otimes\mu^*)$ is complete, there is an $h\in L^2(\mathbb{P}\otimes\mu^*)$ satisfying

$$\lim_{k\to\infty}\mathbb{E}\int (h(x,t)-\mathbb{E}[D_xF_k\mid\eta^{(t)}])^2\mu^*(\mathrm{d}(x,t))=0. \tag{86}$$

On the other hand it follows from Lemma 3 that for any $C\in\mathcal{X}_0$

$$\int_{C\times[0,1]}\mathbb{E}\big|\mathbb{E}[D_xF_k\mid\eta^{(t)}]-\mathbb{E}[D_xF\mid\eta^{(t)}]\big|\mu^*(\mathrm{d}(x,t))$$
$$\le\int_{C\times[0,1]}\mathbb{E}|D_xF_k-D_xF|\mu^*(\mathrm{d}(x,t))\to 0$$

as $k\to\infty$. Comparing this with (86) shows that $h(\omega,x,t)=\mathbb{E}[D_xF\mid\eta^{(t)}](\omega)$ for $\mathbb{P}\otimes\mu^*$-a.e. $(\omega,x,t)\in\Omega\times C\times[0,1]$ and hence also for $\mathbb{P}\otimes\mu^*$-a.e. $(\omega,x,t)\in\Omega\times\mathbb{X}\times[0,1]$. Therefore the fact that $h\in L^2(\mathbb{P}\otimes\mu^*)$ implies (84). Now let G_k, $k\in\mathbb{N}$, be a sequence approximating G. Then Eq. (84) holds with (f_k,G_k) instead of (f,G). But the second summand is just a scalar product in $L^2(\mathbb{P}\otimes\mu^*)$. Taking the limit as $k\to\infty$ and using the L^2-convergence proved above yield the general result. □

A quick consequence of the previous theorem is the *Poincaré inequality* for Poisson processes. The following general version is taken from [30]. A more direct approach can be based on the Fock space representation in Theorem 1, see [13].

Theorem 10 *For any $F\in L^2_\eta$,*

$$\operatorname{Var}F\le\mathbb{E}\int (D_xF)^2\mu(\mathrm{d}x). \tag{87}$$

Proof It is no restriction of generality to assume that η is proper. Take $F=G$ in (84) and apply Jensen's inequality. □

The following extension of (87) (taken from [17]) has been used in the proof of Proposition 5.

Corollary 1 *For $F\in L^1_\eta$,*

$$\mathbb{E}F^2\le(\mathbb{E}F)^2+\mathbb{E}\int (D_xF)^2\mu(\mathrm{d}x). \tag{88}$$

Proof For $s>0$ we define

$$F_s=\mathbb{1}\{F>s\}s+\mathbb{1}\{-s\le F\le s\}F-\mathbb{1}\{F<-s\}s.$$

By definition of F_s we have $F_s \in L^2_\eta$ and $|D_xF_s| \le |D_xF|$ for μ-a.e. $x \in \mathbb{X}$. Together with the Poincaré inequality (87) we obtain that

$$\mathbb{E}F_s^2 \le (\mathbb{E}F_s)^2 + \mathbb{E}\int (D_xF_s)^2\mu(\mathrm{d}x) \le (\mathbb{E}F_s)^2 + \mathbb{E}\int (D_xF)^2\mu(\mathrm{d}x).$$

By the monotone convergence theorem and the dominated convergence theorem, respectively, we have that $\mathbb{E}F_s^2 \to \mathbb{E}F^2$ and $\mathbb{E}F_s \to \mathbb{E}F$ as $s \to \infty$. Hence letting $s \to \infty$ in the previous inequality yields the assertion. □

As a second application of Theorem 9 we obtain the *Harris-FKG inequality* for Poisson processes, derived in [9]. Given $B \in \mathcal{X}$, a function $f \in \mathbf{F}(\mathbf{N}_\sigma)$ is *increasing on* B if $f(\chi + \delta_x) \ge f(\chi)$ for all $\chi \in \mathbf{N}_\sigma$ and all $x \in B$. It is *decreasing on* B if $(-f)$ is increasing on B.

Theorem 11 *Suppose* $B \in \mathcal{X}$. *Let* $f, g \in L^2(\mathbb{P}_\eta)$ *be increasing on* B *and decreasing on* $\mathbb{X} \setminus B$. *Then*

$$\mathbb{E}f(\eta)g(\eta) \ge \mathbb{E}f(\eta)\,\mathbb{E}g(\eta). \tag{89}$$

It was noticed in [30] that the correlation inequality (89) (also referred to as *association*) is a direct consequence of a covariance identity.

Acknowledgements The proof of Proposition 5 is joint work with Matthias Schulte.

Appendix

In this appendix we prove Proposition 1. If $\chi \in \mathbf{N}$ is given by

$$\chi = \sum_{j=1}^{k} \delta_{x_j} \tag{90}$$

for some $k \in \mathbb{N}_0 \cup \{\infty\}$ and some points $x_1, x_2, \ldots \in \mathbb{X}$ (which are not assumed to be distinct) we define, for $m \in \mathbb{N}$, the factorial measure $\chi^{(m)} \in \mathbf{N}(\mathbb{X}^m)$ by

$$\chi^{(m)}(C) = \sum_{i_1,\ldots,i_m \le k}^{\ne} \mathbb{1}\{(x_{i_1}, \ldots, x_{i_m}) \in C\}, \quad C \in \mathcal{X}^m. \tag{91}$$

These measures satisfy the recursion

$$\chi^{(m+1)} = \int \Big[\int \mathbb{1}\{(x_1, \dots, x_{m+1}) \in \cdot\}\, \chi(\mathrm{d}x_{m+1}) \\ - \sum_{j=1}^{m} \mathbb{1}\{(x_1, \dots, x_m, x_j) \in \cdot\} \Big] \chi^{(m)}(\mathrm{d}(x_1, \dots, x_m)). \tag{92}$$

Let $\mathbf{N}_{<\infty}$ denote the set of all $\chi \in \mathbf{N}$ with $\chi(\mathbb{X}) < \infty$. For $\chi \in \mathbf{N}_{<\infty}$ the recursion (92) is solved by

$$\chi^{(m)} = \int \cdots \int \mathbb{1}\{(x_1, \dots, x_m) \in \cdot\} \Big(\chi - \sum_{j=1}^{m-1} \delta_{x_j}\Big)(\mathrm{d}x_m) \cdots \chi(\mathrm{d}x_1), \tag{93}$$

where the integrations are with respect to finite signed measures. Note that $\chi^{(m)}$ is a signed measure such that $\chi^{(m)}(C) \in \mathbb{Z}$ for all $C \in \mathcal{X}^m$. At this stage it might not be obvious that $\chi^{(m)}(C) \geq 0$. If, however, χ is given by (90) with $k \in \mathbb{N}$, then (93) coincides with (91). Hence $\chi^{(m)}$ is a measure in this case. For any $\chi \in \mathbf{N}_{<\infty}$ we denote by $\chi^{(m)}$ the signed measure (93). This is in accordance with the recursion (92). The next lemma shows that $\chi^{(m)}$ is a measure.

Lemma 7 *Let $\chi \in \mathbf{N}_{<\infty}$ and $m \in \mathbb{N}$. Then $\chi^{(m)}(C) \geq 0$ for all $C \in \mathcal{X}^m$.*

Proof Let $B_1, \dots, B_m \in \mathcal{X}$ and let Π_m denote the set of partitions of $[m]$. The definition (93) implies that

$$\chi^{(m)}(B_1 \times \cdots \times B_m) = \sum_{\pi \in \Pi_m} c_\pi \prod_{J \in \pi} \chi(\cap_{i \in J} B_i), \tag{94}$$

where the coefficients $c_\pi \in \mathbb{R}$ do not depend on $B_1, \dots, B_m$ and χ. For instance

$$\chi^{(3)}(B_1 \times B_2 \times B_3) = \chi(B_1)\chi(B_2)\chi(B_3) - \chi(B_1)\chi(B_2 \cap B_3) \\ - \chi(B_2)\chi(B_1 \cap B_3) - \chi(B_3)\chi(B_1 \cap B_2) + 2\chi(B_1 \cap B_2 \cap B_3).$$

It follows that the left-hand side of (94) is determined by the values of χ on the algebra generated by $B_1, \dots, B_m$. The atoms of this algebra are all nonempty sets of the form

$$B = B_1^{i_1} \cap \cdots \cap B_m^{i_m},$$

where $i_1, \dots, i_m \in \{0, 1\}$ and, for $B \subset \mathbb{X}$, $B^1 := B$ and $B^0 := \mathbb{X} \setminus B$. Let $\mathcal{A}$ denote the set of all these atoms. For $B \in \mathcal{A}$ we take $x \in B$ and let $\chi_B := \chi(B)\delta_x$. Then the measure

$$\chi' := \sum_{B \in \mathcal{A}} \chi_B$$

is a finite sum of Dirac measures and (94) implies that

$$(\chi')^{(m)}(B_1 \times \cdots \times B_m) = \chi^{(m)}(B_1 \times \cdots \times B_m).$$

Therefore it follows from (91) (applied to χ') that $\chi^{(m)}(B_1 \times \cdots \times B_m) \geq 0$.

Let $\mathcal{A}_m$ be the system of all finite and disjoint unions of sets $B_1 \times \cdots \times B_m$. This is an algebra; see Proposition 3.2.3 in [2]. From the first step of the proof and additivity of $\chi^{(m)}$ we obtain that $\chi^{(m)}(A) \geq 0$ holds for all $A \in \mathcal{A}_m$. The system $\mathcal{M}$ of all sets $A \in \mathcal{X}^m$ with the property $\chi^{(m)}(A) \geq 0$ is monotone. Hence a monotone class theorem (see e.g. Theorem 4.4.2 in [2]) implies that $\mathcal{M} = \mathcal{X}^m$. Therefore $\chi^{(m)}$ is nonnegative. □

Lemma 8 *Let $\chi, \nu \in \mathbf{N}_{<\infty}$ and assume that $\chi \leq \nu$. Let $m \in \mathbb{N}$. Then $\chi^{(m)} \leq \nu^{(m)}$.*

Proof By a monotone class argument it suffices to show that

$$\chi^{(m)}(B_1 \times \cdots \times B_m) \leq \nu^{(m)}(B_1 \times \cdots \times B_m) \tag{95}$$

for all $B_1, \ldots, B_m \in \mathcal{X}$. Fixing the latter sets we define the system $\mathcal{A}$ of atoms of the generated algebra as in the proof of Lemma 7. For $B \in \mathcal{A}$ we choose $x \in B$ and define $\chi_B := \chi(B)\delta_x$ and $\nu_B := \nu(B)\delta_x$. Then

$$\chi' := \sum_{B \in \mathcal{A}} \chi_B, \quad \nu' := \sum_{B \in \mathcal{A}} \nu_B$$

are finite sums of Dirac measures satisfying $\chi' \leq \nu'$. By (94) we have

$$\chi^{(m)}(B_1 \times \cdots \times B_m) = (\chi')^{(m)}(B_1 \times \cdots \times B_m).$$

A similar identity holds for $\nu^{(m)}$ and $(\nu')^{(m)}$. Therefore (91) (applied to χ' and ν') implies the asserted inequality (95). □

We can now prove a slightly more detailed version of Proposition 1.

Proposition 6 *For any $\chi \in \mathbf{N}_\sigma$ there is a unique sequence $\chi^{(m)}$, $m \in \mathbb{N}$, of symmetric σ-finite measures on $(\mathbb{X}^m, \mathcal{X}^m)$ satisfying $\chi^{(1)} := \chi$ and the recursion (92). Moreover, the mapping $\chi \mapsto \chi^{(m)}$ is measurable. Finally, $\chi^{(m)}(B^m) \leq \chi(B)^m$ for all $m \in \mathbb{N}$ and $B \in \mathcal{X}$.*

Proof For $\chi \in \mathbf{N}_{<\infty}$ the functionals defined by (93) satisfy the recursion (92) and are measures by Lemma 7.

For a general $\chi \in \mathbf{N}_\sigma$ we proceed by induction. For $m = 1$ we have $\chi^{(1)} = \chi$ and there is nothing to prove. Assume now that $m \geq 1$ and that the measures $\chi^{(1)}, \dots, \chi^{(m)}$ satisfy the first $m - 1$ recursions and have the properties stated in the proposition. Then (92) enforces the definition

$$\chi^{(m+1)}(C) := \int K(x_1, \dots, x_m, \chi, C)\, \chi^{(m)}(\mathrm{d}(x_1, \dots, x_m)) \tag{96}$$

for $C \in \mathcal{X}^{m+1}$, where

$$K(x_1, \dots, x_m, \chi, C)$$
$$:= \int \mathbb{1}\{(x_1, \dots, x_{m+1}) \in C\}\, \chi(\mathrm{d}x_{m+1}) - \sum_{j=1}^{m} \mathbb{1}\{(x_1, \dots, x_m, x_j) \in C\}.$$

The function $K\colon \mathbb{X}^m \times \mathbf{N}_\sigma \times \mathcal{X}^m \to (-\infty, \infty]$ is a *signed kernel* in the following sense. The mapping $(x_1, \dots, x_m, \chi) \mapsto K(x_1, \dots, x_m, \chi, C)$ is measurable for all $C \in \mathcal{X}^{m+1}$, while $K(x_1, \dots, x_m, \chi, \cdot)$ is σ-additive for all $(x_1, \dots, x_m, \chi) \in \mathbb{X}^m \times \mathbf{N}_\sigma$. Hence it follows from (96) and the measurability properties of $\chi^{(m)}$ (which are part of the induction hypothesis) that $\chi^{(m+1)}(C)$ is a measurable function of χ.

Next we show that

$$K(x_1, \dots, x_m, \chi, C) \geq 0 \quad \chi^{(m)}\text{-a.e. } (x_1, \dots, x_m) \in \mathbb{X}^m \tag{97}$$

holds for all $\chi \in \mathbf{N}_\sigma$ and all $C \in \mathcal{X}^{m+1}$. Since $\chi^{(m)}$ is a measure (by induction hypothesis) (96), (97) and monotone convergence then imply that $\chi^{(m+1)}$ is a measure. Fix $\chi \in \mathbf{N}_\sigma$ and choose a sequence (χ_n) of finite measures in $\mathbf{N}_\sigma$ such that $\chi_n \uparrow \chi$. Lemma 7 (applied to χ_n and $m + 1$) implies that

$$K(x_1, \dots, x_m, \chi_n, C) \geq 0 \quad (\chi_n)^{(m)}\text{-a.e. } (x_1, \dots, x_m) \in \mathbb{X}^m,\ n \in \mathbb{N}.$$

Indeed, we have for all $B \in \mathcal{X}^m$ that

$$\int_B K(x_1, \dots, x_m, \chi_n, C)\, (\chi_n)^{(m)}(\mathrm{d}(x_1, \dots, x_m)) = (\chi_n)^{(m+1)}((B \times \mathbb{X}) \cap C) \geq 0.$$

Since $K(x_1, \dots, x_m, \cdot, C)$ is increasing, this implies

$$K(x_1, \dots, x_m, \chi, C) \geq 0 \quad (\chi_n)^{(m)}\text{-a.e. } (x_1, \dots, x_m) \in \mathbb{X}^m,\ n \in \mathbb{N}.$$

By induction hypothesis we have that $(\chi_n)^{(m)} \uparrow \chi^{(m)}$ so that (97) follows.

Finally we note that $\chi^{(m)}(B^m) \leq \chi(B)^m$ follows by induction. In particular, $\chi^{(m)}$ is σ-finite. To prove the symmetry of $\chi^{(m)}$ it is then sufficient to show that the restriction of $\chi^{(m)}$ to B^m is symmetric, for any $B \in \mathcal{X}$ with $\chi(B) < \infty$. This fact follows from (94). □

For any $\chi \in \mathbf{N}$, $B \in \mathcal{X}$ with $\chi(B) < \infty$, and $m \in \mathbb{N}$ it follows by induction that

$$\chi^{(m)}(B^m) = \chi(B)(\chi(B) - 1) \cdots (\chi(B) - m + 1).$$

Since χ and $\chi^{(m)}$ are σ-finite, this extends to any $B \in \mathcal{X}$. In particular $\chi^{(m)}$ is the zero measure whenever $\chi(\mathbb{X}) < m$.

References

1. Dellacherie, C., Meyer, P.A.: Probabilities and Potential. North-Holland Mathematics Studies, vol. 29. North-Holland Publishing Company, Amsterdam/New York (1978)
2. Dudley, R.M.: Real Analysis and Probability. Cambridge University Press, Cambridge (2002)
3. Hitsuda, M.: Formula for Brownian partial derivatives. In: Proceedings of the 2nd Japan-USSR Symposium on Probability Theory, pp. 111–114 (1972)
4. Houdré, C., Perez-Abreu, V.: Covariance identities and inequalities for functionals on Wiener space and Poisson space. Ann. Probab. **23**, 400–419 (1995)
5. Houdré, C., Privault, N.: Concentration and deviation inequalities in infinite dimensions via covariance representations. Bernoulli **8**, 697–720 (2002)
6. Itô, K.: Multiple Wiener integral. J. Math. Soc. Jpn. **3**, 157–169 (1951)
7. Itô, K.: Spectral type of the shift transformation of differential processes with stationary increments. Trans. Am. Math. Soc. **81**, 253–263 (1956)
8. Ito, Y.: Generalized Poisson functionals. Probab. Theory Relat. Fields **77**, 1–28 (1988)
9. Janson, S.: Bounds on the distributions of extremal values of a scanning process. Stoch. Process. Appl. **18**, 313–328 (1984)
10. Kabanov, Y.M.: On extended stochastic integrals. Theory Probab. Appl. **20**, 710–722 (1975)
11. Kabanov, Y.M., Skorokhod, A.V.: Extended stochastic integrals. In: Proceedings of the School-Seminar on the Theory of Random Processes, Druskininkai, 25–30 November 1974. Part I. Vilnius (Russian) (1975)
12. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002)
13. Last, G., Penrose, M.D.: Fock space representation, chaos expansion and covariance inequalities for general Poisson processes. Probab. Theory Relat. Fields **150**, 663–690 (2011)
14. Last, G., Penrose, M.D.: Martingale representation for Poisson processes with applications to minimal variance hedging. Stoch. Process. Appl. **121**, 1588–1606 (2011)
15. Last, G., Penrose, M.D.: Lectures on the Poisson Process. Cambridge University Press www.math.kit.edu/stoch/~cost/seite/lehrbuch_poissonlde (2016)
16. Last, G., Penrose, M.D., Schulte, M., Thäle, C.: Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**, 348–364 (2014)
17. Last, G., Peccati, G., Schulte, M.: Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. Probab. Theory Relat. Fields (2014, to appear)
18. Mecke, J.: Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **9**, 36–58 (1967)
19. Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Berlin (2006)
20. Nualart, D., Vives, J.: Anticipative calculus for the Poisson process based on the Fock space. Séminaire Probabilités XXIV. Lecture Notes in Mathematics, vol. 1426, pp. 154–165. Springer, Berlin (1990)

21. Peccati, G., Taqqu, M.S.: Wiener Chaos: Moments, Cumulants and Diagrams. Bocconi & Springer Series, vol. 1, Springer, Milan (2011)
22. Peccati, G., Thäle, C.: Gamma limits and U-statistics on the Poisson space. ALEA Lat. Am. J. Probab. Math. Stat. **10**, 525–560 (2013)
23. Picard, J.: On the existence of smooth densities for jump processes. Probab. Theory Relat. Fields **105**, 481–511 (1996)
24. Privault, N.: Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales. Springer, Berlin (2009)
25. Privault, N.: Combinatorics of Poisson Stochastic Integrals with Random Integrands. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 37–80. Springer, Cham (2016)
26. Skorohod, A.V.: On a generalization of a stochastic integral. Theory Probab. Appl. **20**, 219–233 (1975)
27. Stroock, D.W.: Homogeneous chaos revisited. Séminaire de Probabilités XXI. Lecture Notes in Mathematics, vol. 1247, pp. 1–8. Springer, New York (1987)
28. Surgailis, D.: On multiple Poisson stochastic integrals and associated Markov semigroups. Probab. Math. Stat. **3**, 217–239 (1984)
29. Wiener, N.: The homogeneous chaos. Am. J. Math. **60**, 897–936 (1938).
30. Wu, L.: A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. Probab. Theory Relat. Fields **118**, 427–438 (2000)

Combinatorics of Poisson Stochastic Integrals with Random Integrands

Nicolas Privault

Abstract We present a self-contained account of recent results on moment identities for Poisson stochastic integrals with random integrands, based on the use of functional transforms on the Poisson space. This presentation relies on elementary combinatorics based on the Faà di Bruno formula, partitions and polynomials, which are used together with multiple stochastic integrals, finite difference operators and integration by parts.

1 Introduction

The cumulants $(\kappa_n^X)_{n\geq 1}$ of a random variable X have been defined in [33] and were originally called the "semi-invariants" of X, due to the property $\kappa_n^{X+Y} = \kappa_n^X + \kappa_n^Y$, $n \geq 1$, when X and Y are independent random variables. Precisely, given the moment generating function

$$\mathbb{E}[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n], \tag{1}$$

of a random variable X, where t is in a neighborhood of 0, the *cumulants* of X are defined to be the coefficients $(\kappa_n^X)_{n\geq 1}$ appearing in the series expansion of the logarithmic moment generating function of X, i.e., we have

$$\log(\mathbb{E}[e^{tX}]) = \sum_{n=1}^{\infty} \kappa_n^X \frac{t^n}{n!}, \tag{2}$$

N. Privault (✉)
Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link, Singapore 637371, Singapore
e-mail: nprivault@ntu.edu.sg

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_2

where t is in a neighborhood of 0. In relation with the Faà di Bruno formula, (1) and (2) yield the classical identity

$$\mathbb{E}[X^n] = \sum_{a=0}^{n} \sum_{P_1 \cup \cdots \cup P_a = \{1,\ldots,n\}} \kappa^X_{|P_1|} \cdots \kappa^X_{|P_a|}, \qquad n \in \mathbb{N}, \tag{3}$$

which links the moments $(\mathbb{E}[X^n])_{n\geq 1}$ of a random variable X with its cumulants $(\kappa_n^X)_{n\geq 1}$, cf., e.g., Theorem 1 of [16], and also [15] or §2.4 and Relation (2.4.4) page 27 of [17].

The summation in (3) runs over the partitions $P_1, \ldots, P_a$ of the set $\{1, \ldots, n\}$, i.e., each sequence $P_1, \ldots, P_a$ is a family of nonempty and nonoverlapping subsets of $\{1, \ldots, n\}$ whose union is $\{1, \ldots, n\}$, and $|P_i|$ denotes the cardinal of P_i, cf. §2.2 of [21] for a complete review of the notion of set partition. For example, when X is centered Gaussian we have $\kappa_n^X = 0$, $n \neq 2$, and (3) reads as Wick's theorem for the computation of Gaussian moments of X counting the pair partitions of $\{1, \ldots, n\}$, cf. [10].

In this survey we derive moment identities for Poisson stochastic integrals with random integrands, cf. Theorem 1 below, with application to invariance of Poisson random measures. Our method relies on the tools from combinatorics appearing in [3], i.e., the Faà di Bruno formula and related Stirling numbers, partitions and polynomials, in relation with Poisson random measures, integration by parts on Poisson probability spaces and multiple stochastic integrals. Such moment identities have been recently extended to point processes with Papangelou intensities (see [6] and [5], respectively, for the moments and for the factorial moments of such point processes).

The outline of this survey is as follows. Section 2 starts with preliminaries on combinatorics and the Faà di Bruno formula, providing the needed combinatorial background to rederive the classical identity (3). Then, in Sect. 3 we introduce the Poisson random measures and integration by parts on Poisson probability spaces, along with the tools of $\mathcal{S}$ and $\mathcal{U}$ transforms in view of applications to moment identities. Single and joint moment identities themselves are then detailed in Sect. 4, in relation with set-indexed adaptedness and invariance of Poisson measures.

Our computation of Poisson moments will proceed from the Bismut–Girsanov approach to the stochastic calculus of variations (Malliavin calculus), via the use of functional $\mathcal{S}$ and $\mathcal{U}$-transforms, cf. Sects. 3.3 and 3.4. As an illustration, we start with some informal remarks on that approach in the framework of the Malliavin calculus on the Wiener space. Given $(B_t)_{t\in\mathbb{R}_+}$ a standard Brownian motion and $F(\omega)$ a random functional of the Brownian path $B_t(\omega) = \omega(t)$, $t \in \mathbb{R}_+$, we start from the Girsanov identity

$$\mathbb{E}\left[F\xi(f)\right] = \mathbb{E}\left[F\left(\omega(\cdot) + \int_0^{\cdot} f(s)\mathrm{d}s\right)\right], \tag{4}$$

where $f \in L^2(\mathbb{R}_+)$ and $\xi(f) = X_\infty$ is the terminal value of the (martingale) solution of the stochastic differential equation

$$dX_t = f(t)X_t dB_t, \qquad t \in \mathbb{R}_+. \tag{5}$$

By iteration, the solution of (5) can be written as the series

$$\begin{aligned}
\xi(f) &= X_\infty \\
&= 1 + \int_0^\infty f(t)X_t dB_t \\
&= 1 + \sum_{n=1}^\infty \int_0^\infty \int_0^{t_n} \cdots \int_0^{t_2} f(t_1)\cdots f(t_n) dB_{t_1}\cdots dB_{t_n} \\
&= 1 + \sum_{n=1}^\infty \frac{1}{n!} I_n(f^{\otimes n}),
\end{aligned}$$

of multiple stochastic integrals

$$I_n(f^{\otimes n}) = n! \int_0^\infty \int_0^{t_n} \cdots \int_0^{t_2} f(t_1)\cdots f(t_n) dB_{t_1}\cdots dB_{t_n}, \qquad n \ge 1.$$

We can then rewrite (4) as

$$\begin{aligned}
\mathbb{E}\,[F\xi(f)] &= \mathbb{E}[F] + \sum_{n=1}^\infty \frac{1}{n!}\mathbb{E}[F I_n(f^{\otimes n})] \qquad (6)\\
&= \mathbb{E}\left[F\left(\omega(\cdot) + \int_0^{\cdot} f(s)ds\right)\right] \\
&= \mathbb{E}[F] + \sum_{n=1}^\infty \frac{1}{n!}\frac{\partial^n}{\partial\varepsilon^n}\mathbb{E}\left[F\left(\omega(\cdot) + \varepsilon\int_0^{\cdot} f(s)ds\right)\right]_{\varepsilon=0}.
\end{aligned}$$

By successive differentiations this yields the iterated integration by parts formula

$$\mathbb{E}[F I_n(f^{\otimes n})] = \mathbb{E}[\nabla_f^n F], \tag{7}$$

where ∇_f is the gradient operator defined by

$$\nabla_f F := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(F \left(\omega(\cdot) + \varepsilon \int_0^{\cdot} f(s) \mathrm{d}s \right) - F(\omega(\cdot)) \right).$$

On the other hand, on the Wiener space the above Girsanov shift acts on the paths $(\omega(t))_{t \in \mathbb{R}_+}$ of the underlying Brownian motion $(B_t)_{t \in \mathbb{R}_+}$ as

$$\omega(\cdot) \longmapsto \omega(\cdot) + \varepsilon \int_0^{\cdot} f(s) \mathrm{d}s,$$

which yields

$$\mathbb{E}[\nabla_f^n F] = \mathbb{E}\left[\int_0^\infty \cdots \int_0^\infty f(s_1) \cdots f(s_n) D_{s_1} \cdots D_{s_n} F \mathrm{d}s_1 \cdots \mathrm{d}s_n \right], \tag{8}$$

where $D_s F$ is the Malliavin gradient which satisfies

$$\nabla_f F = \int_0^\infty D_s F f(s) \mathrm{d}s,$$

hence by (7) and (8) we obtain the iterated integration by parts identity

$$\mathbb{E}\left[I_k(f^{\otimes k}) F \right] = \mathbb{E}\left[\int_0^\infty \cdots \int_0^\infty f(s_1) \cdots f(s_k) D_{s_1} \cdots D_{s_k} F \mathrm{d}s_1 \cdots \mathrm{d}s_k \right], \qquad k \geq 1, \tag{9}$$

a relation that can be the basis for the computation of moments. On the Wiener space the operator D also satisfies the identity

$$D_t I_n(g^{\otimes n}) = n g(t) I_{n-1}(g^{\otimes (n-1)}), \qquad t \in \mathbb{R}_+, \tag{10}$$

which can be used to recover (9) as the Stroock's formula [32], cf. Corollary 1 below for the Poisson case.

However, when carrying over this approach to the probability space of a Poisson random measure it turns out that there is no differential operator ∇_f that can satisfy both relations (8) and (10) above. In the sequel we will develop the above approach on the Poisson space via the use of finite difference operators.

2 Combinatorics

In this section we provide the necessary combinatorial background for the derivation of cumulant-type moment identities. We refer the reader to [21] and references therein, cf. also [22], for additional background on combinatorial probability and for the relationships between the moments and cumulants of random variables.

2.1 Faà di Bruno Formula and Bell Polynomials

2.1.1 Faà di Bruno formula

The Faà di Bruno formula plays a fundamental role in the combinatorics of moments, cumulants, and factorial moments. Namely, instead of the multinomial identity

$$\left(\sum_{l=1}^{n} x_l\right)^k = k! \sum_{\substack{d_1+\cdots+d_n=k \\ d_1\geq 0,\ldots,d_n\geq 0}} \frac{x_1^{d_1}}{d_1!}\cdots\frac{x_n^{d_n}}{d_n!}, \tag{11}$$

we will use the combinatorial identity

$$\left(\sum_{n=1}^{\infty} x_n\right)^k = \sum_{n=k}^{\infty} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} x_{d_1}\cdots x_{d_k}, \tag{12}$$

or

$$\left(\sum_{n=1}^{\infty} x_{1,n}\right)\cdots\left(\sum_{n=1}^{\infty} x_{k,n}\right) = \sum_{n=k}^{\infty} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} x_{1,d_1}\cdots x_{k,d_k}. \tag{13}$$

The above identity (12) is equivalent to the Faà di Bruno formula, i.e., given $g(x)$ and $f(y)$ two functions given by the series expansions

$$g(x) = \sum_{n=1}^{\infty} b_n \frac{x^n}{n!}$$

with $g(0) = 0$ and

$$f(y) = \sum_{k=0}^{\infty} a_k \frac{y^k}{k!},$$

the series expansion of $f(g(x))$ is given by

$$f(g(x)) = \sum_{k=0}^{\infty} \frac{a_k}{k!} \left(\sum_{n=1}^{\infty} b_n \frac{x^n}{n!} \right)^k$$
$$= \sum_{k=0}^{\infty} \frac{a_k}{k!} \sum_{n=k}^{\infty} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} b_{d_1} \cdots b_{d_k} \frac{x^{d_1}}{d_1!} \cdots \frac{x^{d_k}}{d_k!}$$
$$= \sum_{n=0}^{\infty} x^n \sum_{k=0}^{n} \frac{a_k}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} \frac{b_{d_1}}{d_1!} \cdots \frac{b_{d_k}}{d_k!}. \qquad (14)$$

In the sequel we will often rewrite (12) using sums over partitions $P_1^n, \ldots, P_k^n$ of $\{1, \ldots, n\}$ into subsets with cardinals $|P_1^n|, \ldots, |P_k^n|$, as

$$\frac{n!}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} \frac{b_{d_1}}{d_1!} \cdots \frac{b_{d_k}}{d_k!} = \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} b_{|P_1^n|} \cdots b_{|P_k^n|}.$$

2.1.2 Bell Polynomials

The Faà di Bruno formula (14) can be rewritten as

$$f(g(x)) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^{n} a_k B_{n,k}(b_1, \ldots, b_{n-k+1}), \qquad (15)$$

where $B_{n,k}(b_1, \ldots, b_{n-k+1})$ is the Bell polynomial of order (n, k) defined by

$$B_{n,k}(b_1, \ldots, b_{n-k+1}) := \frac{1}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} \frac{n!}{d_1! \cdots d_k!} b_{d_1} \cdots b_{d_k}$$
$$= \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} b_{|P_1^n|} \cdots b_{|P_k^n|}$$
$$= n! \sum_{\substack{r_1+2r_2+\cdots+(n-k+1)r_{n-k+1}=n \\ r_1+r_2+\cdots+r_{n-k+1}=k \\ r_1\geq 0,\ldots,r_{n-k+1}\geq 0}} \prod_{l=1}^{n-k+1} \left(\frac{1}{r_l!} \left(\frac{b_l}{l!} \right)^{r_l} \right)$$
$$= \frac{n!}{k!} \sum_{\substack{r_1+2r_2+\cdots+(n-k+1)r_{n-k+1}=n \\ r_1+r_2+\cdots+r_{n-k+1}=k \\ r_1\geq 0,\ldots,r_{n-k+1}\geq 0}} \frac{k!}{r_1! \cdots r_{n-k+1}!} \left(\frac{b_1}{1!} \right)^{r_1} \cdots \left(\frac{b_{n-k+1}}{(n-k+1)!} \right)^{r_{n-k+1}},$$

cf., e.g., Definition 2.4.1 of [21], with $B_{n,0}(b_1, \ldots, b_n) = 0$, $n \geq 1$, and $B_{0,0} = 1$. In particular when $f(y) = e^y$ we have $a_k = 1$, $k \geq 0$, and (15) rewrites as

$$\exp\left(\sum_{n=1}^{\infty} \frac{b_n}{n!}\right) = \sum_{n=0}^{\infty} \frac{1}{n!} A_n(b_1, \ldots, b_n), \tag{16}$$

where

$$A_n(b_1, \ldots, b_n) = \sum_{k=0}^{n} B_{n,k}(b_1, \ldots, b_{n-k+1}) \tag{17}$$

$$= \sum_{k=0}^{n} \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} b_{|P_1^n|} \cdots b_{|P_k^n|}$$

$$= n! \sum_{k=0}^{n} \sum_{\substack{r_1+2r_2+\cdots+(n-k+1)r_{n-k+1}=n \\ r_1+r_2+\cdots+r_{n-k+1}=k \\ r_1 \geq 0,\ldots,r_{n-k+1} \geq 0}} \prod_{l=1}^{n-k+1} \left(\frac{1}{r_l!}\left(\frac{b_l}{l!}\right)^{r_l}\right) \tag{18}$$

$$= n! \sum_{\substack{r_1+2r_2+\cdots+nr_n=n \\ r_1 \geq 0,\ldots,r_n \geq 0}} \prod_{l=1}^{n} \left(\frac{1}{r_l!}\left(\frac{b_l}{l!}\right)^{r_l}\right)$$

is the (complete) Bell polynomial of degree n. Relation (16) is a common formulation of the Faà di Bruno formula and it will be used in the proof of Proposition 5 below on the $\mathcal{U}$-transform on the Poisson space.

2.2 Stirling Inversion

The Stirling numbers will be used for the construction of multiple stochastic integrals, as well as to establish their relations to the Charlier polynomials in Sect. 3.2. Let

$$S(n,k) = \left\{ n \atop k \right\} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^n$$

$$= \frac{1}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1 \geq 1,\ldots,d_k \geq 1}} \frac{n!}{d_1! \cdots d_k!} \tag{19}$$

$$= \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} \mathbf{1},$$

denote the *Stirling number of the second kind* with $S(n,0) = 0, n \geq 1$, and $S(0,0) = 1$, cf. page 824 of [1], i.e., $S(n,k)$ is the number of partitions of a set of n objects into k nonempty subsets, cf. also Relation (3) page 59 of [3], with

$$B_{n,k}(x,\ldots,x) = x^k S(n,k), \qquad 0 \leq k \leq n.$$

Let also

$$s(n,k) = \begin{bmatrix} n \\ k \end{bmatrix} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i \binom{k}{i}(k-i)^n$$

denote the (signed) *Stirling number of the first kind*, cf., e.g., page 824 of [1], i.e., $(-1)^{n-k}s(n,k)$ is the number of permutations of n elements which contain exactly k permutation cycles.

The following Lemma 1, cf., e.g., Relation (3) page 59 of [3], also relies on the Faà di Bruno formula applied to

$$f(t) = \frac{t^k}{k!} \quad \text{and} \quad a_n = \mathbf{1}_{\{n=k\}}$$

and

$$g(t) = \log(1+t) \quad \text{and} \quad b_k = \mathbf{1}_{\{n=k\}}.$$

Lemma 1 *Assume that the function $f(t)$ has the series expansion*

$$f(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} a_n, \qquad t \in \mathbb{R}.$$

Then we have

$$f(e^t - 1) = \sum_{k=0}^{\infty} \frac{t^k}{k!} c_k, \qquad t \in \mathbb{R},$$

with

$$c_n = \sum_{k=0}^{n} a_k S(n,k),$$

and the inversion formula

$$a_n = \sum_{k=0}^{n} c_k s(n,k), \qquad n \in \mathbb{N}.$$

Proof Applying the Faà di Bruno identity (14) to $g(t) = e^t - 1$ and using (19) we have

$$f(e^t - 1) = \sum_{k=0}^{\infty} a_k \frac{(e^t - 1)^k}{k!} = \sum_{k=0}^{\infty} a_k \sum_{n=k}^{\infty} \frac{t^n}{n!} S(n,k)$$
$$= \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{k=0}^{n} a_k S(n,k) = \sum_{n=0}^{\infty} \frac{t^n}{n!} c_n, \qquad t \in \mathbb{R},$$

with

$$c_n = \sum_{k=0}^{n} a_k S(n,k).$$

Conversely we have

$$f(t) = \sum_{k=0}^{\infty} \frac{c_k}{k!} (\log(1+t))^k = \sum_{k=0}^{\infty} c_k \sum_{n=k}^{\infty} \frac{t^n}{n!} s(n,k)$$
$$= \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{k=0}^{n} c_k s(n,k) = \sum_{n=0}^{\infty} \frac{t^n}{n!} a_n, \qquad t \in \mathbb{R},$$

with

$$a_n = \sum_{k=0}^{n} c_k s(n,k).$$

□

As a consequence of Lemma 1, the Stirling transform

$$a_n = \sum_{k=0}^{n} c_k s(n,k), \qquad n \in \mathbb{N},$$

can be inverted as

$$c_n = \sum_{k=0}^{n} a_k S(n,k), \qquad n \in \mathbb{N},$$

i.e., we have the inversion formula

$$\sum_{k=l}^{n} S(n,k) s(k,l) = \mathbf{1}_{\{n=l\}}, \qquad n, l \in \mathbb{N}, \tag{20}$$

for Stirling numbers, cf., e.g., page 825 of [1]. As particular cases of the Stirling transform of Lemma 1 we find that

$$\frac{1}{k!}(e^{\lambda}-1)^k = \frac{1}{k!}\left(\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\right)^k = \frac{1}{k!}\sum_{n=k}^{\infty}\frac{\lambda^n}{n!}\sum_{\substack{d_1+\cdots+d_k=n\\ d_1\geq 1,\ldots,d_k\geq 1}}\frac{n!}{d_1!\cdots d_k!}$$

$$= \sum_{n=k}^{\infty}\frac{\lambda^n}{n!}B_{n,k}(1,\ldots,1) = \sum_{n=k}^{\infty}\frac{\lambda^n}{n!}S(n,k), \qquad k\geq 1. \tag{21}$$

We also have

$$\begin{aligned}\frac{1}{k!}(\log(1+t))^k &= \frac{(-1)^k}{k!}\left(\sum_{n=1}^{\infty}\frac{(-1)^n}{n}t^n\right)^k \\ &= (-1)^k\sum_{n=k}^{\infty}\frac{t^n}{n!}B_{n,k}\left(-1,\frac{1}{2},-\frac{1}{3},\ldots,\frac{(-1)^{n-k+1}}{n-k+1}\right) \\ &= \frac{(-1)^k}{k!}\sum_{n=k}^{\infty}(-1)^n\frac{t^n}{n!}\sum_{\substack{d_1+\cdots+d_k=n\\ d_1\geq 1,\ldots,d_k\geq 1}}\frac{n!}{d_1\cdots d_k} \\ &= \sum_{n=k}^{\infty}\frac{t^n}{n!}s(n,k), \qquad k\geq 1,\end{aligned}$$

which shows the relation

$$s(n,k) = \frac{n!}{k!}\sum_{d_1+\cdots+d_k=n}\frac{(-1)^{n-k}}{d_1\cdots d_k}. \tag{22}$$

In particular, taking $c_k = x^k$ and letting $a_n = x_{(n)}$ be defined by the falling factorial

$$x_{(n)} := x(x-1)\cdots(x-n+1), \qquad k,n\geq 0,$$

i.e.,

$$f(e^t-1) = e^{xt} = \sum_{k=0}^{\infty}\frac{t^k}{k!}x^k,$$

and by Lemma 1 we get

$$f(t) = (1+x)^t = \sum_{n=0}^{\infty}\frac{t^n}{n!}x_{(n)}, \tag{23}$$

which will be used in Lemma 2 below on the Charlier polynomials.

By Stirling inversion we also find the expansion of the falling factorial

$$x_{(n)} = x(x-1)\cdots(x-n+1) = \sum_{k=0}^{n} s(n,k)x^k \tag{24}$$

and

$$x^n = \sum_{k=0}^{n} S(n,k)\, x(x-1)\cdots(x-k+1),$$

cf., e.g., [9] or page 72 of [8].

2.3 Charlier and Touchard Polynomials

2.3.1 Charlier Polynomials

The Charlier polynomials $C_n(x,\lambda)$ of order $n \in \mathbb{N}$ with parameter $\lambda > 0$ are essential in the construction of multiple Poisson stochastic integrals in Sect. 3.2. They can be defined through their generating function

$$\psi_\lambda(x,t) := \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} C_n(x,t) = e^{-\lambda t}(1+\lambda)^x, \qquad x,t \in \mathbb{R}_+, \tag{25}$$

$\lambda \in (-1,1)$, cf., e.g., §4.3.3 of [30].

Lemma 2 *We have*

$$C_n(x,\lambda) = \sum_{k=0}^{n} x^k \sum_{l=0}^{n} \binom{n}{l} (-\lambda)^{n-l} s(l,k), \qquad x,\lambda \in \mathbb{R}. \tag{26}$$

Proof We check that defining $C_n(x,t)$ by (26) yields

$$\begin{aligned}
\psi_\lambda(x,t) &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} C_n(x,t) \\
&= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{k=0}^{n} x^k \sum_{l=k}^{n} \binom{n}{l} (-t)^{n-l} s(l,k) \\
&= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{l=0}^{n} \binom{n}{l} (-t)^{n-l} \sum_{k=0}^{l} x^k s(l,k)
\end{aligned}$$

$$= \sum_{l=0}^{\infty} x_{(l)} \sum_{n=l}^{\infty} \frac{\lambda^n}{n!} \frac{n!}{(n-l)!l!} (-t)^{n-l}$$
$$= \sum_{l=0}^{\infty} x_{(l)} \frac{\lambda^l}{l!} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} (-t)^n$$
$$= e^{-\lambda t} \sum_{l=0}^{\infty} x_{(l)} \frac{\lambda^l}{l!}$$
$$= e^{-\lambda t}(1+\lambda)^x,$$

$\lambda, t > 0$, $x \in \mathbb{N}$, where we applied (23) and (24). □

As a consequence of Lemma 2 and (24), the Charlier polynomial $C_n(x, \lambda)$ can be rewritten in terms of the falling factorial $x_{(n)}$ as

$$C_n(x,\lambda) = \sum_{l=0}^{n} \binom{n}{l} (-\lambda)^{n-l} \sum_{k=0}^{l} x^k s(l,k) = \sum_{l=0}^{n} \binom{n}{l} (-\lambda)^{n-l} x_{(l)}, \quad x, \lambda \in \mathbb{R}. \tag{27}$$

Lemma 3 *We have the orthogonality relation*

$$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} C_n(k,\lambda) C_m(k,\lambda) = n! \lambda^n \mathbf{1}_{\{n=m\}}. \tag{28}$$

Proof We have

$$e^{\lambda ab} = e^{-\lambda(1+a+b)} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (1+a)^k (1+b)^k$$
$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \psi_a(k,\lambda) \psi_b(k,\lambda)$$
$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{a^n}{n!} \frac{b^m}{m!} C_n(k,\lambda) C_m(k,\lambda),$$

which shows that

$$\sum_{p=0}^{\infty} \lambda^p \frac{(ab)^p}{p!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{a^n}{n!} \frac{b^m}{m!} C_n(k,\lambda) C_m(k,\lambda)$$
$$= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(ab)^n}{(n!)^2} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (C_n(k,\lambda))^2,$$

with

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} C_n(k,\lambda) C_m(k,\lambda) = 0$$

for $n \neq m$, and

$$n!\lambda^n = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (C_n(k,\lambda))^2,$$

for $n = m$. □

2.3.2 Touchard Polynomials

The Touchard polynomials can be used to express the moments of a Poisson random variable as a function of its intensity parameter. They can be defined by their generating function

$$e^{\lambda(e^t-1)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} T_n(\lambda), \qquad t \in \mathbb{R},$$

and from (16) or (21) they satisfy

$$\begin{aligned} T_n(\lambda) :&= A_n(\lambda,\ldots,\lambda) = \sum_{k=0}^{n} B_{n,k}(\lambda,\ldots,\lambda) \\ &= \sum_{k=1}^{n} \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} \lambda^k = \sum_{k=0}^{n} \lambda^k S(n,k), \end{aligned} \tag{29}$$

cf., e.g., Proposition 2 of [4] or §3.1 of [20]. Relation (29) above will be used in the proof of the combinatorial Lemma 7 below.

2.4 Moments and Cumulants of Random Variables

Given the identity (1) defining the moment generating function of X, we can write

$$\mathbb{E}[e^{tX}] = 1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] + o(t^2),$$

which allows us to rewrite the cumulant generating function (2) as

$$\begin{aligned}
\log(\mathbb{E}[e^{tX}]) &= \log\left(1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] + o(t^2)\right) \\
&= t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] - \frac{1}{2}\left(t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2]\right)^2 + o(t^2) \\
&= t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] - \frac{t^2}{2}(\mathbb{E}[X])^2 + o(t^2) \\
&= t\mathbb{E}[X] + \frac{t^2}{2}\mathrm{Var}[X] + o(t^2),
\end{aligned}$$

hence $\kappa_1^X = \mathbb{E}[X]$ and $\kappa_2^X = \mathrm{Var}[X]$. More generally, as a consequence of (16), the moment generating function of X expands using the complete Bell polynomials $A_n(b_1, \ldots, b_n)$ of (17) as

$$\begin{aligned}
\mathbb{E}[e^{tX}] &= \exp(\log(\mathbb{E}[e^{tX}])) \\
&= \exp\left(\sum_{n=1}^{\infty} \kappa_n^X \frac{t^n}{n!}\right) \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} A_n(\kappa_1^X, \ldots, \kappa_n^X),
\end{aligned}$$

which shows by comparison with (1) that

$$\begin{aligned}
\mathbb{E}[X^n] &= A_n(\kappa_1^X, \kappa_2^X, \ldots, \kappa_n^X) \\
&= \sum_{k=0}^{n} \frac{n!}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1\geq 1,\ldots,d_k\geq 1}} \frac{\kappa_{d_1}^X}{d_1!} \cdots \frac{\kappa_{d_k}^X}{d_k!} \\
&= \sum_{k=0}^{n} \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \kappa_{|P_1^n|}^X \cdots \kappa_{|P_k^n|}^X, \qquad (30)
\end{aligned}$$

and allows us to recover (3).

The identity (30) can also be recovered from the Thiele [33] recursion formula

$$\mathbb{E}[X^n] = \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-l-1)!} \kappa_{n-l}^X \mathbb{E}[X^l] = \sum_{l=1}^{n} \frac{(n-1)!}{(n-l)!(l-1)!} \kappa_l^X \mathbb{E}[X^{n-l}] \qquad (31)$$

between moments and cumulants of random variables, cf., e.g., §1.3.2 of [22]. Indeed, assuming at the order $n \geq 1$ that

$$\mathbb{E}[X^n] = \sum_{a=0}^{n} \frac{n!}{a!} \sum_{\substack{l_1+\cdots+l_a=n \\ l_1\geq 1,\ldots,l_a\geq 1}} \frac{\kappa_{l_1}^X}{l_1!} \cdots \frac{\kappa_{l_a}^X}{l_a!} = \sum_{a=0}^{n} \sum_{P_1^n\cup\cdots\cup P_a^n=\{1,\ldots,n\}} \kappa_{|P_1^n|}^X \cdots \kappa_{|P_a^n|}^X,$$

and using (31), we have, at the order $n+1$,

$$\begin{aligned}
\mathbb{E}[X^{n+1}] &= \sum_{k=1}^{n+1} \binom{n}{k-1} \kappa_k^X \mathbb{E}[X^{n+1-k}] \\
&= \sum_{k=1}^{n+1} \frac{n!}{(k-1)!} \kappa_k^X \sum_{a=0}^{n+1-k} \frac{1}{a!} \sum_{\substack{l_1+\cdots+l_a=n+1-k \\ l_1\geq 1,\ldots,l_a\geq 1}} \frac{\kappa_{l_1}^X}{l_1!} \cdots \frac{\kappa_{l_a}^X}{l_a!} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} \kappa_k^X \sum_{a=0}^{n+1-k} \sum_{P_1^{n+1-k}\cup\cdots\cup P_a^{n+1-k}=\{1,\ldots,n+1-k\}} \kappa_{|P_1^{n+1-k}|}^X \cdots \kappa_{|P_a^{n+1-k}|}^X \\
&= \sum_{a=0}^{n} \sum_{k=1}^{n+1-a} \binom{n}{k-1} \kappa_k^X \sum_{P_1^{n+1-k}\cup\cdots\cup P_a^{n+1-k}=\{1,\ldots,n+1-k\}} \kappa_{|P_1^{n+1-k}|}^X \cdots \kappa_{|P_a^{n+1-k}|}^X \\
&= \sum_{a=0}^{n} \sum_{P_1^{n+1}\cup\cdots\cup P_{a+1}^{n+1}=\{1,\ldots,n+1\}} \kappa_{|P_1^{n+1}|}^X \cdots \kappa_{|P_{a+1}^{n+1}|}^X \qquad (32) \\
&= \sum_{a=1}^{n+1} \sum_{P_1^{n+1}\cup\cdots\cup P_a^{n+1}=\{1,\ldots,n+1\}} \kappa_{|P_1^{n+1}|}^X \cdots \kappa_{|P_a^{n+1}|}^X \\
&= \sum_{a=0}^{n+1} \frac{(n+1)!}{a!} \sum_{\substack{l_1+\cdots+l_a=n+1 \\ l_1\geq 1,\ldots,l_a\geq 1}} \frac{\kappa_{l_1}^X}{l_1!} \cdots \frac{\kappa_{l_a}^X}{l_a!},
\end{aligned}$$

where in (32) the set P_{a+1}^{n+1} of cardinal $|P_{a+1}^{n+1}| = k$ is built by combining $\{n+1\}$ with $k-1$ elements of $\{1,\ldots,n\}$.

The cumulant formula (30) can also be inverted to compute the cumulant κ_n^X from the moments μ_n^X of X by the inversion formula

$$\kappa_n^X = \sum_{a=1}^{n} (a-1)!(-1)^{a-1} \sum_{P^{n+1}\cup\cdots\cup P_a^n=\{1,\ldots,n\}} \mu_{|P_1^n|}^X \cdots \mu_{|P_a^n|}^X, \qquad n \geq 1, \qquad (33)$$

where the sum runs over the partitions $P_1^n, \ldots, P_a^n$ of $\{1, \ldots, n\}$ with cardinal $|P_i^n|$ by the Faà di Bruno formula, cf. Theorem 1 of [16], and also [15] or §2.4 and Relation (2.4.3) page 27 of [17].

2.4.1 Example: Gaussian Cumulants

When X is centered we have $\kappa_1^X = 0$ and $\kappa_2^X = \mathbb{E}[X^2] = \mathrm{Var}[X]$, and X becomes Gaussian if and only if $\kappa_n^X = 0$, $n \geq 3$, i.e., $\kappa_n^X = \mathbf{1}_{\{n=2\}}\sigma^2$, $n \geq 1$, or

$$(\kappa_1^X, \kappa_2^X, \kappa_3^X, \kappa_4^X, \ldots) = (0, \sigma^2, 0, 0, \ldots).$$

When X is centered Gaussian we have $\kappa_n^X = 0$, $n \neq 2$, and (30) can be read as Wick's theorem for the computation of Gaussian moments of $X \simeq \mathcal{N}(0, \sigma^2)$ by counting the pair partitions of $\{1, \ldots, n\}$, cf. [10], as

$$\mathbb{E}[X^n] = \sigma^n \sum_{k=1}^n \sum_{\substack{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\} \\ |P_1^n|=2,\ldots,|P_k^n|=2}} \kappa_{|P_1^n|}^X \cdots \kappa_{|P_a^n|}^X = \begin{cases} \sigma^n (n-1)!!, & n \text{ even}, \\ 0, & n \text{ odd}, \end{cases} \tag{34}$$

where the double factorial

$$(n-1)!! = \prod_{1 \leq 2k \leq n} (2k-1) = 2^{-n/2} \frac{n!}{(n/2)!}$$

counts the number of pair-partitions of $\{1, \ldots, n\}$ when n is even.

2.4.2 Example: Poisson Cumulants

In the particular case of a Poisson random variable $Z \simeq \mathcal{P}(\lambda)$ with intensity $\lambda > 0$ we have

$$\mathbb{E}[e^{tZ}] = \sum_{n=0}^\infty e^{nt} \mathbb{P}(Z = n) = e^{-\lambda} \sum_{n=0}^\infty \frac{(\lambda e^t)^n}{n!} = e^{\lambda(e^t - 1)}, \qquad t \in \mathbb{R}_+,$$

hence $\kappa_n^Z = \lambda$, $n \geq 1$, or

$$(\kappa_1^Z, \kappa_2^Z, \kappa_3^Z, \kappa_4^Z, \ldots) = (\lambda, \lambda, \lambda, \lambda, \ldots),$$

and by (30) we have

$$\begin{aligned}\mathbb{E}_\lambda[Z^n] &= A_n(\lambda,\ldots,\lambda) = \sum_{k=0}^{n} B_{n,k}(\lambda,\ldots,\lambda)\\ &= \sum_{k=1}^{n} \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \lambda^k = \sum_{k=0}^{n}\lambda^k S(n,k)\\ &= T_n(\lambda),\end{aligned}$$

i.e., the n-th Poisson moment with intensity parameter $\lambda > 0$ is given by $T_n(\lambda)$, where T_n is the Touchard polynomial of degree n.

In the case of centered Poisson random variables, we note that Z and $Z - \mathbb{E}[Z]$ have same cumulants of order $k \geq 2$, hence in case $Z - \mathbb{E}[Z]$ is a centered Poisson random variable with intensity $\lambda > 0$ we have

$$\mathbb{E}[(Z-\mathbb{E}[Z])^n] = \sum_{a=1}^{n} \sum_{\substack{P_1^n\cup\cdots\cup P_a^n=\{1,\ldots,n\}\\ |P_1^n|\geq 2,\ldots,|P_a^n|\geq 2}} \lambda^a = \sum_{k=0}^{n}\lambda^k S_2(n,k), \qquad n \geq 0,$$

where $S_2(n,k)$ is the number of ways to partition a set of n objects into k nonempty subsets of size at least 2, cf. [25].

2.4.3 Example: Compound Poisson Cumulants

Consider the compound Poisson random variable

$$\beta_1 Z_{\alpha_i} + \cdots + \beta_p Z_{\alpha_p} \tag{35}$$

with Lévy measure

$$\alpha_i \delta_{\beta_1} + \cdots + \alpha_p \delta_{\beta_p},$$

where $\beta_1,\ldots,\beta_p \in \mathbb{R}$ are constant parameters and $Z_{\alpha_1},\ldots,Z_{\alpha_p}$ is a sequence of independent Poisson random variables with respective parameters $\alpha_1,\ldots,\alpha_p \in \mathbb{R}_+$. The moment generating function of (35) is given by

$$\mathbb{E}[e^{t(\beta_1 Z_{\alpha_i}+\cdots+\beta_p Z_{\alpha_p})}] = e^{\alpha_1(e^{t\beta_1}-1)+\cdots+\alpha_p(e^{t\beta_p}-1)},$$

which shows that the cumulant of order $k \geq 1$ of (35) is given by

$$\alpha_1\beta_1^k + \cdots + \alpha_p\beta_p^k.$$

As a consequence of the identity (30), the moment of order n of (35) is given by

$$\mathbb{E}\left[\left(\sum_{i=1}^{p}\beta_i Z_{\lambda\alpha_i}\right)^n\right] \tag{36}$$

$$= \sum_{m=0}^{n} \sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\dots,n\}} (\alpha_1\beta_1^{|P_1^n|}+\cdots+\alpha_p\beta_p^{|P_1^n|})\cdots(\alpha_1\beta_1^{|P_m^n|}+\cdots+\alpha_p\beta_p^{|P_m^n|})$$

$$= \sum_{m=0}^{n} \sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\dots,n\}} \sum_{i_1,\dots,i_m=1}^{p} \beta_{i_1}^{|P_1^n|}\alpha_{i_1}\cdots\beta_{i_m}^{|P_m^n|}\alpha_{i_m},$$

where the above sum runs over all partitions $P_1^n,\dots,P_m^n$ of $\{1,\dots,n\}$.

2.4.4 Example: Infinitely Divisible Cumulants

In the case where X is the infinitely divisible Poisson stochastic integral

$$X = \int_0^\infty h(t)\mathrm{d}N_t$$

with respect to a standard Poisson process $(N_t)_{t\in\mathbb{R}_+}$ with intensity $\lambda > 0$ and $h \in \bigcap_{p=1}^\infty L^p(\mathbb{R}_+)$, the logarithmic generating function

$$\log\mathbb{E}\left[\exp\left(\int_0^\infty h(t)\mathrm{d}N_t\right)\right] = \lambda\int_0^\infty (e^{h(t)}-1)\mathrm{d}t = \lambda\sum_{n=1}^\infty \frac{1}{n!}\int_0^\infty h^n(t)\mathrm{d}t$$
$$= \sum_{n=1}^\infty \kappa_n^X\frac{t^n}{n!},$$

shows that the cumulants of $\int_0^\infty h(t)\mathrm{d}N_t$ are given by

$$\kappa_n^X = \lambda\int_0^\infty h^n(t)\mathrm{d}t, \qquad n\geq 1, \tag{37}$$

and (30) becomes the moment identity

$$\mathbb{E}\left[\left(\int_0^\infty h(t)\mathrm{d}N_t\right)^n\right] = \sum_{k=1}^{n}\lambda^k \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\dots,n\}} \int_0^\infty h^{|P_1^n|}(t)\mathrm{d}t\cdots\int_0^\infty h^{|P_k^n|}(t)\mathrm{d}t, \tag{38}$$

where the sum runs over all partitions $P_1^n, \ldots, P_k^n$ of $\{1, \ldots, n\}$, cf. [2] for the non-compensated case and [28], Proposition 3.2 for the compensated case.

3 Analysis of Poisson Random Measures

In this section we introduce the basic definitions and notations relative to Poisson random measures, and we derive the functional transform identities that will be useful for the computation of moments in Sect. 4.

3.1 Poisson Point Processes

From now on we consider a proper Poisson point process η on the space $\mathbf{N}_\sigma(\mathbb{X})$ of all σ-finite counting measures on a measure space $(\mathbb{X}, \mathscr{X})$ equipped with a σ-finite intensity measure $\mu(\mathrm{d}x)$, see [12, 13] for further details and additional notation. The random measure η in $\mathbf{N}_\sigma(\mathbb{X})$ will be represented as

$$\eta = \sum_{n=1}^{\eta(\mathbb{X})} \delta_{x_n},$$

where $(x_n)_{n=1}^{\eta(\mathbb{X})}$ is a (random) sequence in $\mathbb{X}$, δ_x denotes the Dirac measure at $x \in \mathbb{X}$, and $\eta(\mathbb{X}) \in \mathbb{N} \cup \{\infty\}$ denote the cardinality of η identified with the sequence $(x_n)_n$.

Recall that the probability law $\mathbb{P}_\eta$ of η is that of a Poisson probability measure with intensity $\mu(\mathrm{d}x)$ on $\mathbb{X}$: it is the only probability measure on $\mathbf{N}_\sigma(\mathbb{X})$ satisfying

(1) For any measurable subset $A \in \mathscr{X}$ of $\mathbb{X}$ such that $\mu(A) < \infty$, the number $\eta(A)$ of configuration points contained in A is a Poisson random variable with intensity $\mu(A)$, i.e.,

$$\mathbb{P}_\eta(\{\eta \in \mathbf{N}_\sigma(\mathbb{X}) \ : \ \eta(A) = n\}) = e^{-\mu(A)} \frac{(\mu(A))^n}{n!}, \qquad n \in \mathbb{N}.$$

(2) In addition, if $A_1, \ldots, A_n$ are disjoint subsets of $\mathbb{X}$ with $\mu(A_k) < \infty$, $k = 1, \ldots, n$, the $\mathbb{N}^n$-valued random vector

$$\eta \longmapsto (\eta(A_1), \ldots, \eta(A_n)), \qquad \eta \in \mathbf{N}_\sigma(\mathbb{X}),$$

is made of independent random variables for all $n \geq 1$.

When $\mu(\mathbb{X}) < \infty$ the expectation under the Poisson measure $\mathbb{P}_\eta$ can be written as

$$\mathbb{E}[F(\eta)] = e^{-\mu(\mathbb{X})} \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} f_n(x_1, \ldots, x_n)\mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n) \tag{39}$$

for a random variable F of the form

$$F(\eta) = \sum_{n=0}^{\infty} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(x_1, \ldots, x_n) \tag{40}$$

where for each $n \geq 1$, f_n is a symmetric integrable function of $\eta = \{x_1, \ldots, x_n\}$ when $\eta(\mathbb{X}) = n$, cf., e.g., §6.1 of [24].

The next lemma is well known.

Lemma 4 *Given μ and ν two intensity measures on $\mathbb{X}$, the Poisson random measure $\eta_{\mu+\nu}$ with intensity $\mu + \nu$ decomposes into the sum*

$$\eta_{\mu+\nu} \simeq \eta_\mu \oplus \eta_\nu, \tag{41}$$

of a Poisson random measure η_μ with intensity $\mu(\mathrm{d}x)$ and an independent Poisson random measure η_ν with intensity $\nu(\mathrm{d}x)$.

Proof Taking F a random variable of the form (40) we have

$$\mathbb{E}[F(\eta_{\mu+\nu})] = e^{-\mu(\mathbb{X})-\nu(\mathbb{X})} \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} f_n(\{x_1, \ldots, x_n\}) \prod_{k=1}^{n} (\mu(\mathrm{d}x_k) + \nu(\mathrm{d}x_k)),$$

and

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} f_n(\{s_1, \ldots, s_n\}) \prod_{k=1}^{n} (\mu(\mathrm{d}s_k) + \nu(\mathrm{d}s_k)) \\
&\quad = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{l=0}^{n} \binom{n}{l} \int_{\mathbb{X}^n} f_n(\{s_1, \ldots, s_n\})\mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k)\nu(\mathrm{d}s_{k+1}) \cdots \nu(\mathrm{d}s_n) \\
&\quad = \sum_{n=0}^{\infty} \sum_{l=0}^{n} \frac{1}{(n-l)!l!} \int_{\mathbb{X}^n} f_n(\{s_1, \ldots, s_l, \ldots, s_n\})\mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_l)\nu(\mathrm{d}s_{l+1}) \cdots \nu(\mathrm{d}s_n) \\
&\quad = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{l=0}^{\infty} \frac{1}{l!} \int_{\mathbb{X}^{l+m}} f_{l+m}(\{s_1, \ldots, s_l, \ldots, s_{l+m}\}) \\
&\qquad\qquad \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_l)\nu(\mathrm{d}s_{l+1}) \cdots \nu(\mathrm{d}s_{l+m})
\end{aligned}$$

$$= e^{\mu(\mathbb{X})} \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\mathbb{X}^m} \mathbb{E}\left[\epsilon_{\mathfrak{s}_m}^{+} F(\eta_\mu)\right] \nu(\mathrm{d}s_1) \cdots \nu(\mathrm{d}s_m) \tag{42}$$

$$= e^{\mu(\mathbb{X})+\nu(\mathbb{X})} \mathbb{E}[F(\eta_\mu \oplus \eta_\nu)],$$

where $\epsilon_{\mathfrak{s}_m}^{+}$ is the addition operator defined on any random variable $F : \mathbf{N}_\sigma(\mathbb{X}) \to \mathbb{R}$ by

$$\epsilon_{\mathfrak{s}_m}^{+} F(\eta) = F(\eta + \delta_{s_1} + \cdots + \delta_{s_m}), \quad \eta \in \mathbf{N}_\sigma(\mathbb{X}), \quad s_1, \ldots, s_m \in \mathbb{X}, \tag{43}$$

and

$$\mathfrak{s}_m := (s_1, \ldots, s_m) \in \mathbb{X}^m, \quad m \geq 1.$$

□

In the course of the proof of Lemma 4 we have shown in (42) that

$$\mathbb{E}[F(\eta_{\mu+\nu})] = e^{-\mu(\mathbb{X})} \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\mathbb{X}^m} \mathbb{E}\left[\epsilon_{\mathfrak{s}_m}^{+} F(\eta_\nu)\right] \nu(\mathrm{d}s_1) \cdots \nu(\mathrm{d}s_m) = \mathbb{E}[F(\eta_\mu \oplus \eta_\nu)],$$

where $\epsilon_{\mathfrak{s}_k}^{+}$ is defined in (43).

In particular, by applying Lemma 4 above to $\mu(\mathrm{d}x)$ and $\nu(\mathrm{d}x) = f(x)\mu(\mathrm{d}x)$ with $f(x) \geq 0$ $\mu(\mathrm{d}x)$-a.e. we find that the Poisson random measure η with intensity $(1+f)\mathrm{d}\mu$ decomposes into the sum

$$\eta_{(1+f)\mathrm{d}\mu} \simeq \eta_{\mathrm{d}\mu} \oplus \eta_{f\mathrm{d}\mu},$$

of a Poisson random measure $\eta_{\mathrm{d}\mu}$ with intensity $\mu(\mathrm{d}x)$ and an independent Poisson random measure $\eta_{f\mathrm{d}\mu}$ with intensity $f(x)\mu(\mathrm{d}x)$.

In addition we have, using the shorthand notation $\mathbb{E}_\mu$ to denote the Poisson probability measure with intensity μ,

$$\mathbb{E}_{(1+f)\mathrm{d}\mu}[F] = e^{-\mu(\mathbb{X})} \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\mathbb{X}^m} \mathbb{E}_\mu\left[\epsilon_{\mathfrak{s}_m}^{+} F\right] f(s_1) \cdots f(s_n) \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_m). \tag{44}$$

The above identity extends to $f \in L^2(\mathbb{X})$ with $f > -1$, and when $f(x) \in (-1, 0)$, Relation (44) can be interpreted as a thinning of $\eta_{(1+f)\mathrm{d}\mu}$.

3.1.1 Mecke Identity

The following version of Mecke's identity [19], cf. also Relation (1.7) in [12], allows us to compute the first moment of the first order stochastic integral of a random integrand. In the sequel we use the expression "measurable process" to denote a real-valued measurable function from $\mathbb{X} \times \mathbf{N}_\sigma(\mathbb{X})$ into $\mathbb{R}$.

Proposition 1 *For* $u : \mathbb{X} \times \mathbf{N}_\sigma(\mathbb{X}) \longrightarrow \mathbb{R}$ *a measurable process we have*

$$\mathbb{E}_\mu\left[\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right] = \mathbb{E}_\mu\left[\int_{\mathbb{X}} u(x,\eta+\delta_x)\mu(\mathrm{d}x)\right], \tag{45}$$

provided

$$\mathbb{E}_\mu\left[\int_{\mathbb{X}} |u(x,\eta+\delta_x)|\mu(\mathrm{d}x)\right] < \infty.$$

Proof The proof is done when $\mu(\mathbb{X}) < \infty$. We take $u(x,\eta)$ written as

$$u(x,\eta) = \sum_{n=0}^{\infty} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(x; x_1,\ldots,x_n),$$

where $(x_1,\ldots,x_n) \longmapsto f_n(x; x_1,\ldots,x_n)$ is a symmetric integrable function of $\eta = \{x_1,\ldots,x_n\}$ when $\eta(\mathbb{X}) = n$, for each $n \geq 1$. We have

$$\begin{aligned}
&\mathbb{E}_\mu\left[\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right] \\
&= e^{-\mu(\mathbb{X})}\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=1}^{n}\int_{\mathbb{X}^n} f_n(x_i; x_1,\ldots,x_n)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n) \\
&= \sum_{n=1}^{\infty}\frac{e^{-\mu(\mathbb{X})}}{(n-1)!}\int_{\mathbb{X}^n} f_n(x; x_1,\ldots,x_{i-1},x,x_i,\ldots,x_{n-1})\mu(\mathrm{d}x)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_{n-1}) \\
&= e^{-\mu(\mathbb{X})}\sum_{n=0}^{\infty}\frac{1}{n!}\int_{\mathbb{X}^n}\int_{\mathbb{X}} f_{n+1}(x; x, x_1,\ldots,x_n)\mu(\mathrm{d}x)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n) \\
&= \mathbb{E}_\mu\left[\int_{\mathbb{X}} u(x,\eta+\delta_x)\mu(\mathrm{d}x)\right].
\end{aligned}$$

□

3.2 Multiple Stochastic Integrals

In this section we define the multiple Poisson stochastic integral (also called multiple Wiener–Itô integrals) using Charlier polynomials. We denote by "$\circ$" the symmetric tensor product of functions in $L^2(\mathbb{X})$, i.e., given $f_1, \ldots, f_d \in L^2(\mathbb{X})$ and $k_1, \ldots, k_d \geq 1$,

$$f_1^{\circ k_1} \circ \cdots \circ f_d^{\circ k_d}$$

denotes the symmetrization in $n = k_1 + \cdots + k_d$ variables of

$$f_1^{\otimes k_1} \otimes \cdots \otimes f_d^{\otimes k_d},$$

cf. Relation (1.27) in [12].

Definition 1 Consider $A_1, \ldots, A_d$ mutually disjoint subsets of $\mathbb{X}$ with finite μ-measure and $n = k_1+\cdots+k_d$, where $k_1, \ldots, k_d \geq 1$. The multiple Poisson stochastic integral of the function

$$\mathbf{1}_{A_1}^{\circ k_1} \circ \cdots \circ \mathbf{1}_{A_d}^{\circ k_d}$$

is defined by

$$I_n(\mathbf{1}_{A_1}^{\otimes k_1} \otimes \cdots \otimes \mathbf{1}_{A_d}^{\otimes k_d})(\eta) := \prod_{i=1}^{d} C_{k_i}(\eta(A_i), \mu(A_i)). \tag{46}$$

Note that by (27), Relation (46) actually coincides with Relation (1.26) in [12] and this recovers the fact that

$$\eta^{(k)}(A) := \#(\{(i_1, \ldots, i_k) \in \{1, \ldots, \eta(A)\}^k \ : \ i_l \neq i_m, \ 1 \leq l \neq m \leq k\})$$

defined in Relation (9) of [12] coincides with the falling factorial $(\eta(A))_{(k)}$ for $A \in \mathscr{X}$ such that $\mu(A) < \infty$.

See also [7, 31] for a more general framework for the expression of multiple stochastic integrals with respect to Lévy processes based on the combinatorics of the Möbius inversion formula.

From (28) and Definition 1 it can be shown that the multiple Poisson stochastic integral satisfies the isometry formula

$$\mathbb{E}[I_n(f_n)I_m(g_m)] = \mathbf{1}_{\{n=m\}}\langle f_n, g_m\rangle_{L^2(\mathbb{X}^n)}, \tag{47}$$

cf. Lemma 4 in [12], which allows one to extend the definition of I_n to any symmetric function $f_n \in L^2(\mathbb{X}^n)$, cf. also (52) below.

The generating series

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} C_n(\eta(A), \mu(A)) = e^{-\lambda\mu(A)}(1+\lambda)^{\eta(A)} = \psi_\lambda(\eta(A), \mu(A)),$$

cf. (25), admits a multivariate extension using multiple stochastic integrals.

Proposition 2 *For $f \in L^2(\mathbb{X}) \cap L^1(\mathbb{X})$ we have*

$$\xi(f) := \sum_{k=0}^{\infty} \frac{1}{n!} I_n(f^{\otimes n}) = \exp\left(-\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)\right) \prod_{x\in\eta} (1+f(x)). \tag{48}$$

Proof From (47) and an approximation argument it suffices to consider simple functions of the form

$$f = \sum_{k=1}^{m} a_k \mathbf{1}_{A_k},$$

by the multinomial identity (11) we have

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{1}{n!} I_n\left(\left(\sum_{k=1}^{m} a_k \mathbf{1}_{A_k}\right)^{\otimes n}\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{d_1+\cdots+d_m=n} \frac{n!}{d_1!\cdots d_m!} a_1^{d_1}\cdots a_m^{d_m} I_n\left(\mathbf{1}_{A_1}^{\otimes d_1} \circ \cdots \circ \mathbf{1}_{A_m}^{\otimes d_m}\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{d_1+\cdots+d_m=n} \frac{n!}{d_1!\cdots d_m!} a_1^{d_1}\cdots a_m^{d_m} \prod_{i=1}^{m} C_{d_i}(\eta(A_i), \mu(A_i)) \\
&= \prod_{i=1}^{m} \sum_{n=0}^{\infty} \frac{a_i^n}{n!} C_n(\eta(A_i), \mu(A_i)) \\
&= \prod_{i=1}^{m} \left(e^{-a_i\mu(A_i)}(1+a_i)^{\eta(A_i)}\right) \\
&= \exp\left(-\sum_{i=1}^{m} a_i\mu(A_i)\right) \prod_{i=1}^{m} (1+a_i)^{\eta(A_i)} \\
&= \exp\left(\sum_{i=1}^{m} a_i(\eta(A_i)-\mu(A_i))\right) \prod_{i=1}^{m} ((1+a_i)^{\eta(A_i)} e^{-a_i\eta(A_i)}). \qquad \square
\end{aligned}$$

The relation between $\xi(f)$ in (48) and the exponential functional in Lemma 5 of [12] is given by

$$\exp\left(\int_{\mathbb{X}} (e^{f(x)} - 1)\mu(\mathrm{d}x)\right) \xi(e^f - 1) = \exp\left(\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x)\right),$$

provided $e^f - 1 \in L^1(\mathbb{X}) \cap L^2(\mathbb{X})$.

3.3 $\mathcal{S}$-Transform

Given $f \in L^1(\mathbb{X}, \mu) \cap L^2(\mathbb{X}, \mu)$ with $f(x) > -1$ $\mu(\mathrm{d}x)$-a.e., we define the measure $\mathbb{Q}_f$ by its Girsanov density

$$\frac{\mathrm{d}\mathbb{Q}_f}{\mathrm{d}\mathbb{P}_\eta} = \xi(f) = \exp\left(-\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)\right) \prod_{x\in\mathbb{X}} (1 + f(x)), \tag{49}$$

where $\mathbb{P}_\eta$ is the Poisson probability measure with intensity $\mu(\mathrm{d}x)$. From (39), for F a bounded random variable we have the relation

$$\begin{aligned}
\mathbb{E}_\mu[F\xi(f)] &= \mathbb{E}_\mu\left[F \exp\left(-\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)\right) \prod_{x\in\eta}(1 + f(x))\right] \\
&= \exp\left(-\int_{\mathbb{X}} (1 + f(x))\mu(\mathrm{d}x)\right) \\
&\quad \times \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} F(\{s_1, \ldots, s_n\}) \prod_{k=1}^{n} (1 + f(s_k))\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_n) \\
&= \mathbb{E}[F(\eta_{(1+f)\mathrm{d}\mu})],
\end{aligned}$$

which shows the following proposition.

Proposition 3 *Under the probability $\mathbb{Q}_f$ defined by (49), the random measure η is Poisson with intensity $(1 + f)\mathrm{d}\mu$, i.e.,*

$$\mathbb{E}_\mu[F\xi(f)] = \mathbb{E}_{(1+f)\mathrm{d}\mu}[F]$$

for all sufficiently integrable random variables F.

The $\mathcal{S}$-transform (or Segal–Bargmann transform, see [14] for references) on the Poisson space is defined on bounded random variables F by

$$f \longmapsto \mathcal{S}F(f) := \mathbb{E}_{f\mathrm{d}\mu}[F] = \mathbb{E}_\mu[F\xi(f)]$$
$$= \mathbb{E}_\mu\left[F \exp\left(-\int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)\right) \prod_{x\in\eta}(1+f(x))\right],$$

for f bounded and vanishing outside a set of finite σ-measure in $\mathscr{X}$; Lemma 4 and Proposition 3 show that

$$\mathcal{S}F(f) = \mathbb{E}[F(\eta_{\mathrm{d}\mu} \oplus \eta_{f\mathrm{d}\mu})] \tag{50}$$
$$= e^{-\int_{\mathbb{X}} fd\mu}\mathbb{E}_\mu[F]$$
$$+e^{-\int_{\mathbb{X}} fd\mu}\sum_{k=1}^\infty \frac{1}{k!}\int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k}F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k),$$

where $\eta_{f\mathrm{d}\mu}$ is a Poisson random measure with intensity $f\mathrm{d}\mu$, independent of $\eta_{\mathrm{d}\mu}$, by Lemma 4. In the next proposition we use the finite difference operator

$$D_x := \epsilon^+_x - I, \qquad x \in \mathbb{X},$$

i.e.,

$$D_xF(\eta) = F(\eta+\delta_x) - F(\eta),$$

and apply a binomial transformation to get rid of the exponential term in (50). In the next proposition we let

$$D^k_{\mathfrak{s}_k} = D_{s_1}\cdots D_{s_k}, \qquad s_1,\ldots,s_k \in \mathbb{X},$$

and

$$\epsilon^+_{\mathfrak{s}_k} = \epsilon^+_{s_1}\cdots\epsilon^+_{s_k}, \qquad s_1,\ldots,s_k \in \mathbb{X},$$

as in (43), where

$$\mathfrak{s}_k = (s_1,\ldots,s_k) \in \mathbb{X}^k, \quad k \geq 1.$$

Proposition 4 *For any bounded random variable F and f bounded and vanishing outside a set of finite μ-measure in $\mathscr{X}$, we have*

$$\mathcal{S}F(f) = \mathbb{E}_\mu[F(\eta)\xi(f)] = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)\mathbb{E}_\mu\left[D^k_{\mathfrak{s}_k} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k). \tag{51}$$

Proof We apply a binomial transformation to the expansion (50). We have

$$\begin{aligned}
&\mathcal{S}F(f)\\
&= e^{-\int_{\mathbb{X}} f d\mu} \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\left(\int_{\mathbb{X}} f d\mu\right)^n \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\\
&= \sum_{m=0}^{\infty}\sum_{k=0}^{m} \frac{(-1)^{m-k}}{(m-k)!}\left(\int_{\mathbb{X}} f d\mu\right)^{m-k} \frac{1}{k!} \int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\\
&= \sum_{m=0}^{\infty} \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k} \int_{\mathbb{X}^m} f(s_1)\cdots f(s_m)\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_m)\\
&= \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\mathbb{X}^m} f(s_1)\cdots f(s_m)\mathbb{E}_\mu\left[D^m_{\mathfrak{s}_m} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_m).
\end{aligned}$$

□

By identification of terms in the expansions (48) and (51) we obtain the following result, which is equivalent (by (47) and duality) to the Stroock [32] formula, cf. also Theorem 2 in [12].

Corollary 1 *Given a bounded random variable F, for all $n \geq 1$ and all f bounded and vanishing outside a set of finite μ-measure in $\mathscr{X}$ we have*

$$\mathbb{E}_\mu\left[I_n(f^{\otimes n})F\right] = \int_{\mathbb{X}^n} f(s_1)\cdots f(s_n)\mathbb{E}_\mu\left[D^n_{\mathfrak{s}_n} F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_n). \tag{52}$$

Proof We note that (48) yields

$$\mathcal{S}F(f) = \mathbb{E}_{f\mathrm{d}\mu}[F] = \mathbb{E}_\mu[F\xi(f)] = \sum_{n=0}^{\infty} \frac{1}{n!}\mathbb{E}_\mu[FI_n(f^{\otimes n})],$$

and by Proposition 4 we have

$$\mathcal{S}F(f) = \mathbb{E}_{f\mathrm{d}\mu}[F] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} f(s_1)\cdots f(s_n)\mathbb{E}_\mu \left[D^n_{\mathfrak{s}_n} F\right] \mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_n),$$

and identify the respective terms of orders $n \geq 1$ in order to show (52). □

When $k = 1$, we have the integration by parts formula

$$\mathbb{E}_\mu[I_1(f)F] = \mathbb{E}_\mu\left[\int_{\mathbb{X}} f(s)D_sF\mu(\mathrm{d}s)\right].$$

Note that with the pathwise extension $I_k((Ff)^{\otimes k}) = F^k I_k(f^{\otimes k})$ of the multiple stochastic integral, (52) can be rewritten as the identity

$$\mathbb{E}_\mu[I_k((Ff)^{\otimes k})] = \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} f(s_1)\cdots f(s_k)D_{s_1}\cdots D_{s_k}(F^k)\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\right],$$

cf. also Proposition 4.1 of [26].

3.4 $\mathcal{U}$-Transform

The Laplace transform on the Poisson space (also called $\mathcal{U}$-transform, cf., e.g., §2 of [11]), is defined using the exponential functional of Lemma 5 of [12] by

$$f \longmapsto \mathcal{U}F(f) := \mathbb{E}_\mu\left[Fe^{\int_{\mathbb{X}} f\mathrm{d}\eta}\right] = e^{\int_{\mathbb{X}}(e^f-1)\mathrm{d}\mu}\mathbb{E}_\mu[F\xi(e^f-1)],$$

for f bounded and vanishing outside a set of finite μ-measure in $\mathscr{X}$, and will be useful for the derivation of general moment identities in Sect. 4.

Proposition 5 *Let F be a bounded random variable. We have*

$$\mathcal{U}F(f) = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}}\int_{\mathbb{X}^k} f^{|P_1^n|}(s_1)\cdots f^{|P_k^n|}(s_k)$$
$$\mathbb{E}_\mu\left[\epsilon^+_{\mathfrak{s}_k}F\right]\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k), \tag{53}$$

$f \in L^2(\mathbb{X}, \mu)$.

Proof Using the Faà di Bruno identity (13) or (16) we have

$$\sum_{n=0}^{\infty} \frac{1}{n!} \mathbb{E}_\mu \left[F \left(\int_{\mathbb{X}} f \mathrm{d}\eta \right)^n \right] = \mathbb{E}_\mu \left[F e^{\int_{\mathbb{X}} f \mathrm{d}\eta} \right] = e^{\int_{\mathbb{X}} (e^f - 1) \mathrm{d}\mu} \mathbb{E}_\mu \left[F \xi (e^f - 1) \right]$$

$$= e^{\int_{\mathbb{X}} (e^f - 1) \mathrm{d}\mu} \tag{54}$$

$$\times \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} (e^{f(s_1)} - 1) \cdots (e^{f(s_k)} - 1) \mathbb{E}_\mu \left[D^k_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \int_{\mathbb{X}^n} (e^{f(s_1)} - 1) \cdots (e^{f(s_n)} - 1) \mathbb{E}_\mu \left[D^k_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_n)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} (e^{f(s_1)} - 1) \cdots (e^{f(s_n)} - 1) \mathbb{E}_\mu \left[\epsilon^+_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \left(\sum_{n=1}^{\infty} \frac{f^n(s_1)}{n!} \right) \cdots \left(\sum_{n=1}^{\infty} \frac{f^n(s_k)}{n!} \right) \mathbb{E}_\mu \left[\epsilon^+_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=1}^{\infty} \sum_{\substack{d_1 + \cdots + d_k = n \\ d_1, \ldots, d_k \geq 1}} \int_{\mathbb{X}^k} \frac{f^{d_1}(s_1)}{d_1!} \cdots \frac{f^{d_k}(s_k)}{d_k!} \mathbb{E}_\mu \left[\epsilon^+_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{1}{k!}$$

$$\times \sum_{\substack{d_1 + \cdots + d_k = n \\ d_1, \ldots, d_k \geq 1}} \frac{n!}{d_1! \cdots d_k!} \int_{\mathbb{X}^k} f^{d_1}(s_1) \cdots f^{d_k}(s_k) \mathbb{E}_\mu \left[\epsilon^+_{\mathfrak{s}_k} F \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k),$$

where we applied the Faà di Bruno identity (13). □

In particular, by (54) we have

$$\mathbb{E}_\mu \left[I_m(g_m) e^{\int_{\mathbb{X}} f \mathrm{d}\eta} \right] \tag{55}$$

$$= \frac{1}{m!} \int_{\mathbb{X}^m} (e^{f(s_1)} - 1) \cdots (e^{f(s_m)} - 1) \mathbb{E}_\mu \left[D^+_{\mathfrak{s}_m} I_m(g_m) \right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_m)$$

$$= \int_{\mathbb{X}^m} (e^{f(s_1)} - 1) \cdots (e^{f(s_m)} - 1) g_m(s_1, \ldots, s_m) \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_m),$$

cf. Proposition 3.2 of [11].

4 Moment Identities and Invariance

The following cumulant-type moment identities have been extended to the Poisson stochastic integrals of random integrands in [28] through the use of the Skorohod integral on the Poisson space, cf. [23, 27]. These identities and their consequences on invariance have been recently extended to point processes with Papangelou intensities in [6], via simpler proofs based on an induction argument.

4.1 *Moment Identities for Random Integrands*

The moments of Poisson stochastic integrals of deterministic integrands have been derived in [2] by direct iterated differentiation of the Lévy–Khintchine formula or moment generating function

$$\mathbb{E}_\mu \left[\exp\left(\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x) \right) \right] = \exp\left(\int_{\mathbb{X}} (e^{f(x)} - 1)\mu(\mathrm{d}x) \right),$$

for f bounded and vanishing outside a set of finite μ-measure in $\mathscr{X}$. We also note that

$$\begin{aligned}
\mathbb{E}_\mu \left[\exp\left(\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x) \right) \right] &= \exp\left(\int_{\mathbb{X}} (e^{f(x)} - 1)\mu(\mathrm{d}x) \right) \\
&= \sum_{n=0}^\infty \frac{1}{n!} \left(\int_{\mathbb{X}} (e^{f(x)} - 1)\mu(\mathrm{d}x) \right)^n \\
&= \sum_{n=0}^\infty \frac{1}{n!} \int_{\mathbb{X}^n} (e^{f(x_1)} - 1) \cdots (e^{f(x_n)} - 1)\mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n) \\
&= \sum_{n=0}^\infty \frac{1}{n!} \int_{\mathbb{X}^n} (e^{f(x_1)} - 1) \cdots (e^{f(x_n)} - 1)\mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n) \\
&= \sum_{n=0}^\infty \frac{1}{n!} \int_{\mathbb{X}^n} \left(\sum_{k=1}^\infty \frac{f^k(x_1)}{k!} \right) \cdots \left(\sum_{k=1}^\infty \frac{f^k(x_n)}{k!} \right) \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n)
\end{aligned}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=1}^{\infty} \sum_{\substack{d_1+\cdots+d_n=k \\ d_1,\ldots,d_n \geq 1}} \int_{\mathbb{X}^n} \frac{f^{d_1}(x_1)}{d_1!} \cdots \frac{f^{d_n}(x_k)}{d_n!} \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_k)$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=0}^{k} \frac{1}{n!} \sum_{\substack{d_1+\cdots+d_n=k \\ d_1,\ldots,d_n \geq 1}} \frac{k!}{d_1! \cdots d_n!} \int_{\mathbb{X}^n} f^{d_1}(x_1) \cdots f^{d_n}(x_n) \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n),$$

where we applied the Faà di Bruno identity (13), showing that

$$\mathbb{E}_\mu \left[\left(\int_{\mathbb{X}} f(x) \eta(\mathrm{d}x) \right)^n \right] = \sum_{P_1^n,\ldots,P_a^n} \int_{\mathbb{X}^a} f^{|P_1^n|}(s_1) \mu(\mathrm{d}s_1) \cdots \int_{\mathbb{X}^a} f^{|P_a^n|}(s_a) \mu(\mathrm{d}s_a), \tag{56}$$

which recovers in particular (38).

The next Lemma 5 is a moment formula for deterministic Poisson stochastic integrals, and applies in particular in the framework of a change of measure given by a density F.

Lemma 5 *Let $n \geq 1$, $f \in \bigcap_{p=1}^{n} L^p(\mathbb{X}, \mu)$, and consider F a bounded random variable. We have*

$$\mathbb{E}_\mu \left[F \left(\int_{\mathbb{X}} f \mathrm{d}\eta \right)^n \right]$$
$$= \sum_{k=0}^{n} \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} \int_{\mathbb{X}^k} f^{|P_1^n|}(s_1) \cdots f^{|P_k^n|}(s_k) E\left[\epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ F\right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k).$$

Proof We apply Proposition 5 on the $\mathcal{U}$-transform, which reads

$$\mathbb{E}_\mu \left[F \exp\left(\int_{\mathbb{X}} f \mathrm{d}\eta \right) \right] = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbb{E}_\mu \left[F \left(\int_{\mathbb{X}} f \mathrm{d}\eta \right)^n \right]$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{1}{k!} \sum_{\substack{d_1+\cdots+d_k=n \\ d_1,\ldots,d_k \geq 1}} \frac{n!}{d_1! \cdots d_k!} \int_{\mathbb{X}^k} f^{d_1}(s_1) \cdots f^{d_k}(s_k)$$
$$\times \mathbb{E}\left[\epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ F\right] \mu(\mathrm{d}s_1) \cdots \mu(\mathrm{d}s_k).$$

□

Lemma 5 with $F = 1$ recovers the identity (38), and (by means of the complete Bell polynomials $A_n(b_1, \ldots, b_n)$ as in (30)) it can be used to compute the moments of stochastic integrals of deterministic integrands with respect to Lévy processes, cf. [18] for the case of subordinators.

Relation (55) yields

$$\mathbb{E}[FZ^n] = \sum_{k=0}^{n} S(n,k) \int_{A^k} \mathbb{E}\left[\epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ F\right] \mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k), \qquad n \in \mathbb{N}, \tag{57}$$

and when f is a deterministic function, Relation (54) shows that

$$\mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} f(x)\eta(\mathrm{d}x)\right)^n\right] = \sum_{P_1^n,\ldots,P_a^n} \int_{\mathbb{X}^a} f^{|P_1^n|}(s_1)\mu(\mathrm{d}s_1)\cdots\int_{\mathbb{X}^a} f^{|P_a^n|}(s_a)\mu(\mathrm{d}s_a),$$

which recovers (56).

Based on the following version of (57)

$$\mathbb{E}_\mu\left[F\eta(A)^n\right] = \sum_{k=0}^{n} S(n,k)\mathbb{E}\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ F\mathbb{1}_A(s_1)\cdots\mathbb{1}_A(s_k)\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\right] \tag{58}$$

and an induction argument we obtain the following Lemma 6, which can be seen as an elementary joint moment identity obtained by iteration of Lemma 58.

Lemma 6 *For $A_1,\ldots,A_p$ mutually disjoint bounded measurable subsets of $\mathbb{X}$ and $F_1,\ldots,F_p$ bounded random variables we have*

$$\begin{aligned}
&\mathbb{E}_\mu\left[(F_1\eta(A_1))^{n_1}\cdots\left(F_p\eta(A_p)\right)^{n_p}\right] \\
&= \sum_{k_1=0}^{n_1}\cdots\sum_{k_p=0}^{n_p} S(n_1,k_1)\cdots S(n_p,k_p) \\
&\times\mathbb{E}_\mu\left[\int_{\mathbb{X}^{k_1+\cdots+k_p}} \epsilon_{x_1}^+\cdots\epsilon_{x_{k_1+\cdots+k_p}}^+ (F_1^{n_1}\cdots F_p^{n_p}(\mathbb{1}_{A_1^{k_1}}\otimes\cdots\otimes\mathbb{1}_{A_p^{k_p}})\right. \\
&\left.(x_1,\ldots,x_{k_1+\cdots+k_p})\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_{k_1+\cdots+k_p})\right].
\end{aligned}$$

Lemma 6 allows us to recover the following moment identity, which can also be used for the computation of moments under a probability with density F with respect to $\mathbb{P}_\eta$.

Theorem 1 *Given F a random variable and $u : \mathbb{X} \times \mathbf{N}_\sigma(\mathbb{X}) \longrightarrow \mathbb{R}$ a measurable process we have*

$$\mathbb{E}_\mu \left[F \left(\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x) \right)^n \right] \tag{59}$$

$$= \sum_{P_1^n \cup \cdots \cup P_k^n = \{1,\ldots,n\}} \mathbb{E}_\mu \left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ (F u_{s_1}^{|P_1^n|} \cdots u_{s_k}^{|P_k^n|}) \mu(ds_1) \cdots \mu(ds_k) \right],$$

provided all terms in the above summations are $\mathbb{P}_\eta \otimes \mu^{\otimes k}$- integrable, $k = 1, \ldots, n$.

Proof We use the argument of Proposition 4.2 in [5] in order to extend Lemma 6 to (59). We start with $u : \mathbb{X} \times \mathbf{N}_\sigma(\mathbb{X}) \longrightarrow \mathbb{R}$ a simple measurable process of the form $u(x,\eta) = \sum_{i=1}^p F_i(\eta) \mathbb{1}_{A_i}(x)$ with disjoint sets $A_1, \ldots, A_p$. Using Lemma 6 we have

$$\mathbb{E}_\mu \left[\left(\sum_{i=1}^p F_i \int_{\mathbb{X}} \mathbf{1}_{A_i}(x)\ \eta(\mathrm{d}x) \right)^n \right] = \mathbb{E}_\mu \left[\left(\sum_{i=1}^p F_i \eta(A_i) \right)^n \right]$$

$$= \sum_{\substack{n_1+\cdots+n_p=n \\ n_1,\ldots,n_p \geq 0}} \frac{n!}{n_1! \cdots n_p!} \mathbb{E}_\mu \left[(F_1 \eta(A_1))^{n_1} \cdots \left(F_p \eta(A_p)\right)^{n_p} \right]$$

$$= \sum_{\substack{n_1+\cdots+n_p=n \\ n_1,\ldots,n_p \geq 0}} \frac{n!}{n_1! \cdots n_p!} \sum_{k_1=0}^{n_1} \cdots \sum_{k_p=0}^{n_p} S(n_1,k_1) \cdots S(n_p,k_p)$$

$$\mathbb{E}_\mu \left[\int_{\mathbb{X}^{k_1+\cdots+k_p}} \epsilon_{x_1}^+ \cdots \epsilon_{x_{k_1+\cdots+k_p}}^+ \left(F_1^{n_1} \cdots F_p^{n_p} \mathbb{1}_{A_1^{k_1}} \otimes \cdots \otimes \mathbb{1}_{A_p^{k_p}}(x_1, \ldots, x_{k_1+\cdots+k_p}) \right) \right.$$

$$\left. \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_{k_1+\cdots+k_p}) \right]$$

$$= \sum_{m=0}^n \sum_{\substack{n_1+\cdots+n_p=n \\ n_1,\ldots,n_p \geq 0}} \frac{n!}{n_1! \cdots n_p!} \sum_{\substack{k_1+\cdots+k_p=m \\ 1 \leq k_1 \leq n_1,\ldots,1 \leq k_p \leq n_p}} S(n_1,k_1) \cdots S(n_p,k_p)$$

$$\mathbb{E}_\mu \left[\int_{\mathbb{X}^m} \epsilon_{x_1}^+ \cdots \epsilon_{x_m}^+ \left(F_1^{n_1} \cdots F_p^{n_p} \mathbb{1}_{A_1^{k_1}} \otimes \cdots \otimes \mathbb{1}_{A_p^{k_p}}(x_1, \ldots, x_m) \right) \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_m) \right]$$

$$= \sum_{m=0}^n \sum_{\substack{n_1+\cdots+n_p=n \\ n_1,\ldots,n_p \geq 0}} \frac{n!}{n_1! \cdots n_p!} \sum_{\substack{I_1 \cup \cdots \cup I_p = \{1,\cdots,m\} \\ |I_1| \leq n_1,\ldots,|I_p| \leq n_p}} S(n_1,|I_1|) \cdots S(n_p,|I_p|) \frac{|I_1|! \cdots |I_p|!}{m!}$$

$$\mathbb{E}_\mu\left[\int_{\mathbb{X}^m}\epsilon^+_{x_1}\cdots\epsilon^+_{x_m}\left(F_1^{n_1}\cdots F_p^{n_p}\prod_{j\in I_1}\mathbb{1}_{A_1}(x_j)\cdots\prod_{j\in I_p}\mathbb{1}_{A_p}(x_j)\right)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_m)\right]$$

$$=\sum_{m=0}^{n}\sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\ldots,n\}}\sum_{i_1,\ldots,i_m=1}^{p}\mathbb{E}_\mu\left[\int_{\mathbb{X}^m}\epsilon^+_{x_1}\cdots\epsilon^+_{x_m}\left(F_{i_1}^{|P_1^n|}\mathbf{1}_{A_{i_1}}(x_1)\cdots F_{i_m}^{|P_m^n|}\mathbf{1}_{A_{i_m}}(x_m)\right)\right.$$
$$\left.\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_m)\right],$$

where in (60) we made changes of variables in the integral and, in (60), we used the combinatorial identity of Lemma 7 below with $\alpha_{i,j} = \mathbb{1}_{A_i}(x_j)$, $1 \le i \le p, 1 \le j \le m$, and $\beta_i = F_i$. The proof is concluded by using the disjointness of the A_i's in (60), as follows

$$\mathbb{E}_\mu\left[\left(\sum_{i=1}^{p}F_i\int_{\mathbb{X}}\mathbf{1}_{A_i}(x)\ \eta(\mathrm{d}x)\right)^n\right]$$
$$=\sum_{m=0}^{n}\sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\ldots,n\}}$$
$$\mathbb{E}_\mu\left[\int_{\mathbb{X}^m}\epsilon^+_{x_1}\cdots\epsilon^+_{x_m}\left(\sum_{i=1}^{p}\left(F_i^{|P_1^n|}\mathbf{1}_{A_i}(x_1)\right)\cdots\sum_{i=1}^{p}\left(F_i^{|P_m^n|}\mathbf{1}_{A_i}(x_m)\right)\right)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_m)\right]$$
$$=\sum_{m=0}^{n}\sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\ldots,n\}} \qquad (60)$$
$$\mathbb{E}_\mu\left[\int_{\mathbb{X}^m}\epsilon^+_{x_1}\cdots\epsilon^+_{x_m}\left(\left(\sum_{i=1}^{p}F_i\mathbf{1}_{A_i}(x_1)\right)^{|P_1^n|}\cdots\left(\sum_{i=1}^{p}F_i\mathbf{1}_{A_i}(x_m)\right)^{|P_m^n|}\right)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_m)\right].$$

The general case is obtained by approximating $u(x,\eta)$ with simple processes.

□

The next lemma has been used above in the proof of Theorem 1, cf. Lemma 4.3 of [5], and its proof is given for completeness.

Lemma 7 *Let $m, n, p \in \mathbb{N}$, $(\alpha_{i,j})_{1\le i\le p, 1\le j\le m}$ and $\beta_1, \ldots, \beta_p \in \mathbb{R}$. We have*

$$\sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\ge 0}} \frac{n!}{n_1!\cdots n_p!} \sum_{\substack{I_1\cup\cdots\cup I_p=\{1,\ldots,m\}\\ |I_1|\le n_1,\ldots,|I_p|\le n_p}} S(n_1,|I_1|)\cdots S(n_p,|I_p|)$$
$$\times \frac{|I_1|!\cdots|I_p|!}{m!}\beta_1^{n_1}\Big(\prod_{j\in I_1}\alpha_{1,j}\Big)\cdots\beta_p^{n_p}\Big(\prod_{j\in I_p}\alpha_{p,j}\Big)$$
$$= \sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\ldots,n\}} \sum_{i_1,\ldots,i_m=1}^{p} \beta_{i_1}^{|P_1^n|}\alpha_{i_1,1}\cdots\beta_{i_m}^{|P_m^n|}\alpha_{i_m,m}. \tag{61}$$

Proof Observe that (19) ensures

$$S(n,|I|)\beta^n\Big(\prod_{j\in I}\alpha_j\Big) = \sum_{\bigcup_{a\in I} P_a=\{1,\ldots,n\}} \prod_{j\in I}\left(\alpha_j\beta^{|P_j|}\right)$$

for all $\alpha_j, j \in I$, $\beta \in \mathbb{R}$, $n \in \mathbb{N}$. We have

$$\sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\ge 0}} \frac{n!}{n_1!\cdots n_p!} \sum_{\substack{I_1\cup\cdots\cup I_p=\{1,\ldots,m\}\\ |I_1|\le n_1,\ldots,|I_p|\le n_p}} S(n_1,|I_1|)\cdots S(n_p,[I_p|)$$
$$\frac{|I_1|!\cdots|I_p|!}{m!}\beta_1^{n_1}\Big(\prod_{j\in I_1}\alpha_{1,j}\Big)\cdots\beta_p^{n_p}\Big(\prod_{j\in I_p}\alpha_{p,j}\Big)$$
$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\ge 0}} \frac{n!}{n_1!\cdots n_p!} \sum_{\substack{I_1\cup\cdots\cup I_p=\{1,\ldots,m\}\\ |I_1|\le n_1,\ldots,|I_p|\le n_p}} \frac{|I_1|!\cdots|I_p|!}{m!}$$
$$\left(\sum_{\bigcup_{a\in I_1} P_a^1=\{1,\ldots,n_1\}} \prod_{j_1\in I_1}\left(\alpha_{1,j_1}\beta_1^{|P_{j_1}^1|}\right)\right)\cdots\left(\sum_{\bigcup_{a\in I_p} P_a^p=\{1,\ldots,n_p\}} \prod_{j_p\in J_p}\left(\alpha_{p,j_p}\beta_p^{|P_{j_p}^p|}\right)\right)$$
$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\ge 0}} \frac{n!}{n_1!\cdots n_p!} \sum_{\substack{I_1\cup\cdots\cup I_p=\{1,\ldots,m\}\\ |I_1|\le n_1,\ldots,|I_p|\le n_p}} \sum_{\bigcup_{a\in I_1} P_a^1=\{1,\ldots,n_1\}} \cdots \sum_{\bigcup_{a\in I_p} P_a^p=\{1,\ldots,n_p\}}$$
$$\frac{|I_1|!\cdots|I_p|!}{m!}\prod_{l=1}^{p}\prod_{j_l\in I_l}\left(\alpha_{l,j_l}\beta_l^{|P_{j_l}^l|}\right)$$
$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\ge 0}} \frac{n!}{n_1!\cdots n_p!} \sum_{\substack{I_1\cup\cdots\cup I_p=\{1,\ldots,m\}\\ |I_1|\le n_1,\ldots,|I_p|\le n_p}} \sum_{\bigcup_{a\in I_1} P_a^1=\{1,\ldots,n_1\}} \cdots \sum_{\bigcup_{a\in I_p} P_a^p=\{1,\ldots,n_p\}}$$

$$\frac{|I_1|!\cdots|I_p|!}{m!}\prod_{l=1}^{p}\prod_{j_l\in I_l}\alpha_{l,j_l}\prod_{l=1}^{p}\prod_{j_l\in I_l}\beta_l^{|P_{j_l}^l|}$$

$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\geq 0}}\frac{n!}{n_1!\cdots n_p!}\sum_{\substack{k_1+\cdots+k_p=m\\ 1\leq k_1\leq n_1,\ldots,1\leq k_p\leq n_p}}\sum_{i_1,\ldots,i_m=1}^{p}$$

$$\sum_{P_1^1\cup\cdots\cup P_{k_1}^1=\{1,\ldots,n_1\}}\cdots\sum_{P_{k_1+\cdots+k_{p-1}+1}^p\cup\cdots\cup P_{k_1+\cdots+k_p}^p=\{1,\ldots,n_p\}}\prod_{j=1}^{m}\left(\alpha_{i_j,j}\beta_{i_j}^{|P_j^{i_1}|+\cdots+|P_j^{i_m}|}\right)$$

$$= \sum_{P_1\cup\cdots\cup P_m=\{1,\ldots,n\}}\sum_{i_1,\ldots,i_m=1}^{p}\beta_{i_1}^{|P_1|}\alpha_{i_1,1}\cdots\beta_{i_m}^{|P_m|}\alpha_{i_m,m},$$

by a reindexing of the summations and the fact that the reunions of the partitions $P_1^j,\ldots,P_{|I_j|}^j$, $1\leq j\leq p$, of disjoint p subsets of $\{1,\ldots,m\}$ run the partition of $\{1,\ldots,m\}$ when we take into account the choice of the p subsets and the possible length k_j, $1\leq j\leq p$, of the partitions. □

As noted in [5], the combinatorial identity of Lemma 7 also admits a probabilistic proof. Namely given $Z_{\lambda\alpha_1},\ldots,Z_{\lambda\alpha_p}$ independent Poisson random variables with parameters $\lambda\alpha_1,\ldots,\lambda\alpha_p$ we have

$$\sum_{m=0}^{n}\lambda^m\sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\geq 0}}\frac{n!}{n_1!\cdots n_p!}\sum_{\substack{k_1+\cdots+k_p=m\\ k_1\leq n_1,\ldots,k_p\leq n_p}}S(n_1,k_1)\cdots S(n_p,k_p)\beta_1^{n_1}\alpha_1^{k_1}\cdots\beta_p^{n_p}\alpha_p^{k_p}$$

$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\geq 0}}\frac{n!}{n_1!\cdots n_p!}\sum_{k_1=0}^{n_1}S(n_1,k_1)(\lambda\alpha_1)^{k_1}\cdots\sum_{k_p=0}^{n_p}S(n_p,k_p)(\lambda\alpha_p)^{k_p}\beta_1^{n_1}\cdots\beta_p^{n_p}$$

$$= \sum_{\substack{n_1+\cdots+n_p=n\\ n_1,\ldots,n_p\geq 0}}\frac{n!}{n_1!\cdots n_p!}\mathbb{E}[Z_{\lambda\alpha_1}^{n_1}\cdots Z_{\lambda\alpha_p}^{n_p}]\beta_1^{n_1}\cdots\beta_p^{n_p}$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{p}\beta_i Z_{\lambda\alpha_i}\right)^n\right]$$

$$= \sum_{m=0}^{n}\lambda^m\sum_{P_1^n\cup\cdots\cup P_m^n=\{1,\ldots,n\}}\sum_{i_1,\ldots,i_m=1}^{p}\beta_{i_1}^{|P_1^n|}\alpha_{i_1}\cdots\beta_{i_m}^{|P_m^n|}\alpha_{i_m}, \qquad (62)$$

since the moment of order n_i of $Z_{\lambda\alpha_i}$ is given by (29) as

$$\mathbb{E}\left[Z_{\lambda\alpha_i}^{n_i}\right] = \sum_{k=0}^{n_i}S(n_i,k)(\lambda\alpha_i)^k.$$

The above relation (62) being true for all λ, this implies (61). Next we specialize the above results to processes of the form $u = \mathbf{1}_A$ where $A(\eta)$ is a random set.

Proposition 6 *For any bounded variable F and random set $A(\eta)$ we have*

$$\mathbb{E}_\mu\left[F\,(\eta(A))^n\right] = \sum_{k=0}^{n} S(n,k)\mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ (F\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(ds_1)\cdots\mu(ds_k)\right].$$

Proof We have

$$\begin{aligned}
\mathbb{E}_\mu\left[F\,(\eta(A))^n\right] &= \mathbb{E}_\mu\left[F\left(\int_{\mathbb{X}} \mathbf{1}_{A(\eta)}(x)\eta(\mathrm{d}x)\right)^n\right] \\
&= \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ (F\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\right] \\
&= \sum_{k=0}^{n} S(n,k)\mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ (F\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(\mathrm{d}s_1)\cdots\mu(\mathrm{d}s_k)\right].
\end{aligned}$$

□

We also have

$$\begin{aligned}
&\mathbb{E}_\mu\left[F\,(\eta(A))^n\right] \\
&= \sum_{k=0}^{n} S(n,k) \sum_{\Theta\subset\{1,\ldots,k\}} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} D_\Theta(F\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(ds_1)\cdots\mu(ds_k)\right] \\
&= \sum_{k=0}^{n} S(n,k) \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} D_{s_1}\cdots D_{s_l}(F\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(ds_1)\cdots\mu(ds_k)\right].
\end{aligned}$$

When $\mu(A(\eta))$ is deterministic this yields

$$\begin{aligned}
\mathbb{E}_\mu\left[(\eta(A))^n\right] &= \mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} \mathbf{1}_{A(\eta)}(x)\eta(\mathrm{d}x)\right)^n\right] \\
&= \sum_{k=0}^{n} S(n,k)\mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+ \cdots \epsilon_{s_k}^+ (\mathbf{1}_{A(\eta)}(s_1)\cdots\mathbf{1}_{A(\eta)}(s_k))\mu(ds_1)\cdots\mu(ds_k)\right]
\end{aligned}$$

$$= \sum_{k=0}^{n} S(n,k) \sum_{\Theta \subset \{1,\ldots,k\}} \mathbb{E}_\mu \left[\int_{\mathbb{X}^k} D_\Theta (\mathbf{1}_{A(\eta)}(s_1) \cdots \mathbf{1}_{A(\eta)}(s_k)) \mu(ds_1) \cdots \mu(ds_k) \right]$$

$$= \sum_{k=0}^{n} S(n,k) \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}_\mu \left[\int_{\mathbb{X}^k} D_{s_1} \cdots D_{s_l} (\mathbf{1}_{A(\eta)}(s_1) \cdots \mathbf{1}_{A(\eta)}(s_k)) \mu(ds_1) \cdots \mu(ds_l) \right]$$

$$= \sum_{k=0}^{n} S(n,k)$$

$$\times \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}_\mu \left[(\mu(A))^{k-l} \int_{\mathbb{X}^k} D_{s_1} \cdots D_{s_l} (\mathbf{1}_{A(\eta)}(s_1) \cdots \mathbf{1}_{A(\eta)}(s_l)) \mu(ds_1) \cdots \mu(ds_l) \right].$$

4.2 Joint Moment Identities

In this section we derive a joint moment identity for Poisson stochastic integrals with random integrands, which has been applied to mixing of interacting transformations in [29].

Proposition 7 *Let $u : \mathbb{X} \times \mathbf{N}_\sigma(\mathbb{X}) \longrightarrow \mathbb{R}$ be a measurable process and let $n = n_1 + \cdots + n_p$, $p \geq 1$. We have*

$$\mathbb{E}_\mu \left[\left(\int_{\mathbb{X}} u_1(x,\eta)\eta(\mathrm{d}x) \right)^{n_1} \cdots \left(\int_{\mathbb{X}} u_p(x,\eta)\eta(\mathrm{d}x) \right)^{n_p} \right] \tag{63}$$

$$= \sum_{k=1}^{n} \sum_{P_1^n,\ldots,P_k^n = \{1,\ldots,n\}} \mathbb{E}_\mu \left[\int_{\mathbb{X}^k} \epsilon^+_{x_1,\ldots,x_k} \left(\prod_{j=1}^{k} \prod_{i=1}^{p} u_i^{l^n_{i,j}}(x_j,\eta) \right) \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_k) \right],$$

where the sum runs over all partitions $P_1^n, \ldots, P_k^n$ of $\{1,\ldots,n\}$ and the power $l^n_{i,j}$ is the cardinal

$$l^n_{i,j} := |P^n_j \cap (n_1 + \cdots + n_{i-1}, n_1 + \cdots + n_i]|, \qquad i = 1,\ldots,k, \quad j = 1,\ldots,p,$$

for any $n \geq 1$ such that all terms in the right-hand side of (63) are integrable.

Proof We will show the modified identity

$$\mathbb{E}_\mu\left[F\left(\int_{\mathbb{X}} u_1(x,\eta)\eta(\mathrm{d}x)\right)^{n_1}\cdots\left(\int_{\mathbb{X}} u_p(x,\eta)\eta(\mathrm{d}x)\right)^{n_p}\right] \tag{64}$$
$$= \sum_{k=1}^n \sum_{P_1^n,\ldots,P_k^n=\{1,\ldots,n\}} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon^+_{x_1,\ldots,x_k}\left(F\prod_{j=1}^k\prod_{i=1}^p u_i^{l^n_{i,j}}(x_j,\eta)\right)\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k)\right],$$

for F a sufficiently integrable random variable, where $n = n_1 + \cdots + n_p$. For $p = 1$ the identity is Theorem 1. Next we assume that the identity holds at the rank $p \ge 1$. Replacing F with $F\left(\int_{\mathbb{X}} u_{p+1}(x,\eta)\eta(\mathrm{d}x)\right)^{n_{p+1}}$ in (64) we get

$$\mathbb{E}_\mu\left[F\left(\int_{\mathbb{X}} u_1(x,\eta)\eta(\mathrm{d}x)\right)^{n_1}\cdots\left(\int_{\mathbb{X}} u_{p+1}(x,\eta)\eta(\mathrm{d}x)\right)^{n_{p+1}}\right]$$
$$= \sum_{k=1}^n \sum_{P_1^n,\ldots,P_k^n} \int_{\mathbb{X}^k} \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k)$$
$$\mathbb{E}_\mu\left[\epsilon^+_{x_1,\ldots,x_k}\left(F\left(\int_{\mathbb{X}} u_{p+1}(x,\eta)\eta(\mathrm{d}x)\right)^{n_{p+1}}\prod_{j=1}^k\prod_{i=1}^p u_i^{l^n_{i,j}}(x_j,\eta)\right)\right]$$
$$= \sum_{k=1}^n \sum_{P_1^n,\ldots,P_k^n} \int_{\mathbb{X}^k} \mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} \epsilon^+_{x_1,\ldots,x_k} u_{p+1}(x,\eta)\eta(\mathrm{d}x) + \sum_{i=1}^k \epsilon^+_{x_1,\ldots,x_k} u_{p+1}(x_i,\eta)\right)^{n_{p+1}}\right.$$
$$\left.\epsilon^+_{x_1,\ldots,x_k}\left(F\prod_{j=1}^k\prod_{i=1}^p u_i^{l^n_{i,j}}(x_j,\eta)\right)\right]\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k)$$

$$= \sum_{k=1}^n \sum_{P_1^n,\ldots,P_k^n} \sum_{a_0+\cdots+a_k=n_{p+1}} \frac{n_{p+1}!}{a_0!\cdots a_k!}\int_{\mathbb{X}^k} \mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} \epsilon^+_{x_1,\ldots,x_k} u_{p+1}(x,\eta)\eta(\mathrm{d}x)\right)^{a_0}\right.$$
$$\left.\epsilon^+_{x_1,\ldots,x_k}\left(F\prod_{j=1}^k\left(u_{p+1}^{a_j}(x_j,\eta)\prod_{i=1}^p u_i^{l^n_{i,j}}(x_j,\eta)\right)\right)\right]\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k)$$
$$= \sum_{k=1}^n \sum_{P_1^n,\ldots,P_k^n} \sum_{a_0+\cdots+a_k=n_{p+1}} \frac{n_{p+1}!}{a_0!\cdots a_k!}\sum_{j=1}^{a_0}\int_{\mathbb{X}^{k+a_0}} \mathbb{E}_\mu\left[\sum_{Q_j^{a_0},\ldots,Q_j^{a_0}} \epsilon^+_{x_1,\ldots,x_{k+a_0}}\right.$$

$$\left(F \prod_{q=k+1}^{k+a_0} u_{p+1}^{|Q_q^{a_0}|}(x_q,\eta) \prod_{j=1}^{k} \left(u_{p+1}^{a_j}(x_j,\eta) \prod_{i=1}^{p} u_i^{l_{i,j}^n}(x_j,\eta) \right) \right) \Bigg] \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_{k+a_0})$$

$$= \sum_{k=1}^{n+n_{p+1}} \sum_{P_1^{n+n_{p+1}},\ldots,P_k^{n+n_{p+1}}} \mathbb{E}_\mu \Bigg[\int_{\mathbb{X}^k} \epsilon^+_{x_1,\ldots,x_k}$$

$$\left(F \prod_{l=1}^{k} \prod_{i=1}^{p+1} u_i^{l_{i,j}^{n+n_{p+1}}}(x_l,\eta) \right) \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k) \Bigg],$$

where the summation over the partitions $P_1^{n+n_{p+1}},\ldots,P_k^{n+n_{p+1}}$ of $\{1,\ldots,n+n_{p+1}\}$, is obtained by combining the partitions of $\{1,\ldots,n\}$ with the partitions $Q_j^{a_0},\ldots,Q_j^{a_0}$ of $\{1,\ldots,a_0\}$ and $a_1,\ldots,a_k$ elements of $\{1,\ldots,n_{p+1}\}$ which are counted according to $n_{p+1}!/(a_0!\cdots a_k!)$, with

$$l_{p+1,j}^{n+n_{p+1}} = l_{i,j}^n + a_j, \quad 1 \le j \le k, \qquad l_{p+1,j}^{n+n_{p+1}} = l_{i,j}^n + |Q_q^{a_0}|, \quad k+1 \le j \le k+a_0.$$

□

Note that when $n = 1$, (63) coincides with the classical Mecke [19] identity of Proposition 1.

When $n_1 = \cdots = n_p = 1$, the result of Proposition 7 reads

$$\mathbb{E}_\mu \left[\int_{\mathbb{X}} u_1(x,\eta)\eta(\mathrm{d}x) \cdots \int_{\mathbb{X}} u_p(x,\eta)\eta(\mathrm{d}x) \right]$$

$$= \sum_{k=1}^{n} \sum_{P_1^n,\ldots,P_k^n} \mathbb{E}_\mu \left[\int_{\mathbb{X}^k} \epsilon^+_{x_1,\ldots,x_k} \left(\prod_{j=1}^{k} \prod_{i \in P_j^n} u_i(x_j,\eta) \right) \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k) \right],$$

where the sum runs over all partitions $P_1^n,\ldots,P_k^n$ of $\{1,\ldots,n\}$, which coincides with the Poisson version of Theorem 3.1 of [6].

4.3 Invariance and Cyclic Condition

Using the relation $\epsilon_x^+ = D_x + I$, the result

$$\mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right)^n\right] = \sum_{k=1}^n \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} \epsilon_{s_1}^+\cdots\epsilon_{s_k}^+(u_{s_1}^{|P_1^n|}\cdots u_{s_k}^{|P_k^n|})\mu(ds_1)\cdots\mu(ds_k)\right]$$

of Theorem 1 can be rewritten as

$$\begin{aligned}
\mathbb{E}_\mu&\left[\left(\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right)^n\right] \\
&= \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \sum_{l=0}^k \binom{k}{l} \mathbb{E}_\mu\left[\int_{\mathbb{X}^k} D_{s_1}\cdots D_{s_l}(u_{s_1}^{|P_1^n|}\cdots u_{s_k}^{|P_k^n|})\mu(ds_1)\cdots\mu(ds_k)\right] \\
&= \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \sum_{l=0}^k \binom{k}{l} \\
&\quad\times \mathbb{E}_\mu\left[\int_{\mathbb{X}^l} D_{s_1}\cdots D_{s_l}\left(u_{s_1}^{|P_1^n|}\cdots u_{s_l}^{|P_l^n|}\int_{\mathbb{X}} u_{s_{l+1}}^{|P_1^n|}\mu(ds_{l+1})\cdots\int_{\mathbb{X}} u_{s_k}^{|P_k^n|}\mu(ds_k)\right)\right. \\
&\quad\left.\mu(ds_1)\cdots\mu(ds_l)\right].
\end{aligned}$$

Next is an immediate corollary of Theorem 1.

Corollary 2 *Suppose that*

(a) We have

$$D_{s_1}\cdots D_{s_k}(u_{s_1}\cdots u_{s_k}) = 0, \qquad s_1,\ldots,s_k\in\mathbb{X}, \quad k=1,\ldots,n. \tag{65}$$

(b) $\int_{\mathbb{X}} u_s^k\mu(\mathrm{d}s)$ *is deterministic for all* $k=1,\ldots,n$.

Then, $\displaystyle\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)$ *has (deterministic) cumulants* $\displaystyle\int_{\mathbb{X}} u^k(x,\eta)\mu(\mathrm{d}x)$*,* $k = 1,\ldots,n$.

Proof We have

$$\mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right)^n\right]$$
$$= \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \sum_{l=0}^{k-1}\binom{k}{l}\mathbb{E}_\mu\left[\int_{\mathbb{X}^k} D_{s_1}\cdots D_{s_l}(u_{s_1}^{|P_1^n|}\cdots u_{s_k}^{|P_k^n|})\mu(ds_1)\cdots\mu(ds_k)\right]$$
$$= \sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}} \sum_{l=0}^{k-1}\binom{k}{l}\mathbb{E}_\mu\Bigg[\int_{\mathbb{X}^{k-1}} D_{s_1}\cdots D_{s_l}\left(u_{s_1}^{|P_1^n|}\cdots u_{s_{k-1}}^{|P_{k-1}^n|}\int_{\mathbb{X}} u_s^{|P_k^n|}\mu(\mathrm{d}s)\right)$$
$$\mu(ds_1)\cdots\mu(ds_{k-1})\Bigg],$$

hence by a decreasing induction on k we can show that

$$\mathbb{E}_\mu\left[\left(\int_{\mathbb{X}} u(x,\eta)\eta(\mathrm{d}x)\right)^n\right] = \sum_{k=0}^{n}\sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}}\int_{\mathbb{X}^k} u_{s_1}^{|P_1^n|}\cdots u_{s_k}^{|P_k^n|}\mu(ds_1)\cdots\mu(ds_k)$$
$$= \sum_{k=0}^{n}\sum_{P_1^n\cup\cdots\cup P_k^n=\{1,\ldots,n\}}\int_{\mathbb{X}} u_{s_1}^{|P_1^n|}\mu(ds_1)\cdots\int_{\mathbb{X}} u_{s_k}^{|P_k^n|}\mu(ds_k).$$

Hence, by a decreasing induction we can show that the needed formula holds for the moment of order n, and for the moments of lower orders $k=1,..,n-1$. □

Note that from the relation

$$D_\Theta(u(x_1,\eta)\cdots u(x_k,\eta)) = \sum_{\Theta_1\cup\cdots\cup\Theta_k=\Theta} D_{\Theta_1}u(x_1,\eta)\cdots D_{\Theta_k}u(x_k,\eta), \tag{66}$$

where the above sum runs over all (possibly empty) subsets $\Theta_1,\ldots,\Theta_k$ of Θ, in particular when $\Theta=\{1,\ldots,k\}$ we get

$$D_{s_1}\cdots D_{s_k}(u(x_1,\eta)\cdots u(x_k,\eta)) = D_\Theta(u(x_1,\eta)\cdots u(x_k,\eta))$$
$$= \sum_{\Theta_1\cup\cdots\cup\Theta_k=\{1,\ldots,k\}} D_{\Theta_1}u(x_1,\eta)\cdots D_{\Theta_k}u(x_k,\eta),$$

where the sum runs over the (possibly empty) subsets $\Theta_1,\ldots,\Theta_k$ of $\{1,\ldots,k\}$. This shows that we can replace (65) with the condition

$$D_{\Theta_1}u(x_1,\eta)\cdots D_{\Theta_k}u(x_k,\eta) = 0, \tag{67}$$

for all $x_1, \dots, x_k \in \mathbb{X}$ and all (nonempty) subsets $\Theta_1, \dots, \Theta_k \subset \{x_1, \dots, x_n\}$, such that $\Theta_1 \cup \cdots \cup \Theta_n = \{1, \dots, n\}$, $k = 1, 2, \dots, n$. See Proposition 3.3 of [5] for examples of random mappings that satisfy Condition (67).

Acknowledgement This work was supported by the Singapore MOE Tier 2 grant MOE2012-T2-2-033 ARC 3/13.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, vol. 55, 9th edn. Dover Publications, New York (1972)
2. Bassan, B., Bona, E.: Moments of stochastic processes governed by Poisson random measures. Comment. Math. Univ. Carol. **31**(2), 337–343 (1990)
3. Bernstein, M., Sloane, N.J.A.: Some canonical sequences of integers. Linear Algebr. Appl. **226/228**, 57–72 (1995)
4. Boyadzhiev, K.: Exponential polynomials, Stirling numbers, and evaluation of some gamma integrals. Abstr. Appl. Anal. Art. ID 168672, 18 (2009)
5. Breton, J.-C., Privault, N.: Factorial moments of point processes. Stoch. Process. Appl. **124**, 3412–3428 (2014)
6. Decreusefond, L., Flint, I.: Moment formulae for general point processes. J. Funct. Anal. **267**, 452–476 (2014)
7. Farré, M., Jolis, M., Utzet, F.: Multiple Stratonovich integral and Hu-Meyer formula for Lévy processes. Ann. Probab. **38**(6), 2136–2169 (2010)
8. Fristedt, B., Gray, L.: A Modern Approach to Probability Theory. Probability and its Applications. Birkhäuser, Boston (1997)
9. Gerstenkorn, T.: Relations between the crude, factorial and inverse factorial moments. In: Transactions of the Ninth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Prague, 1982), vol. A, pp. 247–252. Reidel, Dordrecht (1983)
10. Isserlis, L.: On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika **12**(1–2), 134–139 (1918)
11. Ito, Y.: Generalized Poisson functionals. Probab. Theory Relat. Fields **77**, 1–28 (1988)
12. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
13. Last, G., Penrose, M.: Poisson process Fock space representation, chaos expansion and covariance inequalities. Probab. Theory Relat. Fields **150**(3–4), 663–690 (2011)
14. Lee, Y.-J., Privault, N., Shih, H.-H.: Clark-Ocone formula by the S-transform on the Poisson white noise space. Commun. Stoch. Anal. **6**, 513–524 (2012)
15. Leonov, V., Shiryaev, A.: On a method of calculation of semi-invariants. Theory Probab. Appl. **4**, 319–329 (1959)
16. Lukacs, E.: Applications of Faà di Bruno's formula in mathematical statistics. Am. Math. Mon. **62**, 340–348 (1955)
17. Lukacs, E.: Characteristic Functions, 2nd edn. Hafner Publishing Co., New York (1970)
18. Madan, D., Yor, M.: Moments of Wiener integrals for subordinators. Electron. Commun. Probab. **17**(55), 8 (2012)
19. Mecke, J.: Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **9**, 36–58 (1967)
20. Nardo, E.D., Senato, D.: Umbral nature of the Poisson random variables. Algebraic Combinatorics and Computer Science, pp. 245–266. Springer, Milan (2001)

21. Peccati, G., Taqqu, M.: Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation. Bocconi & Springer Series, vol. 7, Springer, Milan (2011)
22. Pitman, J.: Combinatorial stochastic processes. Lectures from the 32nd Summer School on Probability Theory Held in Saint-Flour. Lecture Notes in Mathematics, vol. 1875, 7–24 July 2002. Springer, Berlin (2006)
23. Privault, N.: Moment identities for Poisson-Skorohod integrals and application to measure invariance. C. R. Math. Acad. Sci. Paris **347**, 1071–1074 (2009)
24. Privault, N.: Stochastic Analysis in Discrete and Continuous Settings. Lecture Notes in mathematics, vol. 1982. Springer, Berlin (2009)
25. Privault, N.: Generalized Bell polynomials and the combinatorics of Poisson central moments. Electron. J. Comb. **18**(1), Paper 54, 10 (2011)
26. Privault, N.: Girsanov identities for Poisson measures under quasi-nilpotent transformations. Ann. Probab. **40**(3), 1009–1040 (2012)
27. Privault, N.: Invariance of Poisson measures under random transformations. Ann. Inst. Henri Poincare Probab. Stat. **48**(4), 947–972 (2012)
28. Privault, N.: Moments of Poisson stochastic integrals with random integrands. Probab. Math. Stat. **32**(2), 227–239 (2012)
29. Privault, N.: Mixing of Poisson random measures under interacting transformations. arXiv:1301.0672 (2013)
30. Roman, S.: The umbral calculus. Pure and Applied Mathematics, vol. 111. Academic, New York (1984) [Harcourt Brace Jovanovich Publishers]
31. Rota, G.-C., Wallstrom, T.: Stochastic integrals: a combinatorial approach. Ann. Probab. **25**(3), 1257–1283 (1997)
32. Stroock, D.: Homogeneous chaos revisited. In: Séminaire de Probabilités XXI. Lecture Notes in Mathematics, vol. 1247. Springer, Berlin (1987)
33. Thiele, T.: On semi invariants in the theory of observations. Kjöbenhavn Overs. 135–141 (1899)

Variational Analysis of Poisson Processes

Ilya Molchanov and Sergei Zuyev

Abstract The expected value of a functional $F(\eta)$ of a Poisson process η can be considered as a function of its intensity measure μ. The paper surveys several results concerning differentiability properties of this functional on the space of signed measures with finite total variation. Then, necessary conditions for μ being a local minima of the considered functional are elaborated taking into account possible constraints on μ, most importantly the case of μ with given total mass a. These necessary conditions can be phrased by requiring that the gradient of the functional (being the expected first difference $F(\eta + \delta_x) - F(\eta)$) is constant on the support of μ. In many important cases, the gradient depends only on the local structure of μ in a neighbourhood of x and so it is possible to work out the asymptotics of the minimising measure with the total mass a growing to infinity. Examples include the optimal approximation of convex functions, clustering problem and optimal search. In non-asymptotic cases, it is in general possible to find the optimal measure using steepest descent algorithms which are based on the obtained explicit form of the gradient.

1 Preliminaries

The importance of Poisson point processes for modelling various phenomena is impossible to overestimate. Perhaps, this comes from the fact that, despite being among the simplest mathematically tractable models, Poisson point processes enjoy a great degree of flexibility: indeed, the parameter characterising their distribution is a generic "intensity" measure, which roughly describes the density of the process points. It is amazing how many intriguing and deep properties such a seemingly

I. Molchanov
Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
e-mail: ilya.molchanov@stat.unibe.ch

S. Zuyev (✉)
Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden
e-mail: sergei.zuyev@chalmers.se

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_3

simple model enjoys, and how new ones are constantly being discovered, as this monograph readily shows. Because the distribution of a Poisson point process is determined by its intensity measure, altering the measure changes the distribution which, in many cases, is a result of performing a certain transformation of the phase space or of the point configurations. Such approach is taken, for instance, in perturbation analysis of point process driven systems (see, e.g., [10] and the references therein) or in differential geometry of configuration spaces, see, e.g., [1] or [29].

Rather than considering a change of the parameter measure induced by transformations of the phase space, we take a more general approach by changing the parameter measure directly. A control over this change is made possible by a linear structure of the set of measures itself as we describe in detail below.

The main subject of our study is a Poisson point process on a phase space $\mathbb{X}$. Although it can be defined on a very general measurable phase space, for some results below we shall need a certain topological structure, so we assume from now on that $\mathbb{X}$ is a Polish space with its Borel σ-algebra $\mathscr{X}$. The distribution of a point process is a probability measure on $(\mathbf{N}, \mathscr{N})$, where $\mathbf{N}$ is the set of locally finite counting measures on $\mathscr{X}$ called *configurations* and $\mathscr{N}$ is the minimal σ-algebra that makes all the mappings $\varphi \mapsto \varphi(B)$ measurable for any $B \in \mathscr{X}$. Any $\varphi \in \mathbf{N}$ can be represented as a sum of Dirac measures: $\varphi = \sum_i \delta_{x_i}$, where $\delta_x(B) = \mathbb{1}_B(x)$ for every $B \in \mathscr{X}$ and not necessarily all x_i's are distinct.

Let μ be a σ-finite measure on $(\mathbb{X}, \mathscr{X})$. A point process η is Poisson with *intensity measure* μ, if for any sequence of disjoint sets $B_1, \ldots, B_n \in \mathscr{X}$, $n \geq 1$, the counts $\eta(B_1), \ldots, \eta(B_n)$ are independent Poisson $\mathrm{Po}(\mu(B_1)), \ldots, \mathrm{Po}(\mu(B_n))$ distributed random variables. The distribution of the Poisson point process with intensity measure μ will be denoted by $\mathbb{P}_\mu$ with the corresponding expectation $\mathbb{E}_\mu$. The term intensity measure is explained by the fact that, due to the definition, one has $\mathbb{E}\eta(B) = \mu(B)$ for any $B \in \mathscr{X}$. Notice that the Poisson process is finite, i.e. all its configurations with probability 1 contain only a finite number of points, if and only if its intensity measure is finite, that is $\mu(\mathbb{X}) < \infty$.

In what follows, we study the changes in the distributional characteristics of functionals of a configuration, under perturbations of the intensity measure which we first assume finite. Recall that a signed measure ν can be represented as the difference $\nu = \nu^+ - \nu^-$ of two non-negative measures with disjoint supports (the *Lebesgue decomposition*) and that the total variation of ν is defined as $\|\nu\| = \nu^+(\mathbb{X}) + \nu^-(\mathbb{X})$. Consider the set $\tilde{\mathbf{M}}_\mathrm{f}$ of all *signed* measures on $\mathscr{X}$ with a finite total variation, and define operations of addition and multiplication by setting $(\mu + \nu)(B) = \mu(B) + \nu(B)$ and $(t\mu)(B) = t\mu(B)$ for any $B \in \mathscr{X}$. Endowed with the total variation norm, $\tilde{\mathbf{M}}_\mathrm{f}$ becomes a Banach space and the set $\mathbf{M}_\mathrm{f}$ of *finite non-negative measures* is a *pointed cone*, i.e. a set closed under addition and multiplication by non-negative numbers, see, e.g., [6, III.7.4].

Given a function $F : \mathbf{N} \mapsto \mathbb{R}$ of a configuration, its expectation $\mathbb{E}_\mu F(\eta)$ with respect to the distribution $\mathbb{P}_\mu$ of a finite Poisson process η can be regarded as a function of the intensity measure μ and hence as a function on $\mathbf{M}_\mathrm{f}$. Therefore, there is a reason to consider functions on $\tilde{\mathbf{M}}_\mathrm{f}$ and their analytical properties in general.

2 Variational Analysis on Measures

Recall that a function f on a Banach space $\mathbb{B}$ is called *strongly* or *Fréchet differentiable* at $x \in \mathbb{B}$ if

$$f(x+y) = f(x) + L(x)[y] + o(\|y\|),$$

where $L(x)[\cdot] : \mathbb{B} \mapsto \mathbb{B}$ is a bounded linear functional called a *differential*. A function f is called *weakly* or *Gateaux differentiable* at $x \in \mathbb{B}$ if for every $y \in \mathbb{B}$ there exists a limit

$$\partial_y f(x) = \lim_{t \downarrow 0} t^{-1}[f(x+ty) - f(x)]$$

which can be called the *directional derivative* of f along the vector y. Strong differentiability implies that all weak derivatives also exist and that $\partial_y f(x) = L(x)[y]$. The converse is not true even for $\mathbb{B} = \mathbb{R}$. The same definitions apply to functions of a signed measure with finite total variation, since $\tilde{\mathbf{M}}_\mathrm{f}$ is a Banach space. A very wide class of differentiable functions of a measure possess a differential which has a form of an integral so that

$$f(\mu + \nu) = f(\mu) + \int_{\mathbb{X}} g(x; \mu)\nu(\mathrm{d}x) + o(\|\nu\|), \; \nu \in \tilde{\mathbf{M}}_\mathrm{f}$$

for some function $g(\cdot\,; \mu)$ called a *gradient function*. This name comes from the fact that when $\mathbb{X} = \{1, \dots, d\}$ is a finite set, $\tilde{\mathbf{M}}_\mathrm{f}$ is isomorphic to $\mathbb{R}^d$ and $g(\cdot\,; \mu) = (g_1(\mu), \dots, g_d(\mu))$ is a usual gradient, since

$$f(\mu + \nu) = f(\mu) + \langle g(\cdot\,; \mu), \nu \rangle + o(\|\nu\|), \; \mu, \nu \in \mathbb{R}^d.$$

In line with this, we shall use from now on the notation $\langle f, \nu \rangle$ for the integral $\int f \, \mathrm{d}\nu$. Not all differentiable functions of measures possess a gradient function (unless $\mathbb{X}$ is finite), but all practically important functions usually do. Notably, the expectation $\mathbb{E}_\mu F(\eta)$ as a function of $\mu \in \mathbf{M}_\mathrm{f}$ does possess a gradient function, as we will see in the next section. So, it is not a severe restriction to assume that a differentiable function of a measure possesses a gradient function, as we often do below.

The differentiability provides a useful tool for optimisation of functions. Necessary conditions for a local optimum are based on the notion of a tangent cone.

Definition 1 The *tangent cone* to a set $\mathbb{A} \subset \tilde{\mathbf{M}}_\mathrm{f}$ at point $\nu \in \mathbb{A}$ is the set of all signed measures that appear as limits of $\eta_n \in \tilde{\mathbf{M}}_\mathrm{f}$ where $\nu + t_n \eta_n \in \mathbb{A}$ for all n and $t_n \downarrow 0$.

A first order necessary condition for an optimum in a constrained optimisation now takes the following form.

Theorem 1 *Assume $\mathbb{A} \subseteq \tilde{\mathbf{M}}_{\mathrm{f}}$ is closed and convex in the total variation norm and that f is continuous on $\mathbb{A}$ and strongly differentiable at $\nu^* \in \mathbb{A}$. If ν^* provides a local minimum in the constrained optimisation problem*

$$f(\nu) \to \inf \quad \textit{subject to } \nu \in \mathbb{A},$$

then

$$L(\nu^*)[\theta] \geq 0 \quad \textit{for all } \theta \in T_{\mathbb{A}}(\nu^*). \tag{1}$$

The proof of this general fact can be found, e.g., in [3] for the case of a constraint set with non-empty interior. For the purpose of optimisation with respect to the intensity measure, the main constraint set is the cone $\mathbf{M}_{\mathrm{f}}$ of non-negative measures. However, $\mathbf{M}_{\mathrm{f}}$ does *not* have interior points unless $\mathbb{X}$ is finite. The non-emptiness assumption on the interior was first dropped in [4, Theorem 4.1.(i)]. The next result proved in [19] characterises the tangent cone to $\mathbf{M}_{\mathrm{f}}$.

Theorem 2 *The tangent cone to the set $\mathbf{M}_{\mathrm{f}}$ at $\mu \in \mathbf{M}_{\mathrm{f}}$ is the set of signed measures for which the negative part of their Lebesgue decomposition is absolutely continuous with respect to μ:*

$$T_{\mathbf{M}_{\mathrm{f}}}(\mu) = \{\theta \in \tilde{\mathbf{M}}_{\mathrm{f}} : \ \theta^- \ll \mu\}.$$

Assume now that f possesses a gradient function and μ^* provides a local minimum on the constrain set $\mathbb{A} = \mathbf{M}_{\mathrm{f}}$. Applying necessary condition (1) with $\theta = \delta_x$ we immediately get that

$$L(\mu^*)[\delta_x] = g(x; \mu^*) \geq 0 \quad \text{for all } x \in \mathbb{X}.$$

Now letting θ be $-\mu^*$ restricted onto an arbitrary Borel $B \in \mathscr{X}$ leads to

$$L(\mu^*)[\delta_x] = \langle g(\,\cdot\,; \mu^*)\mathbb{1}_B, \mu^*\rangle \leq 0\,.$$

Combining both inequalities proves the following result.

Theorem 3 *Assume that $\mu^* \in \mathbf{M}_{\mathrm{f}}$ provides a local minimum to f on $\mathbf{M}_{\mathrm{f}}$ and that f possesses a gradient function $g(\,\cdot\,; \mu^*)$ at μ^*. Then $g(\,\cdot\,; \mu^*) = 0$ μ^*-almost everywhere on $\mathbb{X}$ and $g(x; \mu^*) \geq 0$ for all $x \in \mathbb{X}$.*

By considering an appropriate Lagrange function, one can generalise this statement to the case of optimisation over $\mathbf{M}_{\mathrm{f}}$ with additional constraints. Before we formulate the result, we need a notion of regularity.

Definition 2 Let Y be a Banach space and $\mathbb{A} \subseteq \mathbf{M}_\mathrm{f}$, $C \subseteq Y$ be closed convex sets. Let $f: \tilde{\mathbf{M}}_\mathrm{f} \mapsto \mathbb{R}$ and $H: \tilde{\mathbf{M}}_\mathrm{f} \mapsto Y$ be strongly differentiable. A measure $\nu \in \tilde{\mathbf{M}}_\mathrm{f}$ is called *regular* for the optimisation problem

$$f(\nu) \to \inf \quad \text{subject to } \nu \in \mathbb{A},\ H(\nu) \in C,$$

if $0 \in \mathrm{core}\big(H(\nu) + L_H(\nu)[\mathbb{A} - \nu] - C\big)$, where L_H is the differential of H and core(B) for $B \subseteq Y$ is the set $\{b \in B : \ \forall y \in Y\ \exists t_1 \text{ such that } b + ty \in B\ \forall t \in [0, t_1]\}$. For $Y = \mathbb{R}^d$, core(B) is just the interior of the set $B \subseteq \mathbb{R}^d$.

Consider the most common case of a finite number of equality and inequality constraints. In this case $Y = \mathbb{R}^k$ and $C = \{0\}^m \times \mathbb{R}_-^{k-m}$, $m \le k$, so that we have the following optimisation problem:

$$f(\mu) \to \inf \quad \text{subject to}$$
$$\begin{cases} \mu \in \mathbf{M}_\mathrm{f} \\ H_i(\mu) = 0,\ i = 1, \dots, m \\ H_j(\mu) \le 0,\ j = m+1, \dots, k \end{cases} \tag{2}$$

for some function $H: \mathbf{M}_\mathrm{f} \mapsto \mathbb{R}^k$. The following result and its generalisations can be found in [19].

Theorem 4 *Let μ^* be a regular (in the sense of Definition 2) local minimum for the problem (2) for a function f which is continuous on $\mathbf{M}_\mathrm{f}$ and strongly differentiable at μ^* with a gradient function $g(x;\mu^*)$. Let $H = (H_1, \dots, H_k)$ also be strongly differentiable at μ^* with a gradient function $h(x;\mu) = (h_1(x;\mu), \dots, h_k(x;\mu))$. Then there exist* Langrange multipliers $u = (u_1, \dots, u_k)$ *with $u_j \le 0$ for those $j \in \{m+1, \dots, k\}$ for which $H_j(\mu^*) = 0$ and $u_j = 0$ if $H_j(\mu^*) < 0$, such that*

$$\begin{cases} g(x;\mu^*) = \sum_{i=1}^k u_i h_i(x;\mu^*) & \mu^* - a.e.\ x \in \mathbb{X}, \\ g(x;\mu^*) \ge \sum_{i=1}^k u_i h_i(x;\mu^*) & \text{for all } x \in \mathbb{X}. \end{cases}$$

When the functions f and H possess gradient functions, as in Theorem 4 above, the regularity condition becomes the so-called *Mangasarian–Fromowitz constraint qualification*, that is, a linear independence of the gradients $h_1(\,\cdot\,;\mu^*), \dots, h_k(\,\cdot\,;\mu^*)$ and the existence of a signed measure $\zeta \in \tilde{\mathbf{M}}_\mathrm{f}$ such that

$$\begin{cases} \langle h_i, \zeta \rangle = 0 & \text{for all } i = 1, \dots, m; \\ \langle h_j, \zeta \rangle < 0 & \text{for all } j \in \{m+1, \dots, k\} \text{ for which } H_j(\mu^*) = 0. \end{cases} \tag{3}$$

Without inequality constraints, (3) holds trivially for ζ being the zero-measure and we come to the following important corollary giving, the first-order necessary condition for optimisation with a fixed total mass.

Theorem 5 *Let f be continuous on $\mathbf{M}_f$ and strongly differentiable at $\mu^* \in \mathbf{M}_f$ with a gradient function $g(x;\mu^*)$. If μ^* is a local minimum in the constrained optimisation problem*

$$f(\mu) \to \inf \quad \text{subject to}$$

$$\begin{cases} \mu \in \mathbf{M}_f \\ \mu(\mathbb{X}) = a > 0, \end{cases} \tag{4}$$

then there exists a real u such that

$$\begin{cases} g(x;\mu^*) = u & \mu^* - a.e.\ x \in \mathbb{X}, \\ g(x;\mu^*) \geq u & \text{for all } x \in \mathbb{X}. \end{cases} \tag{5}$$

3 Analyticity of the Expectation

The linear structure on the set of measures described in the previous section makes it possible to put analysis of variations of the intensity measure in the general framework of differential calculus on a Banach space. In this section we fix a functional $F: \mathbf{N} \mapsto \mathbb{R}$ on the configuration space and regard its expectation $\mathbb{E}_\mu F(\eta)$ as a function of a measure μ. To explain the idea, we first consider a bounded functional F and the Banach space $\mathbf{M}_f$ of finite measures and then discuss extensions to a wider class of functionals and to infinite measures.

It is a well-known fact that for a Poisson process η with a finite intensity measure μ, the conditional distribution of its points given their total number $\eta(\mathbb{X}) = n$ corresponds to n points independently drawn from the distribution $(\mu(\mathbb{X}))^{-1}\mu$. This observation, after applying the total probability formula, gives rise to the following expression for the expectation:

$$\mathbb{E}_\mu F(\eta) = F(\emptyset) + e^{-\mu(\mathbb{X})} \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} F(\delta_{x_1} + \cdots + \delta_{x_n})\, \mu(\mathrm{d}x_1) \ldots \mu(\mathrm{d}x_n), \tag{6}$$

where $\emptyset$ stands for the null measure.

Substituting $\mu \leftarrow (\mu + \nu)$ for a signed measure $\nu \in \tilde{\mathbf{M}}_f$ such that $\mu + \nu \in \mathbf{M}_f$ into (6),

$$\begin{aligned} \mathbb{E}_{\mu+\nu} F(\eta) &= e^{-\mu(\mathbb{X})}(1 - \nu(\mathbb{X}) + o(\|\nu\|)) \\ &\quad \times \Big[F(\emptyset) + \sum_{n=1} \frac{1}{n!} \int_{\mathbb{X}^n} F(\sum_{i=1}^{n} \delta_{x_i})\, (\mu+\nu)(\mathrm{d}x_1) \ldots (\mu+\nu)(\mathrm{d}x_n)\Big] \end{aligned}$$

$$= \mathbb{E}_\mu F + e^{-\mu(\mathbb{X})} \sum_{n=1} \frac{n}{n!} \int\limits_{\mathbb{X}^n} F(\sum_{i=1}^{n} \delta_{x_i})\, \mu(\mathrm{d}x_1) \dots \mu(\mathrm{d}x_{n-1}) \nu(\mathrm{d}x_n)$$

$$- \nu(\mathbb{X}) e^{-\mu(\mathbb{X})} \sum_{n=0} \frac{1}{n!} \int\limits_{\mathbb{X}^n} F(\sum_{i=1}^{n} \delta_{x_i})\, \mu(\mathrm{d}x_1) \dots \mu(\mathrm{d}x_n) + o(\|\nu\|).$$

Thus

$$\begin{aligned}
&\mathbb{E}_{\mu+\nu} F(\eta) - \mathbb{E}_\mu F(\eta) \\
&= e^{-\mu(\mathbb{X})} \sum_{n=0} \frac{1}{n!} \int\limits_{\mathbb{X}^{n+1}} F(\sum_{i=1}^{n} \delta_{x_i} + \delta_x)\, \mu(\mathrm{d}x_1) \dots \mu(\mathrm{d}x_n) \nu(\mathrm{d}x) \\
&\quad - e^{-\mu(\mathbb{X})} \sum_{n=0} \frac{1}{n!} \int\limits_{\mathbb{X}^{n+1}} F(\sum_{i=1}^{n} \delta_{x_i})\, \mu(\mathrm{d}x_1) \dots \mu(\mathrm{d}x_n) \nu(\mathrm{d}x) \ + o(\|\nu\|) \\
&= \mathbb{E}_\mu \int\limits_{\mathbb{X}} [F(\eta + \delta_x) - F(\eta)]\, \nu(\mathrm{d}x) + o(\|\nu\|).
\end{aligned}$$

Denoting by D_x the *difference operator* $D_x F(\eta) = F(\eta + \delta_x) - F(\eta)$, we see that

$$\mathbb{E}_{\mu+\nu} F - \mathbb{E}_\mu F = \langle \mathbb{E}_\mu D.F, \nu \rangle + o(\|\nu\|).$$

Since F is bounded, so is $\mathbb{E}_\mu D.F$, hence $\mathbb{E}_\mu F$ is strongly differentiable on $\mathbf{M}_\mathrm{f}$ with the gradient function $\mathbb{E}_\mu D_x F$.

Using the infinite series Taylor expansion in $\nu(\mathbb{X})$, one can extend the above argument to show not only differentiability, but also *analyticity* of $\mathbb{E}_\mu F$ as a function of μ. Introduce iterations of the operator D_x by setting $D^0 F = F$, $D^1_{x_1} F = D_{x_1} F$, $D^n_{x_1,\dots,x_n} F = D_{x_n}(D^{n-1}_{x_1,\dots,x_{n-1}} F)$ so that

$$D^n_{x_1,\dots,x_n} F(\eta) = \sum_{J \subseteq \{1,2,\dots,n\}} (-1)^{n-|J|} F\big(\eta + \sum_{j\in J} \delta_{x_j}\big),$$

as it can be easily checked.

Theorem 6 *Assume that there exists a constant* $b > 0$ *such that* $|F\big(\sum_{i=1}^n \delta_{x_i}\big)| \le b^n$ *for all* $n \ge 0$ *and* $(x_1, \dots, x_n) \in \mathbb{X}^n$. *Then* $\mathbb{E}_\mu F(\eta)$ *is analytic on* $\mathbf{M}_\mathrm{f}$ *and*

$$\mathbb{E}_{\mu+\nu} F = \sum_{n=0}^{\infty} \frac{1}{n!} \int\limits_{\mathbb{X}^n} \mathbb{E}_\mu D^n_{x_1,\dots,x_n} F\big(\eta + \sum_{i=1}^{n} \delta_{x_i}\big)\, \nu(\mathrm{d}x_1) \dots \nu(\mathrm{d}x_n), \tag{7}$$

where the term corresponding to $n = 0$ *is, by convention,* $\mathbb{E}_\mu F(\eta)$.

The proof can be found in [20]. Notice that the integral above is an n-linear form of the n-th product measure (the n-th differential) and that

$$\mathbb{E}_\mu D^n_{x_1,\dots,x_n} F(\eta) = \sum_{m=0}^{n} (-1)^{n-m} \binom{n}{m} \mathbb{E}_\mu F\big(\eta + \sum_{j=1}^{m} \delta_{x_j}\big)$$

because of the symmetry with respect to permutations of $x_1, \dots, x_n$.

3.1 Margulis–Russo Type Formula for Poisson Process

An important case of perturbations of the intensity measure is when the increment is proportional to the measure itself. So fix a $\mu \in \mathbf{M}_\mathrm{f}$ and consider $\nu = t\mu$ for a small $t \in (-1, 1)$. Substituting this into (7) gives a power series in t:

$$\mathbb{E}_{\mu+t\mu} F = \sum_{n=0}^{\infty} \frac{t^n}{n!} \int_{\mathbb{X}^n} \mathbb{E}_\mu D^n_{x_1,\dots,x_n} F\big(\eta + \sum_{i=1}^{n} \delta_{x_i}\big)\, \mu(\mathrm{d}x_1) \dots \mu(\mathrm{d}x_n).$$

In particular,

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathbb{E}_{s\mu} F(\eta) = \int_{\mathbb{X}} \mathbb{E}_{s\mu} D_x F(\eta)\, \mu(\mathrm{d}x) = \int_{\mathbb{X}} \mathbb{E}_{s\mu} [F(\eta + \delta_x) - F(\eta)]\, \mu(\mathrm{d}x).$$

Let $F(\eta) = \mathbb{1}_\Xi(\eta)$ be an indicator of some event Ξ. The integration in the last expression can be restricted to the (random) set $\Upsilon(\eta) = \{x \in \mathbb{X} : \mathbb{1}_\Xi(\eta + \delta_x) \neq \mathbb{1}_\Xi(\eta)\}$ leading to

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathbb{P}_{s\mu}(\Xi) = \mathbb{E}_{s\mu} \int_{\mathbb{X}} \mathbb{1}_\Xi(\eta + \delta_x) \mathbb{1}_{\Upsilon(\eta)}(x)\, \mu(\mathrm{d}x) - \mathbb{E}_{s\mu} \int_{\mathbb{X}} \mathbb{1}_\Xi(\eta) \mathbb{1}_{\Upsilon(\eta)}(x)\, \mu(\mathrm{d}x).$$

The last term is obviously $\mathbb{E}_{s\mu} \mathbb{1}_\Xi(\eta) \mu(\Upsilon(\eta))$. For the first one, we apply the Refined Campbell theorem together with the Mecke formula

$$\mathbb{E}_\mu \int_{\mathbb{X}} f(x, \eta) \eta(\mathrm{d}x) = \mathbb{E}_\mu \int_{\mathbb{X}} f(x, \eta + \delta_x) \mu(\mathrm{d}x)$$

valid for any measurable $f : \mathbb{X} \times \mathscr{X} \mapsto \mathbb{R}_+$ which characterises the Poisson process, see, e.g., Propositions 13.1.IV and 13.1.VII in [5] and [12, Sect. 1.1]. Using (3.1),

$$\mathbb{E}_{s\mu} \int_{\mathbb{X}} \mathbb{1}_\Xi(\eta + \delta_x) \mathbb{1}_{\Upsilon(\eta)}(x)\, \mu(\mathrm{d}x) = \frac{1}{s} \int_{\mathbb{X}} \mathbb{1}_\Xi(\eta) \mathbb{1}_{\{x:\, \mathbb{1}_\Xi(\eta) \neq \mathbb{1}_\Xi(\eta - \delta_x)\}}(x)\, \eta(\mathrm{d}x).$$

Combining all together,

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{P}_{s\mu}(\Xi) = \frac{1}{s}\mathbb{E}_{s\mu}\mathbb{1}_{\Xi}(\eta)N_{\Xi}(\eta) - \mathbb{E}_{s\mu}\mathbb{1}_{\Xi}(\eta)V_{\Xi}(\eta). \tag{8}$$

Here $V_{\Xi}(\eta) = \mu\{x \in \mathbb{X} : \mathbb{1}_{\Xi}(\eta + \delta_x) \neq \mathbb{1}_{\Xi}(\eta)\}$ is the μ-content of the set where adding a new point to configuration η would change the occurrence of Ξ, so the elements of this set are called *pivotal locations* for event Ξ in configuration η. While $N_{\Xi}(\eta) = \int_{\mathbb{X}} \mathbb{1}\{x \in \eta : \mathbb{1}_{\Xi}(\eta) \neq \mathbb{1}_{\Xi}(\eta - \delta_x)\}\eta(\mathrm{d}x)$, in the case of non-atomic μ, is equal to the number of points in configuration η whose removal would affect the occurrence of Ξ. Such configuration points are called *pivotal points* for event Ξ in configuration η. This geometric interpretation is a key to usefulness of this formula which is a counterpart of the Margulis–Russo formula for Bernoulli fields proved in [13] and independently in [31]. Identity (8) was shown in [34] in more restrictive settings.

Let us mention two useful implications of (8):

$$\frac{\mathrm{d}}{\mathrm{d}s}\log \mathbb{P}_{s\mu}(\Xi) = \frac{1}{s}\mathbb{E}_{s\mu}[N_{\Xi}(\eta) \mid \Xi] - \mathbb{E}_{s\mu}[V_{\Xi}(\eta) \mid \Xi]$$

obtained by dividing both parts by $\mathbb{P}_{s\mu}(\Xi)$, and consequently,

$$\mathbb{P}_{s_2\mu}(\Xi) = \mathbb{P}_{s_1\mu}(\Xi)\exp\Big\{\int_{s_1}^{s_2} \mathbb{E}_{s\mu}[s^{-1}N_{\Xi}(\eta) - V_{\Xi}(\eta) \mid \Xi]\,\mathrm{d}s\Big\}$$

providing a way to control the change in the probability of an event in terms of the control over the number of pivotal points versus the μ-content of the pivotal locations.

3.2 Infinite Measures

To extend the formula (7), or at least its first k-th term expansion, to infinite mass measures one must put additional assumptions on the functional F, as there are examples of a bounded functional whose expectation is, however, not differentiable. A notable example is the indicator that the origin belongs to an infinite cluster in a Boolean model of spheres in $\mathbb{R}^d$, $d \geq 2$. Its expectation is the density of the infinite cluster, which is *not* differentiable at the percolation threshold.

One possible approach for a locally compact phase space is to consider a growing sequence of compact sets $\{\mathbb{X}_n\}$ such that $\cup_n \mathbb{X}_n = \mathbb{X}$, and the corresponding restrictions η_n of the Poisson process η onto $\mathbb{X}_n$ are finite point processes. If $F(\eta_n)$ converges to $F(\eta)$ (such functionals are called *continuous at infinity*), then by controlling this convergence it is possible to ensure that the corresponding derivatives also converge. This approach was adopted in [20] where, in particular, it was shown that if F is bounded and continuous at infinity, then (7) holds for a σ-finite μ and a finite ν such that $\mu + \nu$ is a positive measure, see [20, Theorem 2.2]. Note that the indicator function that the origin is in an infinite cluster is not continuous at infinity.

A more subtle method is based on the Fock space representation (see the survey by Last [12], in this volume) and it makes it possible to extend the expansion formula to square-integrable functionals. Consider two σ-finite non-negative measures μ and another measure ρ dominating their sum $\lambda = \mu + \nu$. Denote by h_λ and h_ν the corresponding Radon–Nikodym densities. The following result is proved in [11].

Theorem 7 *Assume that*

$$\langle (1 - h_\nu)^2, \rho\rangle + \langle (1 - h_\lambda)^2, \rho\rangle < \infty.$$

Let F be such that $\mathbb{E}_\rho F(\eta)^2 < \infty$. Then (7) holds, all the integrals there exist and the series converges absolutely.

Perhaps, the most important case is when the increment measure ν is absolutely continuous with respect to μ with the corresponding density h_ν. Then the above theorem implies that for F such that $\mathbb{E}_{\mu+\nu} F^2(\eta) < \infty$, condition $\langle h_\nu (1 + h_\nu)^{-1}, \mu\rangle < \infty$ is sufficient for (7) to hold.

Note an interesting fact on the validity of the expansion formula. Each general increment measure ν can be represented as $\nu = \nu_1 + \nu_2$, where ν_1 is absolutely continuous with respect to μ and ν_2 is orthogonal to it. In order for (7) to hold for *all bounded F*, it is necessary that $\nu_2(\mathbb{X}) < \infty$! This and other results on the infinite measure case can be found in [11].

4 Asymptotics in the High-Intensity Setting

Consider the minimisation problem

$$f(\mu) = \mathbb{E}_\mu F(\eta) \to \inf \quad \text{subject to } \mu \in \mathbf{M}_\mathrm{f} \text{ and } \mu(\mathbb{X}) = a\,, \tag{9}$$

where F is a functional satisfying the conditions of Theorem 6. For simplicity, we consider only the case of a fixed total mass and refer to [20] for more general cases.

It is rarely possible to find analytic solution to (9), but Theorem 5 opens a possibility to use gradient descent type methods in order to numerically solve it as described later in Sect. 5. However, when the total mass a is large, in many cases it is possible to come up with asymptotic properties of the optimal measure that solves the optimisation problem (9) for a that grows to the infinity.

The key idea is to rescale the optimal measure around some point x, so it looks like proportional to the Lebesgue measure. In the case of a stationary point process, it is then easier to calculate the first difference in order to equate it to a constant, so to satisfy the necessary condition (5) for the minimum.

Assume that $\mathbb{X}$ is a compact subset of $\mathbb{R}^d$ that coincides with the closure of its interior and let $\gamma_a^x(y) = x + a^{1/d}(y - x)$ denote the rescaling around the point $x \in \mathbb{R}^d$, so that the image configuration $\gamma_a^x\eta$ consists of points $\gamma_{a^{-1}}^x x_i$ for $\eta = \{x_i\}$. Consider a solution to (9) which we represent in the form $a\mu_a$ for some probability measure μ_a. In particular,

$$\mathbb{E}_{a\mu_a} F(\eta) = \mathbb{E}_{\hat{\mu}_a^x} F(\gamma_a^x \eta)\,,$$

where $\hat{\mu}_a^x(\cdot) = a\mu_a(\gamma_{a^{-1}}^x \cdot)$. Assume that μ_a is absolutely continuous with density p_a with respect to the Lebesgue measure ℓ_d. Then $\hat{\mu}_a^x(\cdot)$ has density $p_a(\gamma_{a^{-1}}^x y)$ on $\gamma_a^x\mathbb{X}$. The key idea is that in some situations the expected first difference

$$\mathbb{E}_{a\mu_a} D_x(\eta) = \mathbb{E}_{\hat{\mu}_a^x} D_x(\gamma_a^x \eta) \propto g(a)\mathbb{E}_{p(x)\ell_d} \Gamma(x;\eta)$$

for a function $\Gamma(x;\eta)$ that depends on η locally in a possibly random neighbourhood of x, a normalising function g and a function p that corresponds to a limit of p_a in a certain sense. Then, the gradient function used in Theorem 5 can be calculated for a stationary Poisson process with intensity $p(x)$ which is generally easier.

To make precise the local structure of $\Gamma(x;\eta)$, we need the concept of a stopping set, that is a multidimensional analogue of a stopping time, see [35]. Let $\mathscr{A}_B$ be the σ-algebra generated by random variables $\{\eta(C)\}$ for Borel $C \subset B$. A random compact set S is called a *stopping set* if $\{S \subseteq K\} \in \mathscr{A}_K$ for any compact set K in $\mathbb{R}^d$. The stopping σ-algebra is the collection of events $A \in \mathscr{A}$ such that $A \cap \{S \subseteq K\} \in \mathscr{A}_K$ for all compact K.

The following result is proved in [20].

Theorem 8 *Let $a\mu_a$ be a measure solving (4) for the fixed total mass a. Assume that for an interior point x of $\mathbb{X}$ the following condition holds.*

(M) For all sufficiently large a, μ_a is absolutely continuous with respect to ℓ_d with densities p_a, and there exists a finite double limit

$$\lim_{y\to x,\, a\to\infty} p_a(y) = p(x) > 0\,. \tag{10}$$

Furthermore, assume that for the same x, the first difference D_xF satisfies the following conditions.

(D) For some positive function $g(a)$, the random variable

$$\Gamma_a = \Gamma_a(x;\eta) = D_x(\gamma_a^x\eta)/g(a)$$

converges to $\Gamma = \Gamma(x;\eta)$ as $a \to \infty$ for almost all realisations of the stationary Poisson process η with unit intensity, and

$$0 < \mathbb{E}_{p(x)\ell_d}\Gamma(x;\eta) < \infty\,.$$

(L) *There exist a family of stopping sets $S_a = S_a(x;\eta)$ and a stopping set $S(x;\eta)$ such that $\Gamma_a(x;\eta)$ is $\mathscr{A}_{S_a}$-measurable for all sufficiently large a; $\Gamma(x;\eta)$ is $\mathscr{A}_S$-measurable; and for every compact set W containing x in its interior*

$$\mathbb{1}_{S_a(x;\eta)\subseteq W} \to \mathbb{1}_{S(x;\eta)\subseteq W} \quad \text{as } a \to \infty$$

for almost all realisations of a stationary unit intensity Poisson process η.

(UI) *There exists a compact set W containing x in its interior such that*

$$\lim_{a\to\infty,\, n\to\infty} \mathbb{E}_{\hat{\mu}_z^x} |\Gamma_a(x;\eta)| \mathbb{1}_{S_a \subseteq W} = 0$$

and there exists a constant $M = M(W,b)$ such that $|\Gamma_a(x;\eta)| \le M$ for all sufficiently large a and η such that $S_a(x;\eta) \subset \gamma_b^x W$.

Then

$$\lim_{a\to\infty} |\mathbb{E}_{\hat{\mu}_a^x} \Gamma_a(x;\eta) - \mathbb{E}_{p(x)\ell_d} \Gamma(x;\eta)| = 0$$

and

$$\lim_{a\to\infty} \frac{\mathbb{E}_{a\mu_a} D_x F}{\mathbb{E}_{ap(x)\ell_d} D_x F} = 1\,.$$

The uniform integrability condition *(UI)* can be efficiently verified for stopping sets S_a and S that satisfy the condition $\ell_d(B) \ge \alpha \ell_d(S_a)$ for some fixed α and almost all x, see [20, Theorem 5.4]. We now show how this theorem applies to various problems of a practical interest.

4.1 Approximation of Functions

Consider a strictly convex function $f(x)$, $x \in [a,b] \subset \mathbb{R}$, and its linear spline approximation $s(x;\eta)$ built on the grid of points $a \le x_1 \le x_2 \le \cdots \le x_N \le b$, where $\{x_1,\dots,x_N\}$ form a Poisson point process η on $[a,b]$. Since the end-points are included as the spline knots, the spline approximation is well defined even if η is empty. The quality of approximation is measured in the L^1-distance as

$$F(\eta) = \int_a^b (s(x;\eta) - f(x))\mathrm{d}x\,.$$

If, instead of a Poisson process η, one takes a set of deterministic points, the problem of determining the best locations of those points has been considered in [14] (in relation to approximation of convex sets), see also [32]. It is well known that the empirical probability measure generated by the best deterministic points converges weakly to the measure with density proportional to $f''(x)^{1/3}$.

If η is a Poisson process of total intensity a, then the optimisation problem aims to determine the asymptotic behaviour of the measure μ_a such that the intensity measure $\mu = a\mu_a$ minimises $\mathbb{E}_\mu F(\eta)$. The key observation is that the first difference $D_x F(\eta)$ equals the area of the triangle with vertices at $(x,f(x))$, $(x^-,f(x^-))$ and $(x^+,f(x^+))$, where x^- and x^+ are left and right neighbours to x from η. Denoting $r_x^- = x - x^-$ and $r_x^+ = x^+ - x$, we arrive at the expected first difference (the gradient function) given by

$$g(x;\mu) = \mathbb{E}_\mu D_x F = -f(x)[\mathbb{E}_\mu r_x^- + \mathbb{E} r_x^+] + \mathbb{E}_\mu r_x^- \mathbb{E}_\mu f(x + r_x^+) + \mathbb{E}_\mu r_x^+ \mathbb{E}_\mu f(x + r_x^-)\,.$$

If μ is an optimal measure, then the strict convexity and continuity properties imply that (5) holds for all $x \in [a,b]$. It is easy to write down the distributions of r_x^- and r_x^+ in terms of μ. Then the requirement $g(x;\mu) = \text{const}$ turns into a system of four differential equations. However, one is interested in the asymptotic solution when a is large, so the high intensity framework is very much relevant in this setting. Notice that here

$$\Gamma(x;\eta) = -\frac{1}{4} f''(x) r_x^+ r_x^- (r_x^+ + r_x^-)$$

depends only on the stopping set $[x^-, x^+]$ that shrinks to $\{x\}$ as the total mass a of the measure $\mu = a\mu_a$ grows. If μ is proportional to the Lebesgue measure ℓ_1, then it is easy to calculate the first difference explicitly as

$$\mathbb{E}_{p(x)\ell_1} \Gamma(x;\eta) \propto -f''(x) p(x)^{-3}\,.$$

By Theorem 8, if (10) holds, then it is possible to equate the right-hand side to a constant, so that the density of the optimal measure μ_a is asymptotically proportional to $f''(x)^{1/3}$, exactly as it is in the deterministic case. The same argument applies to a strictly convex function $f(x)$, for x taken from a convex compact subset of $\mathbb{R}^d$, and leads to the asymptotically optimal measure with density proportional to $K(x)^{1/(2+d)}$, where $K(x)$ is the Gaussian curvature of f at point x, see [18]. The multidimensional optimal approximation results for deterministic sets of points (including also the Bezier approximation) are also studied in [16].

4.2 *Clustering*

Consider the data set $\{y_1, \dots, y_m\}$ in $\mathbb{R}^d$. One of the objectives in the cluster analysis consists in determining cluster centres $\eta = \{x_1, \dots, x_k\} \subset \mathbb{R}^d$ for some given k. Each cluster centre x_i is associated with the data points (also referred to as daughter points) which are nearest to it, i.e. lie within the corresponding *Voronoi cell* $C_{x_i}(\eta)$

(see, e.g., [26] for the definition and properties of the Voronoi tessellations). The cluster centres can be determined using the Ward-type criterion by minimising

$$F(\eta) = \sum_{x_i \in \eta} \sum_{y_j \in C_{x_i}(\eta)} \|x_i - y_j\|^2 ,$$

which is also the trace of the pooled within groups sum of squares matrix. In view of this criterion function, the optimal set of k cluster centres is also called the k-means of the data, see [25] for further references on this topic. In most applications, the number k is predetermined and then a steepest descent algorithm is employed to find the cluster centres. It should be noted that the functional $F(\eta)$ is not convex and so the descent algorithms might well end up in a local rather than a global minimum.

Alternatively, if the cluster centres are regarded as points of a Poisson point process with intensity measure μ and the mean of $F(\eta)$ is taken as an objective function, then

$$\mathbb{E}_\mu F(\eta) = \mathbb{E}_\mu \left[\sum_{x_i \in \eta} \sum_{y_j \in C_{x_i}(\eta)} \|x_i - y_j\|^2 \right] = \sum_{j=1}^{m} \mathbb{E}_\mu \rho(y_j, \eta)^2 ,$$

where $\rho(y, \eta)$ is the Euclidean distance from y to the nearest point of η. Since η can be empty, we have to assign a certain (typically large) value u to $\rho(y, \emptyset)$. Since η is a Poisson process, it is easy to compute the latter expectation in order to arrive at

$$\mathbb{E}_\mu F(\eta) = \sum_{j=1}^{m} \int_0^{u^2} \exp\{-\mu(B_{\sqrt{t}}(y_j)\}\mathrm{d}t , \tag{11}$$

which is a convex functional of μ. Since taking the expectation in the Poissonised variant of the clustering problem yields a convex objective function, the steepest descent algorithm applied in this situation would always converge to the global minimum. The optimal measure μ can be termed as the solution of the P-means problem.

In the asymptotic setting, it is assumed that the total mass a of the optimal measure $a\mu_a$ is growing to infinity and the data points are sampled from a probability distribution with density p_ν, so that the empty configurations η are no longer relevant and the objective function becomes

$$\mathbb{E}_\mu F(\eta) = \int_{\mathbb{R}^d} \mathbb{E}_\mu [\rho(y, \eta)^2] p_\nu(y)\mathrm{d}y .$$

Adding an extra cluster point x affects only the data points within the so-called Voronoi flower of x, see [26]. The Voronoi flower is a stopping set that satisfies the conditions of Theorem 8. Since $\mathbb{E}_{a\ell_d} D_x F$ is proportional to $p_\nu(x)a^{-1-2/d}$, the high intensity solution has the density proportional to $p_\nu(y)^{d/(d+2)}$.

A similar problem appears in the telecommunication setting, where the data points y_j represent the customers and $x_1, \dots, x_k$ are the locations of server stations. If the connection cost of a customer to the server is proportional to the β-power of the Euclidean distance between them (so that $\beta = 2$ in the clustering application), then the density of the high intensity solution is proportional to $p_\nu(y)^{d/(d+\beta)}$, see [18, 20]. This problem is also known in computational geometry under the name of the mailbox problem, see, e.g., [26, Sect. 9.2.1]. Another similar application is the optimal stratification in Monte Carlo integration, see, e.g., [30, Sect. 5.5].

4.3 Optimal Quantisation

The optimal server placement problem from the previous section can be thought of as a representation of a measure ν on $\mathbb{R}^d$ (that describes the probability distribution of customers) by another (discrete) measure with k atoms. This is a well-known optimal quantisation problem, see [7, 8]. Apart from finding the optimal quantiser, it is important to know the asymptotic behaviour of the quantisation error, which is the infimum of the objective function. The classical quantisation theory concerns the case when the quantiser is deterministic. We follow a variant of this problem for quantising points that form a Poisson point process of total intensity a studied in [17].

Let $p(y)$, $y \in \mathbb{R}^d$ be a Riemann integrable function with bounded support K that is proportional to the density of the probability measure to be approximated by a discrete one. The objective functional for the optimal Poisson quantisation problem is then

$$E(p;\mu) = \int\limits_{\mathbb{R}^d} \mathbb{E}_\mu \rho(y,\eta)^\beta p(y)\mathrm{d}y\,.$$

Denote

$$E_a(p) = n^{\beta/d} \inf_{\mu \in \mathbf{M}_\mathrm{f},\ \mu(\mathbb{R}^d)=a} E(p;\mu)\,.$$

Theorem 9 *The limit of $E_n(p)$ as $n \to \infty$ exists and*

$$\lim_{n\to\infty} E_n(p) = \mathfrak{I}\|p\|_{d/(d+\beta)} = \mathfrak{I}\left(\int\limits_{\mathbb{R}^d} p(y)^{d/(d+\beta)}\mathrm{d}y\right)^{1+\beta/d}$$

for a certain constant $\mathfrak{I}$ *that depends only on* β *and dimension* d*. If* $a\mu_a$ *is supported by* K *and minimises* $E(p;\mu)$ *over all measures with the total mass* a*, then* μ_a *weakly converges as* $a \to \infty$ *to the probability measure with density proportional to* $p(y)^{d/(d+\beta)}$.

The proof from [17] does not rely on Theorem 8. Theorem 9 is proved first for the uniform distribution $p(y) \equiv const$ and then extended to a non-uniform case. The main idea is the firewall construction from [7] that ensures the additivity of the objective functional for indicators of disjoint sets. The main new feature in the Poisson case is that the firewalls constructed by adding extra cluster points in the stochastic case correspond to the changes in the intensity and so may be empty. Bounds on the coverage probabilities from [9] are used in order to ensure that the firewalls are established with a high probability. The constant $\mathfrak{I}$ is the limit of the quantisation error for the uniform distribution on the unit cube.

Note that laws of large numbers for functionals of point processes have been considered in [27]. They make it possible to obtain the limit of a functional of a Poisson process with intensity measure $a\mu$ for any given μ as $a \to \infty$. However, [27] does not contain any results about convergence of minimal values and minimisers. By examining the proof of [27, Lemma 3.1] it is possible to justify the uniform convergence of the rescaled functional of $a\mu_a$ for a measure μ_a with density p_a (and so arrive at the convergence results for minimal values) if

$$a \int_{\|y-x\|\le a^{-1/d}} |p_a(y) - p_a(x)|\mathrm{d}y \to 0 \quad \text{as } a \to \infty \tag{12}$$

for all $x \in \mathbb{R}^d$. If $p_a(x) \to p(x)$ as $a \to \infty$, (12) implies the validity of the double limit condition (10).

4.4 Optimal Search

Let Y be a random closed subset of $\mathbb{R}^d$ that is independent of the Poisson process η. The aim is to determine the intensity measure μ that maximises the coverage probability $\mathbb{P}\{\eta(Y) > 0\}$ meaning that at least one point of η hits Y. Equivalently, it is possible to minimise the avoidance probability

$$\mathbb{E}_\mu \mathbb{1}_{\eta(Y)=0} = \mathbb{E}_\mu e^{-\mu(Y)}\,.$$

The expected first difference is given by

$$g(x;\mu) = \mathbb{E}_\mu D_x F = -\mathbb{E}_\mu \left[e^{-\mu(Y)} \mathbb{1}_{x\in Y}\right]\,.$$

If Y is a subset of a countable space, it is possible to determine μ explicitly, see [18, Sect. 5.5]. Otherwise, the high intensity approach applies. For instance, if $Y = B_\zeta(\xi)$ is a random ball of radius ζ centred at an independent ξ with probability densities p_ζ and p_ξ, then

$$\begin{aligned}&\mathbb{E}_{p(x)\ell_d}\Gamma(x;\eta)\\&\quad\propto -\kappa_d p_\xi(x)\left[\frac{p_\eta(0)(d+1)\Gamma(1+1/d)}{(ap(x)\kappa_d)^{1+1/d}}+\frac{p'_\eta(0)(d+2)\Gamma(1+2/d)}{(ap(x)\kappa_d)^{1+2/d}}+\cdots\right],\end{aligned}$$

where κ_d is the volume of a unit ball in $\mathbb{R}^d$. Thus, the density of the asymptotically optimal measure is proportional to $(p_\xi)^{d/(d+1)}$ if $p_\eta(0) \neq 0$, and to $(p_\xi)^{d/(d+2)}$ if $p_\eta(0) = 0$ and $p'_\eta(0) \neq 0$, etc.

5 Steepest Descent Algorithms

Algorithms of the steepest descent type are widely used in the optimisation literature see, e.g., [28]. The basic steepest descent algorithm consists in moving from a measure μ_n (approximate solution at step n) to $\mu_{n+1} = \mu_n + \nu_n$, where ν_n minimises the directional derivative, which in our context becomes $L(\mu)[\nu] = \langle g(\cdot\,;\mu),\nu\rangle$ with $g(x;\mu) = \mathbb{E}_\mu D_x F(\eta)$.

The general description of the steepest descent direction from [22, Theorem 4.1] in the case of optimisation over intensity measures with a fixed total mass yields the following result.

Theorem 10 *The minimum of $L(\mu)[\nu]$ over all $\nu \in \tilde{\mathbf{M}}_\mathrm{f}$ with $\|\nu\| \le \varepsilon$ is achieved on a signed measure ν such that ν^+ is the positive measure with total mass $\varepsilon/2$ concentrated on the points of the global minima of $g(x;\mu)$ and $\nu^- = \mu|_{M(t_\varepsilon)} + \delta\mu|_{M(s_\varepsilon)\setminus M(t_\varepsilon)}$, where*

$$M(p) = \{x \in \mathbb{X} :\ g(x;\mu) \ge p\},$$

and

$$\begin{aligned}t_\varepsilon &= \inf\{p :\ \mu(M(p)) < \varepsilon/2\}\,,\\ s_\varepsilon &= \sup\{p :\ \mu(M(p)) \ge \varepsilon/2\}\,.\end{aligned}$$

The factor δ is chosen in such a way that $\mu(M(t_\varepsilon)) + \delta\mu(s_\varepsilon)) = \varepsilon/2$.

This result means that the mass of μ is eliminated at high gradient locations, while μ acquires extra atoms at locations where the gradient is the smallest.

In a numeric implementation, the space $\mathbb{X}$ is discretised and the discrete variant of μ is considered. The corresponding steepest descent algorithms are used in

R-libraries *mefista* (for optimisation with a fixed mass) and *medea* (for optimisation with many linear equality constraints) available from the authors' web pages. The increment step size in these algorithms is chosen by either the Armijo method described in [28, Sect. 1.3.2] or by taking into account the difference between the supremum and the infimum of $g(x;\mu_n)$ over the support of μ_n.

Numeric computations of an optimal measure relies on effective evaluation of the gradient function which is possible to obtain in many cases as the next sections demonstrate.

5.1 Design of Experiments

The basic problem in the theory of linear optimal design of experiments [2] aims to find positions of design (observation) points x_i in order to minimise the determinant of the covariance matrix of estimators of coefficients β_j in the linear regression model

$$y_i = \sum_{j=1}^{k} \beta_j r_j(x_i) + \varepsilon_i \,,$$

where $r = (r_1, \dots, r_k)^\top$ is a column vector of linearly independent functions and ε_i are i.i.d. centred errors. If the design points are produced from a probability distribution $\mu(\mathrm{d}x)$ reflecting the frequency of taking x as an observation point, the objective function can be expressed as

$$f(\mu) = -\log\det M(\mu) \,,$$

where the covariance matrix M is given by

$$M(\mu) = \int r(x)^\top r(x)\mu(\mathrm{d}x) \,.$$

For the optimisation purpose, it is possible to discard the logarithm, so that the gradient function in this model becomes

$$g(x;\mu) = -r(x)M^{-1}(x)r^\top(x) \,,$$

see [21, 23]. It is also possible to consider the Poissonised variant of the optimal design problem. It should be noted however that adding an extra design point has a non-local effect and so the high-intensity approach from Sect. 4 does not apply in these problems.

5.2 Mixtures

Let $\{p_x(\cdot)\}$ be a family of probability densities indexed by $x \in \mathbb{X}$. For a probability measure μ on $\mathbb{X}$ define the mixture

$$p_\mu(y) = \int_{\mathbb{X}} p_x(y)\mu(\mathrm{d}x)\,.$$

The estimation of the mixing distribution μ is a well-studied topic in statistics. The steepest descent algorithm in the space of measures yields a pure non-parametric approach to the estimation of μ based on maximising the log-likelihood

$$f(\mu) = \sum_{i=1}^{n} \log p_\mu(y_i)$$

based on a sample $y_1, \ldots, y_n$. The gradient function is

$$g(x;\mu) = \sum_{i=1}^{n} \frac{p_x(y_i)}{\int p_x(y)\mu(\mathrm{d}x)}\,.$$

5.3 P-Means

Recall that measure μ that minimises the functional (11) is called the solution of the P-means problem. A direct computation shows that the gradient of the functional (11) is given by

$$g(x;\mu) = -\sum_{y_j} \int_{\|x-y_j\|^\beta}^{u^2} \exp\{-\mu(B_{\sqrt{t}}(y_j))\}\mathrm{d}t\,.$$

5.4 Maximisation of the Covered Volume

Let η be a Poisson process in $\mathbb{X} \subset \mathbb{R}^d$ with intensity measure μ. If $B_r(x)$ is a ball of radius r centred at x, then

$$\Xi = \bigcup_{x_i \in \eta} B_r(x_i)$$

is called a *Boolean model*, see [15, 33]. The ball of radius r is referred to as the *typical grain*, which can be also a rather general random compact set. Then

$$\mathbb{P}\{x \notin \Xi\} = \exp\{-\mu(B_r(x))\}\,.$$

Fubini's theorem yields that the expected uncovered volume is given by

$$f(\mu) = \int_{\mathbb{X}} \mathbb{P}\{x \notin \Xi\}\mathrm{d}x = \int_{\mathbb{X}} \exp\{-\mu(B_r(x))\}\mathrm{d}x\,.$$

A minimiser of $f(\mu)$ yields the intensity of a Poisson process with the largest coverage. The gradient is directly computed as

$$g(x;\mu) = -\int_{B_r(x)} \exp\{-\mu(B_r(z))\}\mathrm{d}z\,.$$

Further related problems are discussed in [24] in relation to design of materials with given properties. This problem does not admit the high-intensity solution, since adding an extra ball affects the configuration within distance r which does not go to zero as the intensity of the Poisson process grows.

Acknowledgements The authors thank the organisers of the Oberwolfach Mini-Workshop: Stochastic Analysis for Poisson point processes (Feb. 2013) and the hospitality of the Oberwolfach Mathematical Institute for an excellent forum to exchange ideas partly presented in this paper.

References

1. Albeverio, S., Kondratiev, Y.G., Röckner, M.: Analysis and geometry on configuration spaces. J. Funct. Anal. **154**(2), 444–500 (1998)
2. Atkinson, A.C., Donev, A.N.: Optimum Experimental Designs. Clarendon Press, Oxford (1992)
3. Ben-Tal, A., Zowe, J.: A unified theory of first and second order conditions for extremum problems in topological vector spaces. Optimality and stability in mathematical programming. Math. Program. Stud. **19**, 39–76 (1982)
4. Cominetti, R.: Metric regularity, tangent sets and second-order optimality conditions. Appl. Math. Optim. **21**, 265–287 (1990)
5. Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. General Theory and Structure, vol. II, 2nd edn. Springer, New York (2008)
6. Dunford, N., Schwartz, J.T.: Linear Operators. General Theory, vol. 1. Wiley, New York (1988)
7. Graf, S., Luschgy, H.: Foundations of Quantization for Probability Distributions. Lecture Notes in Mathematics, vol. 1730. Springer, Berlin (2000)
8. Gruber, P.M.: Optimum quantization and its applications. Adv. Math. **186**, 456–497 (2004)
9. Hall, P.: On the coverage of k-dimensional space by k-dimensional spheres. Ann. Probab. **13**, 991–1002 (1985)
10. Ho, Y., Cao, X.: Perturbation Analysis of Discrete Event Dynamic Systems. Kluwer Academic Publishers, Boston (1991)

11. Last, G.: Perturbation analysis of Poisson processes. Bernoulli **20**(2), 486–513 (2014)
12. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
13. Margulis, G.A.: Probabilistic characteristics of graphs with large connectivity. Probl. Pereda. Inf. **10**(2), 101–108 (1974) (in Russian)
14. McClure, D.E., Vitale, R.A.: Polygonal approximation of plane convex bodies. J. Math. Anal. Appl. **51**, 326–358 (1975)
15. Molchanov, I.S.: Statistics of the Boolean Model for Practitioners and Mathematicians. Wiley, Chichester (1997)
16. Molchanov, I., Tontchev, N.: Optimal approximation and quantisation. J. Math. Anal. Appl. **325**, 1410–1429 (2007)
17. Molchanov, I., Tontchev, N.: Optimal Poisson quantisation. Stat. Probab. Lett. **77**, 1123–1132 (2007)
18. Molchanov, I.S., Zuyev, S.A.: Variational analysis of functionals of a Poisson process. Tech. Rep. RR-3302, INRIA, Sophia-Antipolis ftp://ftp.inria.fr/INRIA/publication/RR/RR-3302.ps.gz (1997)
19. Molchanov, I., Zuyev, S.: Tangent sets in the space of measures: with applications to variational analysis. J. Math. Anal. Appl. **249**, 539–552 (2000)
20. Molchanov, I., Zuyev, S.: Variational analysis of functionals of Poisson processes. Math. Oper. Res. **25**(3), 485–508 (2000)
21. Molchanov, I., Zuyev, S.: Variational calculus in space of measures and optimal design. In: Atkinson, A., Bogacka, B., Zhigljavsky, A. (eds.) Optimum Experimental Design: Prospects for the New Millennium, pp. 79–90. Kluwer, Dordrecht (2000)
22. Molchanov, I., Zuyev, S.: Steepest descent algorithms in space of measures. Stat. Comput. **12**, 115–123 (2002)
23. Molchanov, I., Zuyev, S.: Optimisation in space of measures and optimal design. ESAIM: Probab. Stat. **8**, 12–24 (2004)
24. Molchanov, I., Chiu, S., Zuyev, S.: Design of inhomogeneous materials with given structural properties. Phys. Rev. E **62**, 4544–4552 (2000)
25. Molchanov, I., van Lieshaut, M., Zuyev, S.: Clustering methods based on variational analysis in the space of measures. Biometrika **88**(4), 1021–1033 (2001)
26. Okabe, A., Boots, B., Sugihara, K., Chiu, S.N.: Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd edn. Wiley, Chichester (2000)
27. Penrose, M.D., Yukich, J.E.: Weak laws in geometric probability. Ann. Appl. Probab. **13**, 277–303 (2003)
28. Polak, E.: Optimization: Algorithms and Consistent Approximations. Springer, New York (1997)
29. Röckner, M.: Stochastic analysis on configuration spaces: basic ideas and recent results. In: New Directions in Dirichlet Forms. AMS/IP Studies in Advanced Mathematics, vol. 8, pp. 157–231. American Mathematical Society, Providence (1998)
30. Rubinstein, R., Croese, D.: Simulation and the Monte-Carlo Method. Wiley, New York (2008)
31. Russo, L.: On the critical percolation probabilities. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **56**, 229–237 (1981)
32. Schneider, R.: Random approximations of convex sets. J. Microsc. **151**, 211–227 (1988)
33. Stoyan, D., Kendall, W.S., Mecke, J.: Stochastic Geometry and its Applications, 2nd edn. Wiley, Chichester (1995)
34. Zuyev, S.A.: Russo's formula for the Poisson point processes and its applications. Diskret. Mate. **4**(3), 149–160 (1992) (In Russian). English translation: Discrete Math. Appl. **3**, 63–73 (1993)
35. Zuyev, S.: Stopping sets: gamma-type results and hitting properties. Adv. Appl. Probab. **31**, 355–366 (1999)

Malliavin Calculus for Stochastic Processes and Random Measures with Independent Increments

Josep Lluís Solé and Frederic Utzet

Abstract Malliavin calculus for Poisson processes based on the *difference operator* or *add-one-cost operator* is extended to stochastic processes and random measures with independent increments. Our approach is to use a Wiener–Itô chaos expansion, valid for both stochastic processes and random measures with independent increments, to construct a Malliavin derivative and a Skorohod integral. Useful derivation rules for smooth functionals given by Geiss and Laukkarinen (Probab Math Stat 31:1–15, 2011) are proved. In addition, characterizations for processes or random measures with independent increments based on the duality between the Malliavin derivative and the Skorohod integral following an interesting point of view from Murr (Stoch Process Appl 123:1729–1749, 2013) are studied.

1 Introduction

This chapter is divided into two parts: the first is devoted to processes with independent increments and the second to random measures with independent increments. Of course, both parts are strongly related to each other and we had doubts about the best order in which to present them in order to avoid repetition. We decided to start with stochastic processes where previous results are better known, and this part is mainly based on Solé et al. [24] where a Malliavin Calculus for Lévy processes is developed. Our approach relies on a chaotic expansion of square integrable functionals of the process, stated by Itô [5], in terms of a vector random measure on the plane; that expansion gives rise to a Fock space structure and enables us to define a Malliavin derivative and a Shorohod integral as an annihilation and creation operator respectively. Later, using an *ad hoc* canonical space, the Malliavin derivative restricted to the jumps part of the process can be conveniently interpreted as an increment quotient operator, extending the idea of the *difference operator* or *add-one-cost operator* of the Poisson processes, see Nualart and Vives [18, 19], Last and Penrose [12], and Last [11] in this volume. We also extend the interesting

J.L. Solé • F. Utzet (✉)
Department of Mathematics, Universitat Autònoma de Barcelona, Campus de Bellaterra, 08193 Barcelona, Catalunya, Spain
e-mail: jllsole@mat.uab.cat; utzet@mat.uab.cat

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_4

formulas of Geiss and Laukkarinen [4] for computing the derivatives of smooth functionals, which widen considerably the practical applications of the calculus. Finally, following Murr [15], we prove that the duality coupling between Malliavin derivative and Skorohod integral characterizes the underlying process and, in this way, extends to stochastic processes some characterizations of Stein's method type. We should point out that in the first part (and also in the second, as we comment below) we use the very general and deep results of Last and Penrose [12] and Last [11] to improve some results and simplify the proofs of Solé et al. [24].

It is worth remarking that there is another approach to a chaos-based Malliavin calculus for jump processes using a different chaos expansion, that we comment in Sect. 2.2. For that development and many applications see Di Nunno et al. [3] and the references therein.

In the second part we extend Malliavin calculus to a random measure with independent increments. We start by recalling a representation theorem of such a random measure in terms of an integral with respect to a Poisson random measure in a product space; a weak version (in law) of that representation was obtained by Kingman [8] (see also Kingman [9]). That representation gives rise to the possibility of building a Malliavin calculus due to the fact that Itô's [5] chaotic representation property also holds here. In this context, the results of Last and Penrose [12] and Last [11] play a central role since, thanks to them, it is not necessary to construct a canonical space, and we can simply interpret the Malliavin derivative as an *add-one-cost* operator. As in the first part, we introduce the smooth functionals of Geiss and Laukkarinen [4], and the characterization of random measures with independent increments by duality formulas of Murr [15].

2 Part 1: Malliavin Calculus for Processes with Independent Increments

2.1 Processes with Independent Increments and Its Lévy–Itô Decomposition

This section contains the notations and properties of processes with independent increments that we use; we mainly follow the excellent book of Sato [22]. In particular, we present the so-called Lévy–Itô decomposition of a process with independent increments as a sum of a continuous function, a continuous Gaussian process with independent increments, and two integrals. One of these integrals is considered with respect to a Poisson random measure whereas the other with respect to a compensated Poisson random measure. These integrals are, respectively, the sum of the big jumps of the process and the compensated sum of small jumps. That decomposition is a masterpiece of stochastic processes theory, and there exist proofs of such a fact that are based on very different tools: see, for example, Sato [22] and Kallenberg [7].

Fix a probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Let $X = \{X_t,\ t \geq 0\}$ be a real process with independent increments, that is, for every $n \geq 1$ and $0 \leq t_1 < \cdots < t_n$, the random variables $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent. We assume that $X_0 = 0$, a.s., and that X is continuous in probability and cadlag. A process with all these properties is also called *an additive process*. We assume that the σ-field $\mathscr{A}$ is generated by X.

The hypothesis that the process is cadlag is not restrictive: every process with independent increments and continuous in probability has a cadlag modification (Sato [22, Theorem 11.5]). The conditions of continuity in probability and cadlag prevent the existence of fixed discontinuities, that is to say, there are no points $t \geq 0$ such that $\mathbb{P}\{X_t \neq X_{t-}\} > 0$.

The system of generating triplets of X is denoted by $\{(m_t, \rho_t, \nu_t),\ t \geq 0\}$. Thus, $m : \mathbb{R}_+ \to \mathbb{R}$, where $\mathbb{R}_+ = [0, \infty)$, is a continuous function that gives a deterministic tendency of the process (see representation (2)); $\rho_t \geq 0$ is the variance of the Gaussian part of X_t, and ν_t is the Lévy measure of the jumps part. More specifically, ν_t is a measure on $\mathbb{R}_0$, where $\mathbb{R}_0 = \mathbb{R}\backslash\{0\}$, such that $\int_{\mathbb{R}_0} (1 \wedge x^2)\, \nu_t(\mathrm{d}x) < \infty$, where $a \wedge b = \min(a, b)$. Observe that for all $t \geq 0$ and $\varepsilon > 0$, $\nu_t\big((-\varepsilon, \varepsilon)^c\big) < \infty$, and hence ν_t is finite on compact sets of $\mathbb{R}_0$, and then σ-finite. Denote by ν the (unique) measure on $\mathscr{B}((0, \infty) \times \mathbb{R}_0)$ defined by

$$\nu\big((0, t] \times B\big) = \nu_t(B),\ B \in \mathscr{B}(\mathbb{R}_0). \tag{1}$$

It is also σ-finite, and moreover, $\nu\big(\{t\} \times \mathbb{R}_0\big) = 0$ for every $t > 0$ (Sato [22, p. 53]); thus it is non-atomic. The measure ν controls the jumps of the process: for $B \in \mathscr{B}(\mathbb{R}_0)$, $\nu((0, t] \times B)$ is the expectation of the number of jumps of the process in the interval $(0, t]$ with size in B. We remark that in a finite time interval the process can have an infinity of jumps of small size, and there are Lévy measures such that, for example, $\nu\big((0, t] \times (0, x_0)\big) = \infty$, for some $x_0 > 0$.

Write

$$N(C) = \#\{t : (t, \Delta X_t) \in C\}, \quad C \in \mathscr{B}((0, \infty) \times \mathbb{R}_0),$$

the jumps measure of the process, where $\Delta X_t = X_t - X_{t-}$. It is a Poisson random measure on $(0, \infty) \times \mathbb{R}_0$ with intensity measure ν (Sato [6, Theorem 19.2]). Let

$$\hat{N} = N - \nu$$

represent the compensated jumps measure.

Theorem 1 (Lévy–Itô Decomposition)

$$X_t = m_t + G_t + \int_{(0,t]\times\{|x|>1\}} x N(\mathrm{d}(s, x)) + \int_{(0,t]\times\{0<|x|\leq 1\}} x \widehat{N}(\mathrm{d}(s, x)), \tag{2}$$

where $\{G_t,\ t \geq 0\}$ is a centered continuous Gaussian process with independent increments and variance $\mathbb{E}[G_t^2] = \rho_t^2$, independent of N.

Sato [22, Theorem 19.2] gives a more precise statement, and instead of the second integral in (2) he writes

$$\lim_{\varepsilon \downarrow 0} \int_{(0,t]\times\{\varepsilon<|x|\leq 1\}} x\,\widehat{N}(\mathrm{d}(s,x)),$$

where the convergence is a.s., uniform in t on every bounded interval.

Remark 1

1. The function $t \mapsto \rho_t$ is continuous and increasing, and $\rho_0 = 0$ (Sato [22, Theorem 9.8]), and hence it defines a σ-finite and non-atomic measure on $\mathbb{R}_+$, denoted by ρ. The Gaussian process $\{G_t,\ t \geq 0\}$ introduced above defines through

$$G\big((s,t]\big) = G_t - G_s,\ 0 \leq s < t,$$

a centered Gaussian random measure G on $\{B \in \mathscr{B}(\mathbb{R}_+),\ \rho(B) < \infty\}$ with control measure ρ (see Peccati and Taqqu [20, p. 63] for this definition). In the Gaussian Malliavin calculus terminology this is called a *white noise measure* (Nualart [17, p. 8]). This will be important when we define Malliavin derivatives with respect to X.
2. Remember that a Lévy process is an additive process with stationary increments. In this case, $m_t = m^\circ t$, for some $m^\circ \in \mathbb{R}$, $\rho_t = \rho^\circ t$, for some $\rho^\circ \geq 0$, and the Gaussian process $\{G_t,\ t \geq 0\}$ can be written as $G_t = \sqrt{\rho^\circ}\, W_t$, where $\{W_t,\ t \geq 0\}$ is a standard Brownian motion. Also, $\nu_t = t\nu^\circ$ for some Lévy measure ν°, and the measure ν is simply the product measure of the Lebesgue measure on $(0,\infty)$ and ν°: $\nu(\mathrm{d}(t,x)) = \mathrm{d}t\,\nu^\circ(\mathrm{d}x)$.
3. The notations are slightly different from Sato [22] and Solé et al. [24], where ν denotes the Lévy measure of a Lévy process, that in the previous point we write ν°. Also, our measure ν on $(0,\infty) \times \mathbb{R}_0$ defined in (1) is denoted by Sato [22] by $\tilde{\nu}$.

2.2 Wiener–Itô Chaos Expansion

The well-known Wiener–Itô chaos expansion of square integrable functionals of a Brownian motion can be extended to the square integrable functionals of a process with independent increments. This was another major contribution made by Itô [5]; indeed, Itô proved that result for Lévy processes, however his proof is written in very general terms and also covers the case of processes with independent increments.

That chaos expansion determines a Fock space structure on $L^2(\mathbb{P})$, which is the basis of our Malliavin calculus development.

With the preceding notations, define a measure μ on $\big(\mathbb{R}_+ \times \mathbb{R}, \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})\big)$ by

$$\mu(\mathrm{d}(t,x)) = \rho(\mathrm{d}t)\,\delta_0(\mathrm{d}x) + x^2 \mathbb{1}_{(0,\infty)\times\mathbb{R}_0}\nu(\mathrm{d}(t,x)). \tag{3}$$

It is non-atomic since ν and ρ are non-atomic. Moreover, for a bounded set $B \in \mathscr{B}(\mathbb{R})$,

$$\mu([0,t] \times B) = \rho_t \delta_0(B) + \int_{B\cap\mathbb{R}_0} x^2\, \nu_t(\mathrm{d}x),$$

and the last integral on the right-hand side is equal to

$$\int_{B\cap\{0<|x|\le 1\}} x^2\, \nu_t(\mathrm{d}x) + \int_{B\cap\{\,|x|>1\}} x^2\, \nu_t(\mathrm{d}x)$$
$$\le \int_{\{0<|x|\le 1\}} x^2\, \nu_t(\mathrm{d}x) + C\, \nu_t\big(\{x:\ |x|>1\}\big) < \infty,$$

where C is a constant. Hence, the measure μ is locally finite, and, in particular, σ-finite.

Extending Itô [5] to this context, we can define a random measure (in the sense of vector measures, see Appendix 2) M on $\big(\mathbb{R}_+ \times \mathbb{R}, \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})\big)$ with control measure μ: for $C \in \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})$, such that $\mu(C) < \infty$, write $C(0) = \{t \ge 0:\ (t,0) \in C\}$ and $C^* = C \cap \big((0,\infty) \times \mathbb{R}_0\big)$, and note that

$$\int_{(0,\infty)\times\mathbb{R}_0} \mathbb{1}_{C^*}(s,x) x^2\, \nu(\mathrm{d}(s,x)) < \infty,$$

that is, $\mathbb{1}_{C^*}(t,x)x \in L^2(\nu)$. So there exists the $L^2(\mathbb{P})$ integral of that function with respect to $\widehat{N}$ (see Appendix 1), and we can define

$$M(C) = G\big(C(0)\big) + \int_{C^*} x\, \widehat{N}(\mathrm{d}(t,x)).$$

We prove that M is a completely random measure; see Appendix 2 where these definitions are recalled.

Proposition 1 *M is a completely random measure with control measure μ.*

Proof In this proof, all sets $C \in \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})$ are assumed to have finite μ-measure. It is clear that $\mathbb{E}[M(C)] = 0$, and by the independence between G and N it follows that

$$\mathbb{E}\big[M(C_1)M(C_2)\big] = \mu(C_1 \cap C_2).$$

From (45) in the appendix it is deduced that the characteristic function of $M(C)$ is

$$\mathbb{E}\Big[\exp\big(iuM(C)\big)\Big] = \exp\left[-\frac{u^2}{2}\rho\big(C(0)\big) + \int_{\mathbb{R}_0}\Big(e^{iux} - 1 - iux\Big)\,\alpha_C(\mathrm{d}x)\right],$$

where α_C is the measure on $\mathbb{R}_0$ defined for $A \in \mathscr{B}(\mathbb{R}_0)$ by

$$\alpha_C(A) = \nu\Big(C \cap \big((0,\infty) \times A\big)\Big).$$

By a standard approximation argument it is proved that if $f : \mathbb{R}_0 \to \mathbb{R}_+$ is measurable, then

$$\int_{\mathbb{R}_0} f(x)\,\alpha_C(\mathrm{d}x) = \int_{C\cap((0,\infty)\times\mathbb{R}_0))} f(x)\,\nu(\mathrm{d}(t,x)).$$

Thus

$$\int_{\mathbb{R}_0} x^2\,\alpha_C(\mathrm{d}x) = \int_{C\cap((0,\infty)\times\mathbb{R}_0))} x^2\nu(\mathrm{d}(t,x)) \le \mu(C) < \infty.$$

Therefore, α_C is a Lévy measure with finite second order moment. Then $M(C)$ has an infinitely divisible law with finite variance and Lévy measure given by α_C. Furthermore, if $C_1, C_2 \in \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})$ are disjoint, $\alpha_{C_1\cup C_2} = \alpha_{C_1} + \alpha_{C_2}$, and it follows that if $C_1, \ldots, C_n \in \mathscr{B}(\mathbb{R}_+ \times \mathbb{R})$, all with finite μ-measure, are disjoint, then $M(C_1), \ldots, M(C_n)$ are independent. □

Hence, we can define multiple Wiener–Itô integrals with respect to M, see Appendix 2. Let $L^2_s(\mu^n)$ be the subset of symmetric functions of $L^2(\mu^{\otimes n})$, and for $f \in L^2_s(\mu^n)$ denote by $I_n(f)$ the multiple integral of f with respect to M.

The chaotic representation Theorem of square integrable functionals of a Lévy process of Itô [5, Theorem 2] is extended to this case with the same proof. So we have the chaotic decomposition property:

$$L^2(\mathbb{P}) = \bigoplus_{n=0}^{\infty} I_n\big(L^2_s(\mu^n)\big),$$

and the (unique) representation of a functional $F \in L^2(\Omega)$,

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L_s^2(\mu^n).$$

From this point, we can apply all the machinery of the annihilation operators (Malliavin derivatives) and creation operators (Skorohod integrals) on Fock spaces, as exposed in Nualart and Vives [18, 19].

Remark 2 For a process without Gaussian part, we can consider the chaos expansion of a square integrable functional in terms of the multiple integrals with respect to the Poisson random measure N rather than M, and then define a Malliavin derivative and a Skorohod integral; see Di Nunno et al. [3] and the references therein. Indeed, in the second part of this paper, dealing with random measures with independent increments, we combine that approach with the multiple integral with respect to M.

2.3 Derivative Operators

Let $F \in L^2(\mathbb{P})$ with a finite chaos expansion

$$F = \sum_{n=0}^{N} I_n(f_n),$$

where $N < \infty$. The Malliavin derivative of F is defined as the element of $L^2(\mu \otimes \mathbb{P})$ given by

$$D_z F = \sum_{n=1}^{N} n I_{n-1}\big(f_n(z, \cdot)\big), \; z \in \mathbb{R}_+ \times \mathbb{R}.$$

This operator is unbounded. However, the set of elements of $L^2(\mathbb{P})$ with finite chaos expansion is dense in $L^2(\mathbb{P})$, and the operator D is closable; the domain of D, denoted by $\operatorname{dom} D$, coincides with the set of $F \in L^2(\mathbb{P})$ with chaotic decomposition

$$F = \sum_{n=0}^{\infty} I_n(f_n),$$

such that

$$\sum_{n=1}^{\infty} n\, n!\|f_n\|^2_{L^2_s(\mu^n)} < \infty. \tag{4}$$

The Malliavin derivative of such an F is given by

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}\Big(f_n\big(z,\cdot\big)\Big), \quad z \in \mathbb{R}_+ \times \mathbb{R},$$

where the convergence of the series is in $L^2(\mu \otimes \mathbb{P})$.

The domain $\operatorname{dom} D$ is a Hilbert space with the scalar product

$$\langle F, G\rangle = \mathbb{E}[F\, G] + \mathbb{E}\left[\int_{\mathbb{R}_+\times\mathbb{R}} D_z F\, D_z G\, \mu(\mathrm{d}z)\right]. \tag{5}$$

For all these properties we refer to Nualart and Vives [18].

Given the form of the measure μ, for $f : (\mathbb{R}_+ \times \mathbb{R})^n \to \mathbb{R}$ measurable, positive or $\mu^{\otimes n}$ integrable, we have

$$\begin{aligned}
&\int_{(\mathbb{R}_+\times\mathbb{R})^n} f\, d\mu^{\otimes n} \\
&\quad = \int_{\mathbb{R}_+\times(\mathbb{R}_+\times\mathbb{R})^{n-1}} f\big((t,0), z_1, \dots z_{n-1})\big)\, \rho(\mathrm{d}t)\, \mu^{\otimes(n-1)}(\mathrm{d}z_1, \dots, \mathrm{d}z_{n-1}) \\
&\qquad + \int_{(0,\infty)\times\mathbb{R}_0\times(\mathbb{R}_+\times\mathbb{R})^{n-1}} f(z_1, z_2, \dots z_n)\, \mu^{\otimes n}(\mathrm{d}z_1, \dots, \mathrm{d}z_n).
\end{aligned}$$

As a consequence, when $\rho \neq 0$ and $\nu \neq 0$, it is natural to consider two more spaces: Let $\operatorname{dom} D^0$ (if $\rho \neq 0$) be the set of $F \in L^2(\mathbb{P})$ with decomposition $F = \sum_{n=0}^{\infty} I_n(f_n)$ such that

$$\sum_{n=1}^{\infty} n\, n! \int_{\mathbb{R}_+\times(\mathbb{R}_+\times\mathbb{R})^{n-1}} f^2((t,0), z_1, \dots, z_{n-1})\, \rho(\mathrm{d}t)\, \mu^{\otimes(n-1)}(\mathrm{d}z_1, \dots, \mathrm{d}z_{n-1}) < \infty.$$

For $F \in \operatorname{dom} D^0$ we can define the square integrable stochastic process

$$D_{t,0}F = \sum_{n=1}^{\infty} n\, I_{n-1}\Big(f_n\big((t,0),\cdot\big)\Big),$$

where the convergence is in $L^2(\rho \otimes \mathbb{P})$. Analogously, if $\nu \neq 0$, let $\operatorname{dom} D^J$ be the set of $F \in L^2(\mathbb{P})$ such that

$$\sum_{n=1}^{\infty} n\, n! \int_{(0,\infty)\times\mathbb{R}_0\times(\mathbb{R}_+\times\mathbb{R})^{n-1}} f_n^2 \, d\mu^{\otimes n} < \infty,$$

and for $F \in \operatorname{dom} D^J$, define

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}\Big(f_n\big(z,\cdot\big)\Big),$$

where the convergence is in $L^2\big((0,\infty) \times \mathbb{R}_0 \times \Omega, x^2\, \nu(\mathrm{d}(t,x)) \otimes \mathbb{P}\big)$.

It is clear that when both $\rho \neq 0$ and $\nu \neq 0$, then $\operatorname{dom} D = \operatorname{dom} D^0 \cap \operatorname{dom} D^J$.

2.4 The Skorohod Integral

Following the scheme of Nualart and Vives [18], we can define a creation operator (Skorohod—or Kabanov–Skorohod—integral) in the following way: let $g \in L^2\big(\mu \otimes \mathbb{P}\big)$, which has a chaotic decomposition

$$g(z) = \sum_{n=0}^{\infty} I_n(f_n(z,\cdot)), \tag{6}$$

where $f_n \in L^2\big(\mu^{\otimes(n+1)}\big)$ is symmetric in the n last variables. Denote by $\hat{f}_n$ the symmetrization in all $n+1$ variables. If

$$\sum_{n=0}^{\infty} (n+1)!\, \|\hat{f}_n\|^2_{L^2_s(\mu^{n+1})} < \infty, \tag{7}$$

define the Skorohod integral of f by

$$\delta(g) = \sum_{n=0}^{\infty} I_{n+1}(\hat{f}_n),$$

where the convergence is in $L^2(\mathbb{P})$. Denote by $\operatorname{dom}\delta$ the set of g that satisfy (7). The operator δ is the dual of the operator D, that is, a process $g \in L^2(\mu \times \mathbb{P})$ belongs to dom δ if and only if there is a constant C such that for all $F \in \operatorname{dom} D$,

$$\Big| \mathbb{E} \int_{\mathbb{R}_+ \times \mathbb{R}} g(z)\, D_z F\, \mu(\mathrm{d}z) \Big| \le C \left(\mathbb{E}[F^2] \right)^{1/2}.$$

If $g \in \operatorname{dom} \delta$, then $\delta(g)$ is the element of $L^2(\mathbb{P})$ characterized by the duality (or integration by parts) formula

$$\mathbb{E}[\delta(g)\, F] = \mathbb{E} \int_{\mathbb{R}_+ \times \mathbb{R}} g(z)\, D_z F\, \mu(\mathrm{d}z), \tag{8}$$

for any $F \in \operatorname{dom} D$.

For more properties of the operator δ in the Lévy processes case, including its relationship with the stochastic integral with respect to the measure M, and a Clark–Ocone–Haussman formula, we refer to Solé et al. [24, 25].

2.5 Derivation of Smooth Functionals

Following an interesting approach of Geiss and Laukkarinen [4] (in the Lévy processes context) we will prove the following formulas of the derivative of smooth functionals: denote by $\mathscr{C}_b^\infty(\mathbb{R}^n)$ the set of infinitely continuous differentiable functions such that the function and all partial derivatives are bounded. Let $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$ and consider

$$F = f\big(X_{t_1}, \dots, X_{t_n}\big). \tag{9}$$

We will prove that $F \in \operatorname{dom} D$ and

$$D_{t,0} F = \sum_{j=1}^n \frac{\partial_j f}{\partial x_j} f\big(X_{t_1}, \dots, X_{t_n}\big) \mathbb{1}_{[0,t_j]}(t), \tag{10}$$

and for $x \neq 0$,

$$D_{t,x} F = \frac{f\big(X_{t_1} + x\mathbb{1}_{[0,t_1]}(t), \dots, X_{t_n} + x\mathbb{1}_{[0,t_n]}(t)\big) - f\big(X_{t_1}, \dots, X_{t_n}\big)}{x}. \tag{11}$$

Note the following relationship between both derivatives of a smooth functional:

$$D_{t,0} F = \lim_{x \to 0} D_{t,x} F, \text{ a.s.}$$

Geiss and Laukkarinen [4] (in the Lévy processes case) give a direct proof of (10) and (11) by using Fourier inversion and a Clark–Ocone–Haussman type formula. They also show that the random variables of form (9) are dense in $L^2(\mathbb{P})$ with respect to the norm induced by (5), and hence it is possible to define the Malliavin derivatives starting with (10) and (11). In order to prove these formulas in our context we will follow an alternative procedure: we will first prove these formulas in a canonical space associated with the process with independent increments and later we will transfer them to the general case.

2.5.1 Malliavin Derivatives in the Canonical Space

Since the Gaussian part and the jumps part of X are independent, we can construct a version of X in a canonical probability space of the form $\big(\Omega_G\times\Omega_N, \mathscr{A}_G\otimes\mathscr{A}_N, \mathbb{P}_G\otimes \mathbb{P}_N\big)$ where

- $(\Omega_G, \mathscr{A}_G, \mathbb{P}_G)$ is the canonical space associated with the Gaussian continuous process G; specifically, $\Omega_G = \mathscr{C}(\mathbb{R}_+)$ is the space of continuous functions on $\mathbb{R}_+$, $\mathscr{A}_G$ the Borel σ-algebra generated by the topology of the uniform convergence on compact sets, and $\mathbb{P}_G$ the probability that makes the projections

$$\begin{aligned} G_t^* :& \Omega_G \to \mathbb{R} \\ & f \mapsto f(t) \end{aligned}$$

  a process with the same law of $\{G_t,\ t\geq 0\}$.
- $(\Omega_N, \mathscr{A}_N, \mathbb{P}_N)$ is a canonical space associated with the Poisson random measure N. Essentially, Ω_N is formed by infinite sequences $\omega = \big((t_1,x_1),(t_2,x_2),\dots\big)\in \big((0,\infty)\times\mathbb{R}_0\big)^{\mathbb{N}}$ (see Appendix 3 for that construction), where t_i are the instants of jump of the process, and x_i the size of the corresponding jump. In this space, under $\mathbb{P}_N$, the mapping defined by

$$N^*(\omega) = \sum \delta_{(t_j,x_j)},\ \text{if } \omega = \big((t_1,x_1),(t_2,x_2),\dots\big)$$

  is a Poisson random measure with intensity measure ν.

  Define

$$J_t^* = \int_{(0,t]\times\{|x|>1\}} x\, N^*(\mathrm{d}(s,x)) + \int_{(0,t]\times\{0<|x|\leq 1\}} x\,\widehat{N}^*(\mathrm{d}(s,x)),$$

  where $\widehat{N}^* = N^* - \nu$. Then $J^* = \{J_t^*,\ t\geq 0\}$ is a process with independent increments with generating triplets $(0,\nu_t,0)$.

- Finally, in the product space $\Omega_G \times \Omega_N$ we write

$$X_t^* = m_t + G_t^* + J_t^*,$$

and call it the canonical version of the process X.

2.5.2 Derivative $D_{t,0}$

In order to compute the derivative $D_{t,0}F$ for $F \in L^2(\Omega_G \times \Omega_N)$, from the isometry

$$L^2(\Omega_G \times \Omega_N) \simeq L^2(\Omega_G; L^2(\Omega_N)),$$

we can consider F as an element of $L^2(\Omega_G; L^2(\Omega_N))$ and apply the theory of Malliavin derivatives of random variables with values in a separable Hilbert space following Nualart [17, p. 31]. This derivative coincides with $D_{t,0}$. This is proved from the fact that, by definition, a $L^2(\Omega_N)$-valued smooth random variable has the form

$$F = \sum_{i=1}^{n} F_i H_i,$$

where F_i are standard smooth variables (see Nualart [17, p. 25]) and $H_i \in L^2(\Omega_N)$. Define the Malliavin derivative of F as

$$D_t^* F = \sum_{i=1}^{n} D_t F_i \otimes H_i. \tag{12}$$

This definition is extended to a subspace $\operatorname{dom} D^*$ by a density argument.

Proposition 2 $\operatorname{dom} D^* \subset \operatorname{dom} D^0$, *and for* $F \in \operatorname{dom} D^*$,

$$D_t^* F = D_{t,0} F. \tag{13}$$

Proof First consider the functionals of the form

$$F = N^*(C_1) \cdots N^*(C_m) G^*(B_1) \cdots G^*(B_k),$$

where $C_1, \dots, C_m \in \mathscr{B}((0,\infty) \times \mathbb{R}_0)$ are bounded, pairwise disjoints, and at strictly positive distance of the t-axis, and $B_1, \dots, B_k \in \mathscr{B}(\mathbb{R}_+)$ are pairwise disjoints, with finite ρ measure. Itô [5] shows that the family of that functionals constitutes a fundamental set in $L^2(\mathbb{P}_G \otimes \mathbb{P}_N)$. Moreover, Itô shows that such an F can be written as a sum of multiple integrals:

$$F = I_0(f_0) + \cdots + I_{m+k}(f_{m+k}),$$

and then the derivatives are easy to compute, proving equality (13), which is extended to dom D^* by density. See Solé et al. [24]. □

From the above proposition and the properties of the Malliavin derivatives in the Gaussian white noise case, it follows that the first rule of differentiation (10) in the canonical space holds:

Proposition 3 *Let* $F = f\big(X^*_{t_1}, \ldots, X^*_{t_n}\big)$ *where* $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$. *Then* $F \in \mathrm{dom} D_{t,0}$ *and*

$$D_{t,0}F = \sum_{j=1}^{n} \frac{\partial_j f}{\partial x_j} f\big(X^*_{t_1}, \ldots, X^*_{t_n}\big)\mathbb{1}_{[0,t_j]}(t).$$

2.5.3 Derivative $D_{t,x}$, $x \neq 0$

Consider $\omega = (\omega^G, \omega^N) \in \Omega_G \times \Omega_N$, $\omega^N = \big((t_1,x_1),(t_2,x_2),\ldots\big) \in \big((0,\infty)\times \mathbb{R}_0\big)^{\mathbb{N}}$. Given $z = (t,x) \in (0,\infty)\times\mathbb{R}_0$, we add to ω_N a jump of size x at instant t, and call the new element $\omega^N_z = \big((t_1,x_1),(t_2,x_2),\ldots,(t,x),\ldots\big)$, and write $\omega_z = (\omega^G, \omega^N_z)$. For a random variable F, we define the quotient operator

$$\boldsymbol{\Psi}_{t,x}F(\omega) = \frac{F(\omega_{t,x}) - F(\omega)}{x}.$$

See Solé et al. [24] for the measurability properties of this function. By iteration, we define

$$\Psi^n_{z_1,\ldots,z_n}F := \Psi_{z_1}\Psi^{n-1}_{z_2,\ldots,z_n}F.$$

Since this function only depends on the part ω^N, we can assume that X does not have a Gaussian part.

In the following lemma we will consider a set Δ of the form $(m, m+1]\times\{x : n < |x| \le n+1\}$ or $(m,m+1]\times\{x : 1/(n+1) < |x| \le 1/n\}$, for some $m \ge 0$ and $n \ge 1$. Then $\nu(\Delta) < \infty$ and for every $k \ge 1$, $\int_\Delta |x|^k\,\nu(\mathrm{d}(t,x)) < \infty$. The Poisson random measure N^* restricted to Δ has finite intensity measure (from now on, in this section, we suppress the $*$ to simplify the notations). The ordinary n-fold product measure is denoted by $N^{\otimes n}$, and by $N^{(n)}$ the measure

$$N^{(n)}(D) = N^{\otimes n}(D_{\neq}),$$

where $D \in \mathscr{B}(\Delta^n)$ and $D_{\neq}$ is the set of elements of $(z_1,\ldots,z_n) \in D$ such that $z_i \neq z_j$ if $i \neq j$. The measure defined by $\mathbb{E}[N^{(n)}(D)]$ is called the n-factorial moment measure of the Poisson random measure N (see Last [11, formula (1.9)] or Schneider and Weil [23, p. 55]). For $F \in L^2(\Omega_N)$, for every $D \in \mathscr{B}(\Delta^n)$, the following integrals are finite and

$$\mathbb{E}\big[FN^{(n)}(D)\big] = \int_D \mathbb{E}\big[F(\omega_{z_1,\dots,z_n}))\big]\,\nu(\mathrm{d}z_1)\cdots\nu(\mathrm{d}z_n). \tag{14}$$

To deduce that equality, note that we can write $F = f(N)$, for some f defined on the set of integer valued (including ∞) locally finite measures (see Part 2). For $\omega = (z_1, z_2, \dots)$, $N(\omega) = \sum_i \delta_{z_i}$, and hence $N(\omega_z) = \sum_i \delta_{z_i} + \delta_z$. Then, equality (14) is just a reformulation of a generalized Mecke formula (see Last [11, formula (1.10)]).

As a consequence, we have

Lemma 1 *Let $F_k, F \in L^2(\mathbb{P}_N)$ such that* $\lim_k F_k = F$ *in $L^2(\mathbb{P}_N)$. Then for every $D \in \mathscr{B}(\Delta^n)$*

$$\lim_k \mathbb{E}\int_D \big|\Psi_{z_1,\dots,z_n}F_k - \Psi_{z_1,\dots,z_n}F\big|\,\nu(\mathrm{d}z_1)\cdots\nu(\mathrm{d}z_n) = 0.$$

Proof The proof is very similar to the proof of Lemma 2 of Last [11]. It suffices to show that for every $m = 1, \dots, n$,

$$\lim_k \mathbb{E}\int_D \left|\frac{F_k(\omega_{z_1,\dots,z_m}) - F(\omega_{z_1,\dots,z_m})}{x_1\cdots x_m}\right|\nu(\mathrm{d}z_1)\cdots\nu(\mathrm{d}z_n) = 0,$$

where $z_i = (t_i, x_i)$. Since Δ is bounded and far from 0, the $x_1, \dots, x_n$ in the denominator can be suppressed. By (14),

$$\mathbb{E}\int_D \big|F_k(\omega_{z_1,\dots,z_m}) - F(\omega_{z_1,\dots,z_m})\big|\,\nu(\mathrm{d}z_1)\cdots\nu(\mathrm{d}z_n) = \mathbb{E}\Big[\big|F_k - F\big|N^{(n)}(D)\Big]$$

$$\le \Big(\mathbb{E}\big[(F_k - F)^2\big]\mathbb{E}\big[N^{(n)}(D)^2\big]\Big)^{1/2},$$

which goes to 0 as $k \to \infty$. □

Proposition 4 *Let $F \in L^2(\mathbb{P}_G \otimes \mathbb{P}_N)$ such that*

$$E\Big[\int_{\mathbb{R}_+\times\mathbb{R}_0} \big(\boldsymbol{\Psi}_z F\big)^2\,\mu(dz)\Big] < \infty. \tag{15}$$

Then $F \in$ $\operatorname{dom} D^J$ *and*

$$D_z F(\omega) = \boldsymbol{\Psi}_z F(\omega), \quad \mu \otimes \mathbb{P} - \text{a.e. } (z, \omega) \in (0,\infty) \times \mathbb{R}_0 \times \Omega.$$

Proof To simplify the notations we write $\mu(\mathrm{d}(x,t))$ rather than $x^2\,\nu(\mathrm{d}(x,t))$. First it is proved that for $f \in L_s^2(\mu^n)$,

$$DI_n(f) = \boldsymbol{\Psi} I_n(f),\ \mu \otimes \mathbb{P} - \text{a.e.} \tag{16}$$

To prove this, first, instead of f we consider $f\mathbb{1}_{\Delta}^{\otimes n}$. Thus, as before, the multiple integrals (with respect to M) can be computed pathwise, and the above equality is easily checked, and then extended to f. Moreover, it is proved that the operator Ψ is closed, again working first with the restriction on Δ. Hence, if $F \in \operatorname{dom} D^J$, then $DF = \Psi F$. For the details see Solé et al. [24].

Note that as a consequence of (16),

$$f_n(z_1,\ldots,z_n) = \frac{1}{n!}\,\mathbb{E}\Big[\Psi^n_{z_1,\ldots,z_n} I_n(f_n)\Big],\ \mu^{\otimes n} - \text{a.e.}$$

This property is extended to a general $F = \sum_{n=0}^{\infty} I_n(f_n) \in L^2(\mathbb{P})$ to get a Stroock type formula

$$f_n = \frac{1}{n!}\,\mathbb{E}\big[\Psi^n F\big]\,\mu^{\otimes n} - \text{a.e.} \tag{17}$$

This is proved considering $F_k = \sum_{n=0}^{k} I_n(f_n)$. We have that for $k \geq n$, $f_n = \frac{1}{n!}\,\mathbb{E}\big[\Psi^n F_k\big]$. By Lemma 1, for every $D \in \mathscr{B}(\Delta^n)$

$$\lim_k \int_D \big|E\big[\Psi^n F_k\big] - E\big[\Psi^n F\big]\big|\,\mathrm{d}\nu^{\otimes n} = \int_D \big|n! f_n - E\big[\Psi^n F\big]\big|\,\mathrm{d}\nu^{\otimes n} = 0.$$

Then

$$f_n = \frac{1}{n!}\,\mathbb{E}\big[\Psi^n F\big],\ \nu^{\otimes n} - \text{a.e. on } \Delta,$$

and also $\mu^{\otimes n}$-a.e. And hence the equality holds on $(0,\infty) \times \mathbb{R}_0$ because it is a countable union of sets of type Δ.

Now assume that condition (15) holds. Then

$$\Psi_z F = \sum_{n=0}^{\infty} I_n(g_n(z,\cdot)),$$

with

$$\sum_{n=0}^{\infty} n! \int g_n^2\,\mathrm{d}\mu^{\otimes(n+1)} < \infty. \tag{18}$$

However, thanks to (17), the kernel g_n is related to the kernel f_{n+1} due to

$$g_n(z,z_1,\ldots,z_n) = \frac{1}{n!}\,\mathbb{E}\Big[\Psi^n_{z_1,\ldots,z_n}\Psi_z F\Big] = \frac{1}{n!}\,\mathbb{E}\Big[\Psi^{n+1}_{z_1,\ldots,z_n,z}F\Big] = (n+1)f_{n+1}(z,z_1,\ldots,z_n),$$

and by (18),

$$\sum_{n=1}^{\infty} n n! \int f_n^2\,\mathrm{d}\mu^{\otimes n} = \sum_{n=0}^{\infty} n! \int g_n^2\,\mathrm{d}\mu^{\otimes(n+1)} < \infty,$$

which is the condition for $F \in \operatorname{dom} D^J$. □

We can deduce the second rule of differentiation (11):

Proposition 5 *Let $F = f\big(X^*_{t_1},\ldots,X^*_{t_n}\big)$ where $f \in \mathscr{C}^\infty_b(\mathbb{R}^n)$. Then $F \in \operatorname{dom} D^J$ and for $x \neq 0$,*

$$D_{t,x}F = \frac{f\big(X^*_{t_1} + x\mathbb{1}_{[0,t_1]}(t),\ldots,X^*_{t_n} + x\mathbb{1}_{[0,t_n]}(t)\big) - f\big(X^*_{t_1},\ldots,X^*_{t_n}\big)}{x}.$$

Proof To shorten the notations we suppress the star in X^*_t. We consider the case $F = f\big(X_s\big)$; the general case is similar. We have

$$\Psi_{t,x}F = \frac{f\big(X_s + x\mathbb{1}_{[0,s]}(t)\big) - f\big(X_s\big)}{x}.$$

By Proposition 4 it suffices to prove that the following integral is finite:

$$\begin{aligned}
\mathbb{E}\left[\int_{(0,\infty)\times\mathbb{R}_0} \big(\boldsymbol{\Psi}_z F\big)^2\,\mu(\mathrm{d}z)\right] &= \mathbb{E}\left[\int_{(0,\infty)\times\mathbb{R}_0} \big(\boldsymbol{\Psi}_{t,x} F\big)^2\,x^2\nu(\mathrm{d}(t,x))\right] \\
&= \mathbb{E}\left[\int_{(0,\infty)\times\mathbb{R}_0} \left(\frac{f\big(X_s + x\mathbb{1}_{[0,s]}(t)\big) - f\big(X_s\big)}{x}\right)^2 x^2\nu(\mathrm{d}(t,x))\right] \\
&= \mathbb{E}\left[\int_{(0,s]\times\mathbb{R}_0} \Big(f\big(X_s + x\big) - f\big(X_s\big)\Big)^2\,\nu(\mathrm{d}(t,x))\right] \\
&= \mathbb{E}\left[\int_{\mathbb{R}_0} \Big(f\big(X_s + x\big) - f\big(X_s\big)\Big)^2\,\nu_s(\mathrm{d}x)\right].
\end{aligned}$$

To this end, by the mean value Theorem, there is a random point Y such that

$$f\big(X_s + x\big) - f\big(X_s\big) = xf'(Y).$$

Since f' is bounded, say by C, using that ν_s is a Lévy measure, for every ω,

$$\int_{\{|x|\le 1\}} \Big(F\big(X_s + x\big) - f\big(X_s\big)\Big)^2 \nu_s(\mathrm{d}x) \le C^2 \int_{\{|x|\le 1\}} x^2 \nu_s(\mathrm{d}x) = C' < \infty,$$

where C' is a constant independent of ω. Similarly,

$$\int_{\{|x|>1\}} \Big(f\big(X_s + x\big) - f\big(X_s\big)\Big)^2 \nu_s(\mathrm{d}x) \le C'''\nu_s\{x : |x| > 1\} = C''' < \infty,$$

where C'' and C''' are constants independent of ω. □

2.5.4 Transfer of the Derivative Rules from the Canonical Space to an Arbitrary Space

Recall that we write a star to denote random variables, measures, processes, or operators in the canonical space. We consider a process with independent increments X on $(\Omega, \mathscr{A}, \mathbb{P})$ with Poisson measure N and independent Gaussian part G, with the same law as N^* and G^* respectively, related to the additive process X^* constructed in the canonical space $\big(\Omega_G \times \Omega_N, \mathscr{A}_G \otimes \mathscr{A}_N, \mathbb{P}_G \otimes \mathbb{P}_N\big)$. Note that the generating triplets of X and X^* coincide, and hence the measures μ and μ^* (see (3)) are the same. Moreover, the Fock space structure of $L^2(\mathbb{P})$ allows us to transfer some properties of the derivatives and Skorohod integrals in the canonical space to the space $(\Omega, \mathscr{A}, \mathbb{P})$. This can be done thanks to the fact that to a square integrable random variable $F \in L^2(\mathbb{P})$ with

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L_s^2(\mu^n),$$

we can associate $F^* \in L^2(\mathbb{P}_G \times \mathbb{P}_N)$ given by

$$F^* = \sum_{n=0}^{\infty} I_n^*(f_n).$$

That is, the kernels of F and F^* are the same. In a similar way, since, given that $g \in L^2\big(\mathbb{R}_+ \times \mathbb{R} \times \Omega, \mathscr{B}(\mathbb{R}_+ \times \mathbb{R}) \otimes \mathscr{A}, \mu \otimes \mathbb{P}\big)$ has a chaotic decomposition

$$g(z) = \sum_{n=0}^{\infty} I_n(f_n(z,\cdot)), \tag{19}$$

$f_n \in L^2_{n+1}$ is symmetric in the n last variables, we can transfer from g to g^*, and if $g \in \operatorname{dom}\delta$, then $g^* \in \operatorname{dom}\delta^*$. More specifically,

Lemma 2 *With the previous notations, for every $t_1, \dots, t_n \in \mathbb{R}_+$ and $F \in L^2(\mathbb{P})$, we have that*

$$\big(X_{t_1}, \dots, X_{t_n}, F\big) \stackrel{\mathscr{L}}{=} \big(X^*_{t_1}, \dots, X^*_{t_n}, F^*\big), \tag{20}$$

where $\stackrel{\mathscr{L}}{=}$ means equality in law.

Proof We undertake the proof in several steps:

Step 1 Let $F = \sum_{n=0}^{\infty} I_n(f_n) \in L^2(\mathbb{P})$. We first prove that F and F^* have the same law:

$$\sum_{n=0}^{\infty} I_n(f_n) \stackrel{\mathscr{L}}{=} \sum_{n=0}^{\infty} I^*_n(f_n). \tag{21}$$

In fact, if the sum has a finite number of terms, and the f_n are simple (see the appendix) then the equality in law is clear. Equality (21) for finite sums with arbitrary kernels follows by $L^2(\mathbb{P})$-convergence. The infinite sum case is proved in a similar fashion.

Step 2 For $F, G \in L^2(\mathbb{P})$ we prove that

$$(F, G) \stackrel{\mathscr{L}}{=} (F^*, G^*).$$

We use Cramer–Wold device. Let $F = \sum_{n=0}^{\infty} I_n(f_n)$ and $G = \sum_{n=0}^{\infty} I_n(g_n)$. For $a, b \in \mathbb{R}$,

$$aF + bG = \sum_{n=0}^{\infty} I_n\big(af_n + bg_n\big) \stackrel{\mathscr{L}}{=} \sum_{n=0}^{\infty} I^*_n\big(af_n + bg_n\big) = aF^* + bG^*.$$

Step 3 To prove (20) we consider $n = 1$; the general case is similar. First assume that the process X is square integrable, then $\int_{R_0} x^2\, \nu_t(\mathrm{d}x) < \infty$, and thus $\int_{(0,t]\times\mathbb{R}_0} x^2\, \nu(\mathrm{d}(s,x)) < \infty$. This implies that the representation (2) admits the form:

$$X_t = m_t + G_t + \int_{(0,t]\times\{|x|>1\}} x\,\nu(\mathrm{d}(s,x)) + \int_{(0,t]\times\mathbb{R}_0} x\,\widehat{N}(\mathrm{d}(s,x))$$
$$= m_t + \int_{(0,t]\times\{|x|>1\}} x\,\nu(\mathrm{d}(s,x)) + I_1\big(\mathbb{1}_{[0,t]\times\{0\}} + \mathbb{1}_{(0,t]\times\mathbb{R}_0}\big),$$

and the property follows from step 2. In the general case, define

$$X_t^{(n)} = m_t + G_t + \int_{(0,t]\times\{1<|x|\le n\}} x\,N(\mathrm{d}(s,x)) + \int_{(0,t]\times\{0<|x|\le 1\}} x\,\widehat{N}(\mathrm{d}(s,x)) \tag{22}$$

$$= m_t + \int_{(0,t]\times\{1<|x|\le n\}} x\,\nu(\mathrm{d}(s,x)) + I_1\big(\mathbb{1}_{[0,t]\times\{0\}} + \mathbb{1}_{(0,t)\times\{0<|x|\le n\}}\big). \tag{23}$$

By expression (23) and Step 2, $\big(X_t^{(n)}, F\big) \stackrel{\mathscr{L}}{=} \big(X_t^{(n)*}, F^*\big)$. Since $\nu\big((0,t]\times\{|x|>1\}\big) < \infty$, we can apply Proposition 10 in the appendix to the first integral in the expression (22), and we deduce that when $n\to\infty$, $X_t^{(n)} \to X_t$ in probability, and the lemma follows. □

To transfer the derivative rules we will use the duality coupling (8). By construction, $F \in L^2(\mathbb{P})$ belongs to $\operatorname{dom} D$ if and only if there is a constant C such that for all $g \in \operatorname{dom}\delta$,

$$\Big|\mathbb{E}\big[F\delta(g)\big]\Big| \le C\left(\mathbb{E}\left[\int_{\mathbb{R}_+\times\mathbb{R}} g^2\,\mathrm{d}\mu\right]\right)^{1/2}. \tag{24}$$

If $F \in \operatorname{dom} D$, then DF is the element of $L^2(\mu\otimes\mathbb{P})$ characterized by

$$\mathbb{E}[\delta(g)\,F] = \mathbb{E}\int_{\mathbb{R}_+\times\mathbb{R}} g(z)\,D_zF\,d\mu(z), \tag{25}$$

for every $g \in \operatorname{dom}\delta$. That is, we use (8) to prove a property of the derivative from the Skorohod integral.

Proposition 6 *Let* $F = f\big(X_{t_1}, \dots, X_{t_n}\big)$ *with* $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$. *Then* $F \in \operatorname{dom} D$ *and*

$$D_{t,0}F = \sum_{j=1}^n \frac{\partial_j f}{\partial x_j} f\big(X_{t_1}, \dots, X_{t_n}\big) \mathbb{1}_{[0,t_j]}(t), \tag{26}$$

and for $x \neq 0$,

$$D_{t,x}F = \psi_{t,x}F = \frac{f\big(X_{t_1} + x\mathbb{1}_{[0,t_1]}(t), \dots, X_{t_n} + x\mathbb{1}_{[0,t_n]}(t)\big) - f\big(X_{t_1}, \dots, X_{t_n}\big)}{x}. \tag{27}$$

Proof We are going to prove that $F \in \operatorname{dom} D$. For this objective, let $g \in \operatorname{dom}\delta$ and consider g^* to have the same kernels as g, and then $g^* \in \operatorname{dom}\delta^*$ and satisfies inequality (24). Then, since we have proved that $f\big(X^*_{t_1}, \dots, X^*_{t_n}\big) \in \operatorname{dom} D$, by Lemma 2,

$$\begin{aligned}\Big|\mathbb{E}\big[f\big(X_{t_1}, \dots, X_{t_n}\big)\delta(g)\big]\Big| &= \Big|\mathbb{E}^*\big[f\big(X^*_{t_1}, \dots, X^*_{t_n}\big)\delta^*(g^*)\big]\Big| \\ &\leq C\left(\mathbb{E}^*\left[\int_{\mathbb{R}_+\times\mathbb{R}} (g^*)^2 \, \mathrm{d}\mu\right]\right)^{1/2} = C\left(\mathbb{E}\left[\int_{\mathbb{R}_+\times\mathbb{R}} g^2 \, \mathrm{d}\mu\right]\right)^{1/2} < \infty,\end{aligned}$$

where $\mathbb{E}^*$ is the expectation in $\Omega_G \times \Omega_N$. Now in an identical way, we can show that

$$\begin{aligned}Y_{t,x} := &\sum_{j=1}^n \frac{\partial_j f}{\partial x_j} f\big(X_{t_1}, \dots, X_{t_n}\big) \mathbb{1}_{[0,t_j]}(t)\mathbb{1}_{\{0\}}(x) \\ &+ \frac{f\big(X_{t_1} + x\mathbb{1}_{[0,t_1]}(t), \dots, X_{t_n} + x\mathbb{1}_{[0,t_n]}(t)\big) - f\big(X_{t_1}, \dots, X_{t_n}\big)}{x} \mathbb{1}_{\{x\neq 0\}}(x)\end{aligned}$$

satisfies formula (25). □

2.6 Characterization of Processes with Independent Increments by Duality Formulas

Following Murr [15] we prove that the duality formula (8) characterizes the law of a process with independent increments. We restrict ourselves to real processes, while Murr [15] studies the vector case. Like Murr [15] we assume that the process is integrable. The fact that the process is integrable is equivalent to

$\int_{\{|x|>1\}} |x| \,\mathrm{d}\nu_t(x) < \infty$. Then, as in the proof of Lemma 2, we can write the following representation:

$$X_t = b_t + G_t + \int_{(0,t]\times\{|x|>1\}} x\,\widehat{N}(\mathrm{d}(s,x)) + \int_{(0,t]\times\{0<|x|\le 1\}} x\,\widehat{N}(\mathrm{d}(s,x)),$$

where $b_t = m_t + \int_{(0,t]\times\{|x|>1\}} x\,\nu(\mathrm{d}(s,x))$, the first integral belongs to $L^1(\mathbb{P})$ and the second to $L^2(\mathbb{P})$ (see Theorem 7 in the appendix).

Consider the system of generating triplets of X (with respect to the cutoff function $\chi(x) = x$) $\{(b_t, \rho_t, \nu_t),\ t \ge 0\}$. As we commented in Sect. 2.1 (see Sato [22, Theorem 9.8]):

1. $b_0 = 0$ and the function $t \mapsto b_t$ is continuous.
2. $\rho_0 = 0$, $\rho_t \ge 0$ and the function $t \mapsto \rho_t$ is increasing and continuous.
3. For every $t \ge 0$, ν_t is a Lévy measure, and $\lim_{s\to t} \nu_s(B) = \nu_t(B)$ for every $B \in \mathscr{B}(\mathbb{R})$ such that $B \subset \{x : |x| > \varepsilon\}$ for some $\varepsilon > 0$.
4. For every $t \ge 0$, $\int_{\{|x|>1\}} |x| \,\mathrm{d}\nu_t(x) < \infty$.

Denote by $\mathbb{S}$ the set of random variables of the form $F = f\big(X_{t_1}, \ldots, X_{t_n}\big)$ with $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$, and by $\mathscr{E}$ the set of real step functions $g = \sum_{j=1}^k a_j \mathbb{1}_{(s_j, s_{j+1}]}$, with $0 \le s_1 < \cdots < s_{k+1}$. In the next theorem we add conditions regarding the regularity of the trajectories to agree with our definitions.

Theorem 2 (Murr) *Let X be an integrable process, cadlag and continuous in probability, and $\{(b_t, \rho_t, \nu_t),\ t \ge 0\}$ be such that (1)–(4) above are satisfied. Then X is a process with independent increments with system of generating triplets $\{(b_t, \rho_t, \nu_t),\ t \ge 0\}$ if and only if for every $F \in \mathbb{S}$, and every step function $g \in \mathscr{E}$,*

$$\mathbb{E}\left[F \int_{\mathbb{R}_+} g(t)\,\mathrm{d}(X_t - b_t)\right] = \mathbb{E}\left[\int_{\mathbb{R}_+} D_{t,0}F\, g(t)\,\rho(\mathrm{d}t)\right] + \mathbb{E}\left[\int_{(0,t]\times\mathbb{R}_0} \Psi_{t,x}F\, g(t) x^2\,\nu(\mathrm{d}(t,x))\right], \tag{28}$$

where ν is defined in (1).

Proof Assume that X is a process with independent increments. To prove (28), by linearity, it suffices to consider $g = \mathbb{1}_{[0,u]}$, So we will check

$$\mathbb{E}\left[F\big(X_u - b_u\big)\right] = \mathbb{E}\left[\int_{[0,u]} D_{t,0}F\,\rho(\mathrm{d}t)\right] + \mathbb{E}\left[\int_{(0,u]\times\mathbb{R}_0} \Psi_{t,x}F\, x^2\,\nu(\mathrm{d}(t,x))\right]. \tag{29}$$

Note that for a deterministic function $h \in L^2(\mathbb{R} \times \mathbb{R}_+, \mu)$ the duality formula (8) gives

$$\mathbb{E}\left[F \int_{\mathbb{R}_+ \times \mathbb{R}} h \,\mathrm{d}M\right] = \mathbb{E}\left[\int_{\mathbb{R}_+} D_{t,0}F\, h(t,0)\, \rho(\mathrm{d}t)\right] + \mathbb{E}\left[\int_{(0,\infty)\times\mathbb{R}_0} \Psi_{t,x}F\, h(t,x)x^2\, \nu(\mathrm{d}(t,x))\right].$$

Set

$$h_n(t,x) = \mathbb{1}_{[0,u]\times\{0\}} + x\,\mathbb{1}_{[0,u]\times\{1<|x|\le n\}}(t,x) + x\,\mathbb{1}_{[0,u]\times\{0<|x|\le 1\}}(t,x)$$

that belongs to $L^2(\mu)$, then

$$\int_{\mathbb{R}_+\times\mathbb{R}} h_n \,\mathrm{d}M = G_u + \int_{(0,u]\times\{1<|x|\le n\}} x\,\widehat{N}(\mathrm{d}(s,x)) + \int_{(0,t]\times\{0<|x|\le 1\}} x\,\widehat{N}(\mathrm{d}(s,x)).$$

In relation to the first integral in the right-hand side, note that $x\mathbb{1}_{(0,u]\times\{1<|x|\le n\}}$ belongs to $L^1(\nu)\cap L^2(\nu)$, and

$$\lim_n \int_{(0,u]\times\{1<|x|\le n\}} x\,\widehat{N}(\mathrm{d}(s,x)) = \int_{(0,u]\times\{|x|>1\}} x\,\mathrm{d}\widehat{N}(\mathrm{d}(s,x))$$

in $L^1(\mathbb{P})$, and hence $\int_{\mathbb{R}_+\times\mathbb{R}} h_n\,\mathrm{d}M$ converges in $L^1(\mathbb{P})$ to $X_u - b_u$. Since F is bounded, it follows (29).

To prove the reciprocal implication, Murr [15] fixes $g = \sum_{j=1}^k a_j \mathbb{1}_{(s_j,s_{j+1}]}$, with $0 \le s_1 < \cdots < s_{k+1}$, and for $u \in \mathbb{R}$, defines

$$\varphi(u) = \mathbb{E}\left[\exp\left\{iu\int_{\mathbb{R}_+} g\,dX\right\}\right].$$

Since

$$\varphi'(u) = i\mathbb{E}\left[\exp\left\{iu\int_{\mathbb{R}_+} g\,dX\right\}\int_{\mathbb{R}_+} g\,dX\right],$$

applying the duality formula (28) with $F = \exp\{iu\int_{\mathbb{R}_+} g\,dX\}$ it is deduced a differential equation, which for $u = 1$ determines the characteristic function of $(X_{s_1}, X_{s_2} - X_{s_1}, \ldots, X_{s_{k+1}} - X_{s_k})$, which determines the law of the process, and the theorem follows. □

Remark 3 Murr [15] defines $\Psi_{t,x}F$ as

$$\Psi_{t,x}F = f\big(X_{t_1} + x\mathbb{1}_{[0,t_1]}(t), \ldots, X_{t_n} + x\mathbb{1}_{[0,t_n]}(t)\big) - f\big(X_{t_1}, \ldots, X_{t_n}\big),$$

whereas in our definition of Ψ given in (27) we divide by x. However in the second term in the right-hand side of formula (28) Murr puts x rather than x^2. Of course, both formulations are equivalent.

3 Part 2: Random Measures

The context of this part is one of the random measures a.s. locally finites on a locally compact second countable Hausdorff space; the main references here are Kallenberg [6] and Schneider and Weyl [23]. In this part we use standard notations of random measures.

3.1 Random Measures

Let $\mathbb{X}$ be a locally compact second countable Hausdorff space; it can be proved that this space is Polish (complete separable metrizable space). Denote by $\mathcal{X}$ its Borel σ-field. A measure χ on $(\mathbb{X}, \mathcal{X})$ is *locally finite* if $\chi(K) < \infty$ for every compact set K; note that such a measure is σ-finite.

Denote by $\mathbf{M}$ (or $\mathbf{M}(\mathbb{X})$ if we want to stress the underlying space) the set of locally finite measures on $(\mathbb{X}, \mathcal{X})$ and endow this space with the σ-field $\mathscr{M}$ generated by the evaluation maps. We also denote by $\mathbf{N}$ the subset of locally finite measures taking values in $\{0, 1, \ldots\} \cup \{\infty\}$. This notation is consistent with the one adopted in the survey [14] in this volume.

Given a random measure ξ on $(\mathbb{X}, \mathcal{X})$ with intensity λ, remember that it is said that $s \in \mathbb{X}$ is a fixed atom of ξ if $\mathbb{P}\{\xi\{s\} > 0\} > 0$. Note that if ξ has no fixed atoms, then for every $s \in \mathbb{X}$, $\lambda\{s\} = \mathbb{E}\big[\xi\{s\}\big] = 0$, so the intensity measure is non-atomic.

3.2 Infinitely Divisible Random Measures and Random Measures with Independent Increments

It is said that the random measure ξ has **independent increments** if for any family of pairwise disjoint sets $A_1, \ldots, A_k \in \mathcal{X}$, the random variables $\xi(A_1), \ldots, \xi(A_k)$ are independent. Matthes et al. [13, p. 16] call these random measures *free from after-effects*, and Kingman [8, 9] *completely random measures*.

A random measure ξ is said to be **infinitely divisible** if for every $n \geq 1$ there are random measures $\xi_1, \ldots, \xi_n$ such that they are independent, and ξ has the same law as $\xi_1 + \cdots + \xi_n$. Indeed, every random measure with independent increments without fixed atoms is infinitely divisible (Kallenberg [6, Chap. 7]). The nice Lévy–Itô decomposition of processes with independent increments in terms of a Poisson random measures (Theorem 1) is transferred to random measures with independent increments; general infinitely divisible random measures have a representation in law (Kallenberg [6, Theorem 8.1]).

Before the representation theorem it is convenient to comment that since the number of fixed atoms of a random measure is at most countable (Kallenberg, [6, p. 56]), if ξ is a random measure with independent increments it can be written as

$$\xi = \sum_{n=1}^{N} \xi(\{s_n\})\, \delta_{s_n} + \xi',$$

with $N \leq \infty$, where $\{s_n,\, n \geq 1\}$ is the set of fixed atoms of ξ, and ξ' is a random measure without fixed atoms with independent increments. So, as Kingman [9] graphically says, fixed atoms can be removed by simple surgery.

Theorem 3 *Let ξ be a random measure with independent increments with intensity measure λ, without fixed atoms. Then it can be represented uniquely in the form*

$$\xi(A) = \beta(A) + \int_{A\times(0,\infty)} x\, \eta(\mathrm{d}(s,x)), \tag{30}$$

for $A \in \mathcal{X}$, where $\beta \in \mathbf{M}(\mathbb{X})$ is non-atomic, and η is a Poisson random measure on $\mathbb{X} \times (0,\infty)$ which intensity measure $\nu \in \mathbf{M}(\mathbb{X} \times (0,\infty))$ non-atomic. Moreover, for $A \in \mathcal{X}$, we have $\xi(A) < \infty,$ a.s. if and only if $\beta(A) < \infty$ and

$$\int_{A\times(0,\infty)} (1 \wedge x)\, \nu(\mathrm{d}(s,x)) < \infty.$$

For a proof see Kallenberg [7, Corollary 12.11] in the context of Borel spaces or Daley and Vere–Jones [2, Theorem 10.1.III] for Polish spaces.

Remark 4 We comment some key points used in the proof that we need later:

1. The measure ν on $\mathbb{X} \times (0, \infty)$ comes from

$$\nu(A \times B) = \nu_A(B),$$

where $A \in \mathcal{X}$ with $\lambda(A) < \infty$, and $B \in \mathscr{B}((0, \infty))$, and ν_A is a Lévy measure on $(0, \infty)$. That Lévy measure is associated with the positive infinitely divisible random variable with finite expectation $\xi(A)$, and then it integrates the function $f(x) = x$. So, for $A \in \mathcal{X}$ with $\lambda(A) < \infty$, we have

$$\int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(s, x)) < \infty,$$

and

$$\mathbb{E}[\xi(A)] = \lambda(A) = \beta(A) + \int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(s, x)). \tag{31}$$

2. The Poisson random measure η is given by

$$\eta = \sum_{s\in\mathbb{X}} \delta_{(s,\xi\{s\})}.$$

Since it is measurable (see Kallenberg [7], proof of Corollary 12.11) it follows that the σ-fields generated by ξ and η coincide. We will assume that $\mathscr{A}$ is that σ-field.

3. The Laplace functional of ξ at $h : \mathbb{X} \to \mathbb{R}_+$ is

$$\mathbb{E}\left[\exp\left\{-\int_{\mathbb{X}} h\, \mathrm{d}\xi\right\}\right] = \exp\left\{-\int_{\mathbb{X}} h \mathrm{d}\beta - \int_{\mathbb{X}\times(0,\infty)} \left(1 - e^{-xh(s)}\right)\nu(\mathrm{d}(s, x))\right\}. \tag{32}$$

Example: Subordinators A subordinator $X = \{X_t,\ t \geq 0\}$ is a Lévy process such that the trajectories are increasing a.s. Then it defines a random measure on $\mathbb{X} = \mathbb{R}_+$. Representation (3) corresponds to the Lévy–Itô decomposition of X (Theorem 1) which, with the notations of Part 1, is reduced to (see Sato [22, Theorem 21.5])

$$X_t = \gamma^\circ t + \int_{(0,t]\times(0,\infty)} x N(\mathrm{d}(s, x)),$$

where $\gamma^\circ \geq 0$ and N is a Poisson random measure on $(0,\infty)\times(0,\infty)$ with intensity $\nu(\mathrm{d}(t,x)) = \mathrm{d}t\,\nu^\circ(\mathrm{d}x)$, where ν° is the Lévy measure of X (see Remark 1.3), and the Gaussian part is 0. For every $t > 0$, $\int_{(0,t]\times(0,\infty)} (1\wedge x)\,\nu^\circ(\mathrm{d}x) < \infty$, and the intensity measure of the random measure is given by

$$\lambda\big([0,t]\big) = \gamma_0 t + t \int_{(0,\infty)} x\,\nu^\circ(\mathrm{d}x),$$

which, in general, can be infinite.

3.3 Mecke Formula for Random Measures with Independent Increments

We prove Mecke formula for random measures with independent increments which is inspired in Murr [15]. We first recall classical Mecke formula for Poisson processes (Last [11, formula (1.7)], Privault [21, formula (2.44)]); see Schneider and Weil [23, Theorem 3.2.5] for the following version of the formula, which we use later.

Theorem 4 (Mecke Formula for Poisson Random Measures) *Let γ be a point process with non-atomic intensity measure $\lambda \in \mathbf{M}(\mathbb{X})$. Then γ is a Poisson random measure if and only if for every measurable function $h : \mathbf{N}(\mathbb{X})\times\mathbb{X} \to \mathbb{R}_+$ we have*

$$\mathbb{E}\left[\int_{\mathbb{X}} h(\gamma,s)\,\gamma(\mathrm{d}s)\right] = \int_{\mathbb{X}} \mathbb{E}\left[h(\gamma+\delta_s,s)\right]\,\lambda(\mathrm{d}s). \tag{33}$$

Theorem 5 (Mecke Formula for Random Measures with Independent Increments) *Let ξ be a random measure without fixed atoms and let $\beta \in \mathbf{M}(\mathbb{X})$ be non-atomic and $\nu \in \mathbf{M}(\mathbb{X}\times(0,\infty))$ be non-atomic. Then ξ is a random measure with independent increments with associated measures β and ν if and only if for every measurable function $h : \mathbf{M}(\mathbb{X})\times\mathbb{X}\to\mathbb{R}_+$ we have*

$$\mathbb{E}\left[\int_{\mathbb{X}} h(\xi,s)\,\xi(\mathrm{d}s)\right] = \int_{\mathbb{X}} \mathbb{E}\big[h(\xi,s)\big]\,\beta(\mathrm{d}s) + \int_{\mathbb{X}\times(0,\infty)} \mathbb{E}\big[h(\xi+x\delta_s,s)\big]\,x\,\nu(\mathrm{d}(s,x)). \tag{34}$$

Proof

1. Let ξ be a random measure with independent increments with associated measures β and ν. First note that since β is a deterministic measure, changing ξ by $\xi-\beta$, and changing the function h conveniently, we can assume that $\beta = 0$.

We will reduce the proof to an easy case, and later we prove formula (34) in that case.

By standard arguments, it suffices to prove formula (34) for $h(x, s) = f(x)g(s)$ where $f : \mathbf{M}(\mathbb{X}) \to \mathbb{R}_+$ is bounded and $g = \mathbb{1}_C$ for some $C \in \mathcal{X}$ with $\lambda(C) < \infty$. Now, given that $\mathscr{M}(\mathbb{X})$ is generated by the projections π_A, for $A \in \mathcal{X}$, there is a countable family $\{A_n,\ n \geq 1\} \subset \mathcal{X}$ and a measurable function $F : \mathbb{R}^\infty \to \mathbb{R}_+$ such that

$$f = F\big(\pi_{A_1}, \pi_{A_2}, \dots\big).$$

(See Chow and Teicher [1, p. 17].) Hence,

$$f(\xi) = F\big(\xi(A_1), \xi(A_2), \dots,\big).$$

Denote by $\mathscr{A}_n$ the σ-field generated by $\xi(A_1), \dots, \xi(A_n)$, and define

$$F_n = \mathbb{E}\big[f(\xi)\,|\mathscr{A}_n\big].$$

By the convergence of martingales theorem we have that

$$\lim_n F_n = f, \text{ a.s.}$$

and since f is bounded, the convergence is also in L^p, for all $p \geq 1$. Hence, there is enough to consider the case

$$f(\xi) = f\big(\xi(A_1), \dots, \xi(A_n)\big).$$

With a monotone class argument, we can restrict to

$$f(\xi) = f_1\big(\xi(A_1)\big)\cdots f_n\big(\xi(A_n)\big),$$

with bounded $f_1, \dots, f_n \geq 0$, and $A_1, \dots, A_n$ pairwise disjoint. Using that ξ has independent increments, in formula (34) with such an $f(\xi)$ and $g = \mathbb{1}_C$, it is clear that we need only to consider two cases: when C is disjoint with all A_j, $j = 1, \dots, n$, or when C coincides with one of the A_j. In the first case equality (34) is reduced to check that if $A \cap C = \emptyset$, then

$$\mathbb{E}\big[f\big(\xi(A)\big)\xi(C)\big] = \int_{C\times(0,\infty)} \mathbb{E}\big[f(\xi(A) + x\delta_s(A))\big]\, x\, \nu(\mathrm{d}(s,x)),$$

that is evident since, thanks to (31) and the independence between $\xi(A)$ and $\xi(C)$, both sides are equal to $\mathbb{E}\big[f\big(\xi(A)\big)\big]\,\lambda(C)$.

In the second case (remember that here $\beta = 0$), equality (34) simplifies as

$$\mathbb{E}\big[f\big(\xi(A)\big)\xi(A)\big] = \int_{A\times(0,\infty)} \mathbb{E}\big[f(\xi(A) + x\delta_s(A))\big]\, x\, \nu(\mathrm{d}(s,x)). \tag{35}$$

Changing $\xi(A)$ by its expression in representation Theorem 3, in the left-hand side of (35) we have

$$\mathbb{E}\left[f\Big(\int_{A\times(0,\infty)} x\,\eta(\mathrm{d}(s,x))\Big)\int_{A\times(0,\infty)} x\,\eta(\mathrm{d}(s,x))\right], \tag{36}$$

where η is a Poisson random measure on $\mathbb{X}\times(0,\infty)$ with intensity measure ν. By Mecke formula for Poisson random measures (33),

$$(36) = \int_{A\times(0,\infty)} \mathbb{E}\left[f\left(\int_{A\times(0,\infty)} x\,\eta(\mathrm{d}(s,x)) + x\delta_{(s,x)}\big(A\times(0,\infty)\big)\right)\right] x\,\nu(\mathrm{d}(s,x)),$$

that is exactly the right-hand side of (35).

2. We prove the reciprocal implication. This proof is also inspired in Murr [15]. Note that applying formula (34) to the function $h(\mu,s) = f(s)$ we have

$$\int_{\mathbb{X}} f\,\mathrm{d}\lambda = \mathbb{E}\left[\int_{\mathbb{X}} f\,\mathrm{d}\xi\right] = \int_{\mathbb{X}} f\,\mathrm{d}\beta + \int_{\mathbb{X}\times(0,\infty)} xf(s)\,\nu(\mathrm{d}(s,x)).$$

Fix $g:\ \mathbb{X}\to\mathbb{R}_+$ measurable with $\int_{\mathbb{X}} g\,\mathrm{d}\lambda < \infty$. and define, for $u>0$,

$$G(u) = \mathbb{E}\left[\exp\{-u\int_{\mathbb{X}} g\,\mathrm{d}\xi\}\right].$$

Since $\mathbb{E}\big[\int_{\mathbb{X}} g\,\mathrm{d}\xi\big] < \infty$, by differentiation we get

$$G'(u) = -\mathbb{E}\left[\exp\{-u\int_{\mathbb{X}} g\,\mathrm{d}\xi\}\int_{\mathbb{X}} g\,\mathrm{d}\xi\right].$$

Now in formula (34) take

$$h(\mu, s) = \exp\{-u \int_{\mathbb{X}} g \, \mathrm{d}\mu\} \, g(s),$$

and then,

$$G'(u) = -\int_{\mathbb{X}} G(u) g(s) \, \beta(\mathrm{d}s) - \int_{\mathbb{X}\times(0,\infty)} G(u) \exp\{-uxg(s)\} g(s) x \, \nu(\mathrm{d}(s,x)),$$

or

$$\frac{G'(u)}{G(u)} = -\int_{\mathbb{X}} g(s) \, \beta(\mathrm{d}s) - \int_{\mathbb{X}\times(0,\infty)} \exp\{-uxg(s)\} g(s) x \, \nu(\mathrm{d}(s,x)).$$

The function on the right-hand side is continuous in u, and given that $G(0) = 1$ we have the

$$\begin{aligned} G(u) &= \exp\left\{ -\int_0^u \left(\int_{\mathbb{X}} g(s) \, \beta(\mathrm{d}s) + \int_{\mathbb{X}\times(0,\infty)} \exp\{-zxg(s)\} g(s) x \, \nu(\mathrm{d}(s,x)) \right) \mathrm{d}z \right\} \\ &= \exp\left\{ -u \int_{\mathbb{X}} g(s) \, \beta(\mathrm{d}s) - \int_{\mathbb{X}\times(0,\infty)} \left(\int_0^u \exp\{-zxg(s)\} \, \mathrm{d}z \right) g(s) x \, \nu(\mathrm{d}(s,x)) \right\}. \end{aligned}$$

In particular, for $u = 1$ we get

$$\int_0^1 \exp\Big\{ -zxg(s) \Big\} \, \mathrm{d}z = \mathbb{1}_{\{s:\, g(s)>0\}}(s) \frac{1}{xg(s)} \Big(1 - e^{-xg(s)} \Big),$$

and then the Laplace functional of ξ is

$$G(1) = \exp\left\{ -\int_{\mathbb{X}} g(s) \, \mathrm{d}\beta(s) - \int_{\mathbb{X}\times(0,\infty)} \left(1 - e^{-xg(s)} \right) \nu(\mathrm{d}(s,x)) \right\},$$

which corresponds to the claimed random measure (see (32)) □

3.4 Malliavin Calculus

From now on, we consider the random measure with independent increments given by

$$\xi(A) = \beta(A) + \int_{A\times(0,\infty)} x\,\eta(\mathrm{d}(s,x)), \tag{37}$$

where η is a Poisson random measure with intensity ν.

As in Part 1, we construct a completely random measure on $\mathbb{X}\times(0,\infty)$. With that purpose, define a new measure μ on $\mathbb{X}\times(0,\infty)$ by

$$\mu(\mathrm{d}(s,x)) = x^2\,\nu(\mathrm{d}(s,x)).$$

For $C \in \mathcal{X}\times\mathscr{B}(0,\infty)$ such that $\mu(C) < \infty$, the function $\mathbb{1}_C(s,x)x$ is in $L^2(\nu)$; hence the following random variable is well defined (as a limit in $L^2(\mathbb{P})$):

$$M(C) := \int_{\mathbb{X}\times(0,\infty)} \mathbb{1}_C(s,x)x\hat{\eta}(\mathrm{d}(s,x)) = \int_C x\,\hat{\eta}(\mathrm{d}(s,x)),$$

where $\hat{\eta} = \eta - \nu$. It is a completely random measure. As before, consider the set of symmetric functions

$$L_s^2(\mu^n) = L_s^2\Big(\big(\mathbb{X}\times(0,\infty)\big)^n, \mu^{\otimes n}\Big).$$

The multiple Itô integral of order n with respect to M of a function $f \in L_s^2(\mu^n)$ is denoted by $I_n(f)$. Itô chaotic representation property is also true in this context, and we have that $F \in L^2(\mathbb{P})$ admits a representation of the form

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L_s^2(\mu^n). \tag{38}$$

So we can define as in Part 1 a Malliavin derivative D with domain $\operatorname{dom} D$ and its dual, the Skorohod integral δ in $\operatorname{dom}\delta$.

3.4.1 Malliavin Derivatives with Respect to the Underlying Poisson Random Measure

In the present context of random measures, the absence of the Gaussian part and the fact that the integral in the representation (37) is pathwise make things easier, and we do not need to introduce a canonical space. As we commented in the

Introduction, we rely on the very general construction of Last and Penrose [12] and Last [11] (see also Privault [21] for multiple Poisson integrals). Denote by $I_n^{\hat{\eta}}(f)$ the multiple integral of order n with respect to $\hat{\eta}$ of a function $f \in L_s^2(\nu^n)$. For $f : \big(\mathbb{X} \times (0,\infty)\big)^n \to \mathbb{R}$ write

$$f^*\big((s_1,x_1),\ldots,(s_n,x_n)\big) = x_1 \cdots x_n f\big((s_1,x_1),\ldots,(s_n,x_n)\big).$$

Obviously we have that $f \in L_s^2(\mu^n)$ if and only if $f^* \in L_s^2(\nu^n)$. In this case,

$$I_n(f) = I_n^{\hat{\eta}}(f^*).$$

This is proved by standard techniques by considering first the case of elementary functions and by using a density argument.

Hence, for $F \in L^2(\mathbb{P})$ with an expansion (38) (remember that the σ-field generated by ξ and η coincide) we have also the expansion

$$F = \sum_{n=0}^{\infty} I_n^{\hat{\eta}}(f_n^*).$$

Last and Penrose [12] (see Last [11, Theorem 3]) introduce two derivative operators, the first one as an *add-one-cost operator*, that we comment in next subsection, and a Malliavin derivative D^η (Last denotes it by D') as an annihilation operator on the chaos expansion. The relation between our derivative D and D^η is the following:

Proposition 7 *We have* $\operatorname{dom} D = \operatorname{dom} D^\eta$, *and for* $F \in \operatorname{dom} D$,

$$D_{(s,x)}F = \frac{1}{x} D^\eta_{(s,x)}F, \;\; \mu \otimes \mathbb{P} - a.e.$$

Proof The proof is direct from the chaos expansion of F. □

3.4.2 Derivation of Smooth Functionals

We first prove a property for the Poisson process case: following Last and Penrose [12] and Last [11], consider a square integrable random variable $F \in L^2(\mathbb{P})$; since it is measurable with respect to the σ-field generated by η, there is a measurable function $f : \mathbf{N}(\mathbb{X} \times (0,\infty)) \to \mathbb{R}$ such that $F = f(\eta)$ and $\mathbb{E}[f^2(\eta)] < \infty$. Define

$$\mathbb{D}_z^\eta f(\eta) = f(\eta + \delta_z) - f(\eta).$$

By iteration, let

$$\mathbb{D}^{\eta,n}_{z_1,\ldots,z_n} f(\eta) = \mathbb{D}^\eta_{z_1} \mathbb{D}^{\eta,n-1}_{z_2,\ldots,z_n} f(\eta).$$

Now define $T_0 f = \mathbb{E}\big[f(\eta)\big]$, and for $n \geq 1$,

$$T_n f(z_1, \dots, z_n) = \mathbb{E}\big[\mathbb{D}^n_{z_1,\dots,z_n} f(\eta)\big].$$

This operator verifies that $T_n f \in L^2_s(\nu^n)$, and in the (Poisson) chaotic decomposition of $F = f(\eta)$

$$F = \sum_{n=0}^{\infty} I_n^{\hat{\eta}}(f_n),$$

the kernels are

$$f_n = \frac{1}{n!} T_n f.$$

See Last [11, Theorem 2].

Proposition 8 *Let* $F \in L^2(\mathbb{P})$. *Then* $F \in \operatorname{dom} D^\eta$ *if and only if* $\mathbb{D}^\eta F \in L^2(\Omega \times \mathbb{X} \times (0,\infty), \mathbb{P} \otimes \nu)$.

Proof If $F \in \operatorname{dom} D^\eta$ then the property follows from the coincidence between D^η and $\mathbb{D}^\eta$ (Last [11, equality (1.48)]). The proof of the reciprocal implication is analogous to the proof of the second part of Proposition 4. □

Now we return to Malliavin derivatives with respect to the random measure ξ.

Proposition 9 *Let* $A_1, \dots . A_n \in \mathbb{X}$*, with finite* λ *measure. Let* $F = f(\xi(A_1), \dots, \xi(A_n))$*, with* $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$. *Then* $F \in \operatorname{dom} D$ *and*

$$D_{s,x} F = \frac{1}{x}\Big(f\big(\xi(A_1) + x\delta_s(A_1), \dots, \xi(A_n) + x\delta_s(A_n)\big) - f\big(\xi(A_1), \dots \xi(A_n)\big)\Big).$$

The idea of the proof is the same of that of Proposition 5.

3.4.3 Characterization of Random Measures by Duality Formulas

Following Murr [15] we present another version of the Mecke formula to characterize random measures with independent increments by duality formulas. Indeed, Murr [15] gives a characterization of infinitely random measures so it is more general than our result. The interest of our characterization is that the proof is based on Malliavin calculus for random measures with independent increments, specifically, the duality coupling between D and δ: For $F \in \operatorname{dom} D$ and $g \in \operatorname{dom} \delta$,

$$\mathbb{E}\big[F\delta(g)\big] = \mathbb{E}\left[\int_{\mathbb{X}\times(0,\infty)} D_z F g(z)\, \mu(\mathrm{d}z)\right]. \tag{39}$$

Denote by $\mathscr{U}$ the ring of relatively compact sets of $\mathbb{X}$. Every locally finite measure is finite on the sets of $\mathscr{U}$. Let $\mathbb{S}$ be the set of functions $f : \mathbf{M}(\mathbb{X}) \to \mathbb{R}$ of the form $f(\mu) = h\big(\mu(A_1), \dots, \mu(A_n)\big)$ with $h \in \mathscr{C}_b^\infty(\mathbb{R}^n)$ and $A_1, \dots, A_n \in \mathscr{U}$; also let $\mathscr{E}$ be the set of simple functions $g = \sum_{j=1}^k a_j \mathbb{1}_{A_j}$, with $a_1, \dots, a_n > 0$ and $A_1, \dots, A_n \in \mathscr{U}$.

Theorem 6 (Murr) *Let $\beta \in \mathbf{M}(\mathbb{X})$ be non-atomic and $\nu \in \mathbf{M}\big(\mathbb{X} \times (0,\infty)\big)$ be non-atomic and such that for $A \in \mathscr{U}$, $\int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(x,s)) < \infty$. A random measure ξ has independent increments with characteristics β and ν if and only if for all $f \in \mathbb{S}$ and $g \in \mathscr{E}$,*

$$\begin{aligned}\mathbb{E}\left[f(\xi)\int_{\mathbb{X}} g(s)\, \xi(\mathrm{d}s)\right] &= \mathbb{E}\big[f(\xi)\big]\int_{\mathbb{X}} g(s)\, \beta(\mathrm{d}s)\\ &\quad + \int_{\mathbb{X}\times(0,\infty)} \mathbb{E}\big[f(\xi + x\delta_s)\big] g(s) x\, \nu(\mathrm{d}(s,x)).\end{aligned} \tag{40}$$

Proof Assume that ξ is a random measure with independent increments. Formula (40) is the particular case of formula (34) for $h(\mu, s) = f(\mu)g(s)$. However, as we commented, we will see that (40) is also consequence of the duality coupling (39).

To prove (40), by linearity, it suffices to consider the case $g = \mathbb{1}_A$ for $A \in \mathscr{U}$. By construction (see Remark 4) $\int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(s,x)) < \infty$. Assume first that also

$$\int_{A\times(0,\infty)} x^2\, \nu(\mathrm{d}(s,x)) < \infty.$$

Then $x\mathbb{1}_{A\times(0,\infty)} \in L^1\big(\nu\big) \cap L^2\big(\nu\big)$, and by the representation (30),

$$\begin{aligned}\delta(g) &= \int_{A\times(0,\infty)} x\, \hat{\eta}(\mathrm{d}(s,x)) = \int_{A\times(0,\infty)} x\, \eta(\mathrm{d}(s,x)) - \int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(s,x))\\ &= \int_{\mathbb{X}} g(s)\, \xi(\mathrm{d}s) - \int_{\mathbb{X}} g(s)\, \mathrm{d}\beta(s) - \int_{A\times(0,\infty)} x\, \nu(\mathrm{d}(s,x)).\end{aligned}$$

Further, the right-hand side of formula of duality (39) for $F = f(\xi)$ and $g = \mathbb{1}_A$ is

$$\mathbb{E}\left[\int_{A\times(0,\infty)} \big(f(\xi + x\delta_s) - f(\xi)\big)\, x\, \nu(\mathrm{d}(s,x))\right],$$

and formula (40) follows. When $\int_{A\times(0,\infty)} x^2\, \nu(\mathrm{d}(s,x)) = \infty$, then the result is obtaining approximating $\int_{A\times(0,\infty)} x\, \eta(\mathrm{d}(s,x))$ by $\int_{A\times\{0<x<n\}} x\, \eta(\mathrm{d}(s,x))$ as in the proof of Theorem 2.

The reciprocal implication is also proved as in Theorem 2. □

Remark 5 For an infinitely divisible random measure Murr [15] writes formula (40) as

$$\mathbb{E}\left[f(\xi)\int_{\mathbb{X}} g(s)\, \xi(\mathrm{d}s)\right]$$
$$= \mathbb{E}\big[f(\xi)\big]\int_{\mathbb{X}} g(s)\, \beta(\mathrm{d}s) + \mathbb{E}\left[\int_{\mathbf{M}_0(\mathbb{X})} f(\xi+\chi)\Big(\int_{\mathbb{X}} g(s)\, \chi(\mathrm{d}s)\Big)\, \Gamma(\mathrm{d}\chi)\right], \tag{41}$$

where $\mathbf{M}_0(\mathbb{X}) = \mathbf{M}(\mathbb{X})\backslash\{0\}$, here 0 is the zero measure, and Γ is a σ-finite measure on $\mathbf{M}_0(\mathbb{X})$. Kallenberg [6, Lemma 7.3] proves that ξ has independent increments if and only if Γ is concentrated on the set of degenerate measures in $\mathbf{M}_0(\mathbb{X})$, that are the measures of the form $\chi = x\,\delta_s$, for some $x > 0$ and $s \in \mathbb{X}$. In this case, consider the (measurable) mapping

$$\mathbf{M}_0(\mathbb{X}) \to \mathbb{X}\times(0,\infty)$$
$$x\,\delta_s \mapsto (s,x)$$

and then ν is the image measure of Γ by this mapping. Thus, by the image measure Theorem,

$$\int_{\mathbf{M}_0(\mathbb{X})} f(\xi+\chi)\left(\int_{\mathbb{X}} g(s)\, \chi(\mathrm{d}s)\right)\, \Gamma(\mathrm{d}\chi) = \int_{\mathbb{X}\times(0,\infty)} f(\xi + x\delta_s) x g(s)\, \nu(\mathrm{d}(s,x)),$$

so formula (41) and (40) are the same in case of a random measure with independent increments.

Acknowledgements The authors were partially supported by grants MINECO reference MTM2012-33937 and UNAB10-4E-378 co-funded by FEDER "A way to build Europe," and Generalitat de Catalunya reference 2014-SGR422.

Appendix 1: Pathwise and $L^2(\mathbb{P})$ Integrals with Respect to Poisson Random Measures

For reader convenience we review the definition and properties of the integrals with respect to Poisson random measures. For these properties there is no need of topological conditions on the state space.

Pathwise Integrals with Respect to a Poisson Random Measure

Let $(\mathbb{X}, \mathcal{X}, \nu)$ be a σ-finite measure space and η a Poisson random measure with intensity ν. For a measurable mapping $f : \mathbb{X} \to \mathbb{R}$ we can consider the integral $\int_{\mathbb{X}} f(x)\, \eta(\omega, \mathrm{d}x)$ assuming that f is positive or $\int_{\mathbb{X}} |f(x)|\, \eta(\omega, \mathrm{d}x) < \infty$, and if this happens for all $\omega \in \Omega$ a.s., the mapping $\omega \mapsto \int_{\mathbb{X}} f(x)\, \eta(\omega, \mathrm{d}x)$ defines a random variable. The following theorem summarizes the main properties of that integral. See also Privault [21, Sect. 2.4.1] for additional properties.

Theorem 7 (Kyprianou [10, Theorem 2.7]) *Let $f : \mathbb{X} \to \mathbb{R}$ be a measurable mapping. Then*

1. *The integral $\int_{\mathbb{X}} f(x)\, \eta(\omega, \mathrm{d}x)$ is absolutely convergent for every $\omega \in \Omega$ a.s. if and only if*

$$\int_{\mathbb{X}} \big(1 \wedge |f(x)|\big)\, \nu(\mathrm{d}x) < \infty. \tag{42}$$

   In this case, the characteristic function of $\int_{\mathbb{X}} f\, \mathrm{d}\eta$ is

$$\mathbb{E}\left[\exp\left\{iu \int_{\mathbb{X}} f\, \mathrm{d}\eta\right\}\right] = \exp\left\{\int_{\mathbb{X}} \left(e^{iuf(x)} - 1\right) \nu(\mathrm{d}x)\right\}.$$

2. *If $f \in L^1(\nu)$, then $\int_{\mathbb{X}} f\, \mathrm{d}\eta \in L^1(\mathbb{P})$ and*

$$\mathbb{E}\left[\int_{\mathbb{X}} f\, \mathrm{d}\eta\right] = \int_{\mathbb{X}} f\, \mathrm{d}\nu.$$

3. If $f \in L^1(\nu) \cap L^2(\nu)$*, then* $\int_{\mathbb{X}} f \, d\eta \in L^2(\mathbb{P})$ *and*

$$\mathbb{E}\left[\left(\int_{\mathbb{X}} f \, d\eta\right)^2\right] = \int_{\mathbb{X}} f^2 \, d\nu + \left(\int_{\mathbb{X}} f \, d\nu\right)^2. \tag{43}$$

Note that $f \in L^1(\nu)$ implies (42) because $1 \wedge |f| \leq |f|$.
We need the following property.

Proposition 10 *Assume that* $\nu(\mathbb{X}) < \infty$*. Let* $\{f_n,\, n \geq 1\}$ *and* f *be measurable functions on* $\mathbb{X}$ *such that* $\lim_n f_n = f$*. Then* $\lim_n \int_{\mathbb{X}} f_n \, d\eta = \int_{\mathbb{X}} f \, d\eta$ *in probability.*

Proof Observe that since $\nu(\mathbb{X}) < \infty$ all the integrals are well defined. Set $g_n = |f_n - f|$. The characteristic function of $\int_{\mathbb{X}} g_n \, d\eta$ is

$$\varphi_n(u) = \exp\left\{\int_{\mathbb{X}} \left(e^{iu\, g_n(x)} - 1\right) \nu(dx)\right\},$$

that converges to 1 by dominated convergence. Hence $\int_{\mathbb{X}} g_n \, d\eta$ converges to 0 in law, and thus in probability. □

$L^2(\mathbb{P})$*-Integral with Respect to the Compensed Poisson Random Measure*

Again with the preceding notations, consider the ring $\mathcal{X}_0 = \{C \in \mathcal{X} : \nu(C) < \infty\}$. The compensated Poisson measure is defined on $\mathcal{X}_0$ by $\hat{\eta}(C) = \eta(C) - \nu(C),\ C \in \mathcal{X}_0$. Recall that the simple functions of the form

$$f = \sum_{i=1}^{n} c_i \mathbb{1}_{C_i}, \text{ with } C_1, \ldots, C_n \in \mathcal{X}_0,$$

are dense in $L^p(\nu)$ $(p \geq 1)$. Denote by $\mathscr{D}$ the set of such functions. For $f \in \mathscr{D}$ define

$$\int_{\mathbb{X}} f \, d\hat{\eta} = \sum_{i=1}^{n} c_i\big(\eta(C_i) - \nu(C_i)\big).$$

It is clear that $\int_{\mathbb{X}} f \, d\hat{\eta} \in L^2(\mathbb{P})$ is centered, and for $f, g \in \mathscr{D}$,

$$\mathbb{E}\left[\int_{\mathbb{X}} f \, d\hat{\eta} \int_{\mathbb{X}} g \, d\hat{\eta}\right] = \int_{\mathbb{X}} fg \, d\nu. \tag{44}$$

Now, for a general $f \in L^2(\nu)$ the definition of $\int_{\mathbb{X}} f \, d\hat{\eta}$ follows by the standard procedure, and equality (44) is true for $f, g \in L^2(\nu)$. The characteristic function of $\int_{\mathbb{X}} f \, d\hat{\eta}$ is

$$\mathbb{E}\left[\exp\{iu \int_{\mathbb{X}} f \, d\hat{\eta}\}\right] = \exp\left\{\int_{\mathbb{X}} \left(e^{iuf(x)} - 1 - iuf(x)\right) \nu(dx)\right\}. \tag{45}$$

Relation Between Pathwise and $L^2(\mathbb{P})$ Integrals, and Definition of the $L^1(\mathbb{P})$ Integral

If $f \in L^1(\nu) \cap L^2(\nu)$, both integrals of f with respect to η and $\hat{\eta}$ are defined and we have

$$\int_{\mathbb{X}} f \, d\hat{\eta} = \int_{\mathbb{X}} f \, d\eta - \int_{\mathbb{X}} f \, d\nu, \text{ a.s.}$$

This is proved in a standard way.

Even if we only have $f \in L^1(\nu)$, both integrals on the right-hand side above are well defined, and then, abusing of the language, we also write $\int_{\mathbb{X}} f \, d\hat{\eta}$ to denote that difference of integrals. As a consequence of Theorem 7, that integral belongs to $L^1(\mathbb{P})$.

Appendix 2: Completely Random Measures

We recall the notion of completely random measures (in the sense of vector measures) and multiple integrals following Peccati and Taqqu [20]; for the properties presented here there are no topological conditions on the phase space. We restrict ourselves to the $L^2(\mathbb{P})$-valued completely random measures.

Let $(\mathbb{X}, \mathcal{X}, \lambda)$ be a measure space where λ is σ-finite and non-atomic. As before, set $\mathcal{X}_0 = \{C \in \mathcal{X} : \lambda(C) < \infty\}$. A centered completely random measure in $L^2(\mathbb{P})$, for short a **completely random measure**, with control measure λ is a mapping

$\varphi : \mathcal{X}_0 \times \Omega \to \mathbb{R}$ such that

1. Fixed $C \in \mathcal{X}_0$, $\varphi(\cdot, C) : \Omega \to \mathbb{R}$ is a centered square integrable random variable. We denote this random variable by $\varphi(C)$.
2. If $C_1, \ldots, C_n \in \mathcal{X}_0$ are disjoint, the random variables $\varphi(C_1), \ldots, \varphi(C_n)$ are independent.
3. For every $C_1, C_2 \in \mathcal{X}_0$,

$$\mathbb{E}[\varphi(C_1)\varphi(C_2)] = \lambda(C_1 \cap C_2).$$

As pointed out by Peccati and Taqqu [20, p. 52], φ is additive and σ-additive on $\mathbb{X}_0$ in the sense of vector measures on $L^2(\mathbb{P})$, that means, for every finite sequence of disjoint sets $C_1, \ldots, C_n \in \mathcal{X}_0$,

$$\varphi\Big(\bigcup_{i=1}^n C_i\Big) = \sum_{i=1}^n \varphi(C_i), \text{ a.s.}$$

and the same is true for an infinite sequence of pairwise disjoints sets $\{C_n,\ n \geq 1\} \subset \mathcal{X}_0$ such that $\bigcup_{n=1}^\infty C_n \in \mathcal{X}_0$. However, we stress that in general, fixed $\omega \in \Omega$, $\varphi(\omega, \cdot)$ is not σ-additive, that means, in general a completely random measure is not a random measure in the sense used in Part 2 of this paper.

Multiple Integrals with Respect to a Completely Random Measure

Itô construction of multiple integrals [5] can be extended to the case that the integrator is a general completely random measure; see Peccati and Taqqu [20, Chap. 5], and note the comment on page 83 when φ is an $L^2(\mathbb{P})$ completely random measure.

The multiple stochastic integral of order n with respect to φ, $I_n(f)$, is defined through the same steps as in the Wiener case: For

$$f = \mathbb{1}_{C_1 \times \cdots \times C_n},$$

where $C_1, \ldots, C_n \in \mathcal{X}_0$, pairwise disjoints, set

$$I_n(f) = \varphi(C_1) \cdots \varphi(C_n).$$

Therefore, I_n is extended to $L^2(\lambda^{\otimes n})$ by linearity and continuity. This integral has the usual properties:

1. $I_n(f) = I_n(\tilde{f})$, where $\tilde{f}$ is the symmetrization of f:

$$\tilde{f}(s_1, \ldots, s_n) = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} f(s_{\pi(1)}, \ldots, s_{\pi(n)}),$$

   where $\mathfrak{S}_n$ is the set of permutations of $\{1, 2, \ldots, n\}$.
2. $I_n(af + bg) = aI_n(f) + bI_n(g)$.
3. $\mathbb{E}[I_n(f)I_m(g)] = \delta_{n,m} n! \int_{\mathbb{X}^n} \tilde{f}\, \tilde{g}\, \mathrm{d}\lambda^{\otimes n}$, where $\delta_{n,m} = 1$, if $n = m$, and 0 otherwise.

Appendix 3: Canonical Space of the Jumps Part of a Process with Independent Increments

As in the Lévy processes case, we use a nice construction by Neveu [16] to build a Poisson random measure on $(0, \infty) \times \mathbb{R}_0$ with intensity measure ν defined in (1). It is worth remarking that this measure is locally finite on $(0, \infty) \times \mathbb{R}_0$. We separate the construction in two steps:

Step 1 For $m \geq 0$ and $k \geq 1$, set

$$\Delta_{m,1} = (m, m+1] \times \{x \in \mathbb{R} : 1 < |x|\},$$
$$\Delta_{m,k} = (m, m+1] \times \{x \in \mathbb{R} : 1/k < |x \leq 1/(k-1)|\}, \quad k \geq 2.$$

Since for every $t > 0$, ν_t is a Lévy measure, we have that $\nu(\Delta_{m,k}) < \infty$. Denote by $\nu_{m,k}$ the restriction of ν to $\Delta_{m,k}$. We consider the space of the finite sequences of elements of $\Delta_{m,k}$, including the empty sequence; specifically, let

$$\Omega_{m,k} = \bigcup_{n \geq 0} (\Delta_{m,k})^n,$$

where $(\Delta_{m,k})^0 = \{\alpha\}$, α being a distinguished element that represents the empty sequence. Let

$$\mathscr{A}_{m,k} = \Big\{B \subset \Omega_{m,k} : B = \bigcup_{n \geq 0} B_n,\ B_n \in \mathscr{B}(\Delta_{m,k})^n\Big\}.$$

Since $\nu_{m,k}(\Delta_{m,k}) < \infty$, there is a probability $\mathbb{Q}_{m,k}$ on $\Delta_{m,k}$ such that $\nu_{m,k} = c_{m,k}\, \mathbb{Q}_{m,k}$, for some constant $c_{m,k} > 0$. Now define a probability $\mathbb{P}_{m,k}$ on $(\Omega_{m,k}, \mathscr{A}_{m,k})$ in the following way: for $B = \bigcup_n B_n$,

$B_n \in \mathscr{B}(\Delta_{m,k})^n$ set

$$\mathbb{P}_{m,k}(B) = e^{-c_{m,k}} \sum_{n=0}^{\infty} \frac{c_{m,k}^n}{n!} \mathbb{Q}_{m,k}^{\otimes n}(B_n),$$

where $\mathbb{Q}_{m,k}^{\otimes 0} = \delta_\alpha$. Then, Neveu [16, Proposition I.6] proves that under $\mathbb{P}_{m,k}$, the mapping given by

$$N'_{m,k}(\omega) = \sum_{j=1}^{n} \delta_{(t_j,x_j)}, \text{ if } \omega = \big((t_1,x_1),\dots,(t_n,x_n)\big),$$

and $N'_{m,k}(\alpha) = 0$, is a Poisson random measure with intensity $\nu_{m,k}$.

Step 2 Now superpose the Poisson random measures $N'_{m,k}$: Let

$$(\Omega_N, \mathscr{A}_N, \mathbb{P}_N) = \bigotimes_{m\geq 1, k\geq 1} (\Omega_{m,k}, \mathscr{A}_{m,k}, \mathbb{P}_{m,k}).$$

For $\omega = (\omega_{m,k}, m \geq 1, k \geq 1)$, define

$$N^*_{m,k}(\omega) = N'_{m,k}(\omega_{m,k})$$

and finally

$$N^*(\omega) = \sum_{m,k} N^*_{m,k}(\omega),$$

which is a Poisson random measure with intensity measure ν.

References

1. Chow, Y.S., Teicher H.: Probability Theory. Independence, Interchangeability, Martingales. Springer, New York (1978)
2. Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure, 2nd edn. Springer, New York (2008)
3. Di Nunno, G., Øksendal, B., Proske, F.: Malliavin Calculus for Lévy Processes with Applications to Finance. Springer, Berlin (2009)
4. Geiss, C., Laukkarinen, E.: Denseness of certain smooth Lévy functionals in $\mathbb{D}_{1,2}$. Probab. Math. Stat. **31**, 1–15 (2011)
5. Itô, K.: Spectral type of the shift transformation of differential processes with stationary increments. Trans. Am. Math. Soc. **81**, 253–263 (1956)
6. Kallenberg, O.: Random Measures, 3rd edn. Akademie-Verlag, Berlin (1983)
7. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002)
8. Kingman, J.F.C.: Completely random measures. Pac. J. Math. **21**, 59–78 (1967)

9. Kingman, J.F.C.: Poisson Processes. Oxford Science Publications/Clarendon Press, Oxford (1993)
10. Kyprianou, A.E.: Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer, Berlin (2006)
11. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
12. Last, G., Penrose, M.D.: Poisson process Fock space representation, chaos expansion and covariance inequalities. Probab. Theory Relat. Fields **150**, 663–690 (2011)
13. Matthes, K., Kerstan, J., Mecke, J.: Infinite Divisible Point Processes. Wiley, Chichester (1978)
14. Molchanov, I., Zuyev, S.: Variational analysis of Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 81–101. Springer, Cham (2016)
15. Murr, R.: Characterization of infinite divisibility by duality formulas. Application to Lévy processes and random measures. Stoch. Process. Appl. **123**, 1729–1749 (2013)
16. Neveu, J.: Processus Pontuels. In: École d'Eté de Probabilités de Saint Flour, VI-1976. Lecture Notes in Mathematics, vol. 598, pp. 247–447. Springer, Berlin (1977)
17. Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Springer, Berlin (2006)
18. Nualart, D., Vives, J.: Anticipative calculus for the Poisson process based on the Fock space. In: Séminaire de Probabilités XXIV. Lecture Notes in Mathematics, vol. 1426, pp. 154–165. Springer, Berlin (1990)
19. Nualart, D., Vives, J.: A duality formula on the Poisson space and some applications. In: Proceedings of the Ascona Conference on Stochastic Analysis. Progress in Probability, vol. 36, pp. 205–213. Birkhäuser, Basel (1995)
20. Peccati, G., Taqqu M.S.: Wiener Chaos: Moments, Cumulants and Diagrams. Bocconi & Springer Series, vol. 1, Springer, Milan (2011)
21. Privault, N.: Combinatorics of Poisson stochastic integrals with random integrals. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 37–80. Springer, Cham (2016)
22. Sato, K.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (1999)
23. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Springer, Berlin (2008)
24. Solé, J.L., Utzet, F., Vives, J.: Canonical Lévy process and Malliavin calculus. Stoch. Process. Appl. **117**, 165–187 (2007)
25. Solé, J.L., Utzet, F., Vives, J.: Chaos expansions and Malliavin calculus for Lévy processes. In: Abel Symposium 2005, Stochastic Analysis and Applications, pp. 595–612. Springer, Berlin (2007)

Introduction to Stochastic Geometry

Daniel Hug and Matthias Reitzner

Abstract This chapter introduces some of the fundamental notions from stochastic geometry. Background information from convex geometry is provided as far as this is required for the applications to stochastic geometry.

First, the necessary definitions and concepts related to geometric point processes and from convex geometry are provided. These include Grassmann spaces and invariant measures, Hausdorff distance, parallel sets and intrinsic volumes, mixed volumes, area measures, geometric inequalities and their stability improvements. All these notions and related results will be used repeatedly in the present and in the subsequent chapters of the book.

Second, a variety of important models and problems from stochastic geometry will be reviewed. Among these are the Boolean model, random geometric graphs, intersection processes of (Poisson) processes of affine subspaces, random mosaics, and random polytopes. We state the most natural problems and point out important new results and directions of current research.

1 Introduction

Stochastic geometry is a branch of probability theory which deals with set-valued random elements. It describes the behavior of random configurations such as random graphs, random networks, random cluster processes, random unions of convex sets, random mosaics, and many other random geometric structures. Due to its strong connections to the classical field of stereology, to communication theory, and to spatial statistics it has a large number of important applications.

The connection between probability theory and geometry can be traced back at least to the middle of the eighteenth century when Buffon's needle problem (1733), and subsequently questions related to Sylvester's four point problem (1864)

D. Hug
Department of Mathematics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany
e-mail: daniel.hug@kit.edu

M. Reitzner (✉)
Institut für Mathematik, Universität Osnabrück, Albrechtstraße 28a, 49086 Osnabrück, Germany
e-mail: matthias.reitzner@uni-osnabrueck.de

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*,
Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_5

and Bertrand's paradox (1889) started to challenge prominent mathematicians and helped to advance probabilistic modeling. Typically, in these early contributions a fixed number of random objects of a fixed shape was considered and their interaction was studied when some of the objects were moved randomly. For a short historical outline of these early days of *Geometric Probability* see [104, Chap. 8] and [105, Chap. 1].

Since the 1950s, the framework broadened substantially. In particular, the focus mainly switched to models involving a random number of randomly chosen geometric objects. As a consequence, the notion of a point process started to play a prominent role in this field, which since then was called *Stochastic Geometry*.

In this chapter we describe some of the classical problems of stochastic geometry, together with their recent developments and some interesting open questions. For a more thorough treatment we refer to the seminal book on "Stochastic and Integral Geometry" by Schneider and Weil [104].

2 Geometric Point Processes

A point process η is a measurable map from some probability space $(\Omega, \mathscr{A}, \mathbb{P})$ to the locally finite subsets of a Polish space $\mathbb{X}$ (endowed with a suitable σ-algebra), which is the state space. The intensity measure of η, evaluated at a measurable set $A \subset \mathbb{X}$, is defined by $\mu(A) = \mathbb{E}\eta(A)$ and equals the mean number of elements of η lying in A. 32

In many examples considered in this chapter, $\mathbb{X}$ is either $\mathbb{R}^d$, the space of compact (convex) subsets of $\mathbb{R}^d$, or the space of flats (affine subspaces) of a certain dimension in $\mathbb{R}^d$. More generally, $\mathbb{X}$ could be the family $\mathcal{F}(\mathbb{R}^d)$ of all closed subsets of $\mathbb{R}^d$ endowed with the hit-and-miss topology (which yields a compact Hausdorff space with countable basis).

In this section, we start with processes of flats. In the next section, we discuss particle processes in connection with Boolean models.

2.1 Grassmannians and Invariant Measures

Let $\mathbb{X}$ be the space of linear or affine subspaces (flats) of a certain dimension in $\mathbb{R}^d$. More specifically, for $i \in \{0, \dots, d\}$ we consider the linear Grassmannian

$$G(d, i) = \{L \text{ linear subspace of } \mathbb{R}^d : \ \dim L = i\}$$

and the affine Grassmannian

$$A(d, i) = \{E \text{ affine subspace of } \mathbb{R}^d : \ \dim E = i\}.$$

These spaces can be endowed with a canonical topology and with a metric inducing this topology. In both cases, we work with the corresponding Borel σ-algebra. Other examples of spaces $\mathbb{X}$ are the space of compact subsets or the space of compact convex subsets of $\mathbb{R}^d$. All these spaces are subspaces of $\mathcal{F}(\mathbb{R}^d)$ and are endowed with the subspace topology.

In each of these examples, translations and rotations act in a natural way on the elements of $\mathbb{X}$ as well as on subsets (point configurations) of $\mathbb{X}$. It is well known and an often used fact that there is—up to normalization—only one translation invariant and locally finite measure on $\mathbb{R}^d$, the Lebesgue measure $\ell_d(\cdot)$. It is also rotation invariant and normalized in such a way that the unit cube $C^d = [0,1]^d$ satisfies $\ell_d(C^d) = 1$.

Analogously, there is only one rotation invariant probability measure on $G(d,i)$, which we denote by ν_i^d and which by definition satisfies $\nu_i^d(G(d,i)) = 1$. Observe that ν_{d-1}^d coincides (up to normalization) with (spherical) Lebesgue measure σ^d on the unit sphere S^{d-1}, by identifying a unit vector $u \in S^{d-1}$ with its orthogonal complement $u^\perp = L \in G(d, d-1)$. A corresponding remark applies to ν_1^d on $G(d,1)$ where a unit vector is identified with the one-dimensional linear subspace it spans.

In a similar way, there is—up to normalization—only one rotation and translation invariant measure on $A(d,i)$, the Haar measure μ_i^d, which is normalized in such a way that $\mu_i^d(\{E \in A(d,i) : E \cap B^d \neq \emptyset\}) = \kappa_{d-i}$, where B^d is the unit ball in $\mathbb{R}^d$ and κ_d denotes its volume. Since the space $A(d,i)$ is not compact, its total μ_i^d-measure is infinite.

It is often convenient to describe the Haar measure μ_i^d on $A(d,i)$ in terms of the Haar measure ν_i^d on $G(d,i)$. The relation is

$$\mu_i^d(A) = \int_{G(d,i)} \int_{L^\perp} \mathbb{1}_A(L+x)\, \ell_{d-i}(\mathrm{d}x)\, \nu_i^d(\mathrm{d}L), \tag{1}$$

for measurable sets $A \subset A(d,i)$. This is based on the obvious fact that each i-flat $E \in A(d,i)$ can be uniquely written in the form $E = L + x$ with $L \in G(d,i)$ and $x \in L^\perp$, the orthogonal complement of L. If a locally finite measure μ on $A(d,i)$ is only translation invariant, then it can still be decomposed into a probability measure σ on $G(d,i)$ and, given a direction space $L \in G(d,i)$, a translation invariant measure on the orthogonal complement of L, which then coincides up to a constant with Lebesgue measure on $L^\perp$. In fact, a more careful argument shows the existence of a constant $t \geq 0$ such that

$$\mu(A) = t \int_{G(d,i)} \int_{L^\perp} \mathbb{1}_A(L+x)\, \ell_{d-i}(\mathrm{d}x)\, \sigma(\mathrm{d}L),$$

for all measurable sets $A \subset A(d,i)$. In this situation, $\sigma = \nu_i^d$ if and only if μ is also rotation invariant and therefore $\mu = \mu_i^d$, at least up to a constant factor.

The Haar measures ℓ_d, ν_i^d, and μ_i^d are the basis of the most natural constructions of point processes on $\mathbb{X} = \mathbb{R}^d, G(d,i)$ and $A(d,i)$, if some kind of invariance is involved.

2.2 Stationary Point Processes

Next we describe point processes on these spaces in a slightly more formal way than at the beginning of this section and refer to [71] for a general detailed introduction. A point process (resp. simple point process) η on $\mathbb{X}$ is a measurable map from the underlying probability space $(\Omega, \mathscr{A}, \mathbb{P})$ to the set of locally finite (resp. locally finite and simple) counting measures $\mathbf{N}(\mathbb{X})$ (resp., $\mathbf{N}_s(\mathbb{X})$) on $\mathbb{X}$, which is endowed with the smallest σ-algebra, so that the evaluation maps $\omega \mapsto \eta(\omega)(A)$ are measurable, for all Borel sets $A \subset \mathbb{X}$. For $z \in \mathbb{X}$, let δ_z denote the unit point measure at z. It can be shown that a point process can be written in form

$$\eta = \sum_{i=1}^{\tau} \delta_{\zeta_i},$$

where τ is a random variable taking values in $\mathbb{N}_0 \cup \{\infty\}$ and $\zeta_1, \zeta_2, \dots$ is a sequence of random points in $\mathbb{X}$. In the following, we will only consider simple point processes, where $\zeta_i \neq \zeta_j$ for $i \neq j$. If η is simple and identifying a simple measure with its support, we can think of η as a locally finite random set $\eta = \{\zeta_i : i = 1, \dots, \tau\}$.

Taking the expectation of η yields the intensity measure

$$\mu(A) = \mathbb{E}\eta(A)$$

of η. As indicated above, the most convenient point processes from a geometric point of view are those where the intensity measure equals the Haar measure, or at least a translation invariant measure, times a constant $t > 0$, the intensity of the point process. If we refer to this setting, we write η_t and μ_t to emphasize the dependence on the intensity t. In the following, we make this precise under the general assumption that the intensity measure is locally finite. As usual we say that a point process η is stationary if any translate of η by a fixed vector has the same distribution as the process η.

Let us discuss the consequences of the assumptions of stationarity or some additional distributional invariance in some particular cases. If η is a stationary point process on $\mathbb{X} = \mathbb{R}^d$, then $\mu_t(A) = t\ell_d(A)$ for all Borel sets $A \subset \mathbb{R}^d$. Clearly, this measure is also rotation invariant.

Furthermore, if η is a stationary flat process on $\mathbb{X} = A(d,i)$ and $A \subset \mathbb{R}^d$ is a Borel set, we set $[A] = \{E \in A(d,i) : E \cap A \neq \emptyset\}$. Then the number of i-flats of the process meeting A is given by $\eta([A])$ and its expectation can be written as

$$\mu_t([A]) = t \int_{G(d,i)} \int_{L^\perp} \mathbb{1}_{[A]}(L + x)\, \ell_{d-i}(\mathrm{d}x)\, \sigma(\mathrm{d}L),$$

where σ is a probability measure on $G(d, i)$ and $t \geq 0$ is the intensity. This follows from what we said in the previous subsection, since the intensity measure is translation invariant by the assumption of stationarity of η. Here, the indicator function $\mathbb{1}_{[A]}(L + x)$ equals 1 if and only if x is in the orthogonal projection $A|L^\perp$ of A to $L^\perp$. Thus

$$\mu_t([A]) = t \int_{G(d,i)} \ell_{d-i}(A|L^\perp)\, \sigma(\mathrm{d}L).$$

A special situation arises if η is also isotropic (its distribution is rotation invariant). In this case and for a convex set A, the preceding formula can be expressed as an intrinsic volume, which will be introduced in the next section.

2.3 *Tools from Convex Geometry*

We work in the d-dimensional Euclidean space $\mathbb{R}^d$ with Euclidean norm $\|x\| = \sqrt{\langle x, x \rangle}$, unit ball B^d and unit sphere S^{d-1}. The set of all convex bodies, i.e., compact convex sets in $\mathbb{R}^d$, is denoted by $\mathcal{K}^d$. The Hausdorff distance between two sets A, B is defined as $d_H(A, B) = \inf\{\varepsilon \geq 0 : A \subset B + \varepsilon B^d \text{ and } B \subset A + \varepsilon B^d\}$ where "+" denotes the usual vector or Minkowski addition. When equipped with the Hausdorff distance, $\mathcal{K}^d$ is a metric space. The elements of the convex ring $\mathcal{R}^d$ are the polyconvex sets, which are defined as finite unions of convex bodies.

If Lebesgue measure is applied to elements of $\mathcal{K}^d$, we usually write V_d instead of ℓ_d. Using the Minkowski addition on $\mathcal{K}^d$, we can define the surface area of a convex body by

$$\lim_{\varepsilon \to 0+} \frac{V_d(K + \varepsilon B^d) - V_d(K)}{\varepsilon}.$$

Classical results in convex geometry imply that the limit exists. The mean width of a convex body K is the mean length of the projection $K|L$ of the set onto a uniform random line L through the origin,

$$\int_{G(d,1)} V_1(K|L)\, \nu_1^d(\mathrm{d}L).$$

These two quantities, which describe natural geometric properties of convex bodies, are just two examples of a sequence of characteristics associated with convex bodies.

2.3.1 Intrinsic Volumes

More generally, we now introduce *intrinsic volumes* V_i of convex bodies, $i = 1, \dots, d$. These can be defined through the Steiner formula which states that, for any convex body $K \in \mathcal{K}^d$, the volume of $K + \varepsilon B^d$ is a polynomial in $\varepsilon \geq 0$ of degree d. The intrinsic volumes are the suitably normalized coefficients of this polynomial, namely,

$$V_d(K + \varepsilon B^d) = \sum_{i=0}^{d} \kappa_i V_{d-i}(K) \varepsilon^i, \qquad \varepsilon \geq 0,$$

where κ_i is the volume of the i-dimensional unit ball. Clearly, the functional $2V_{d-1}$ is the surface area, V_1 is a multiple of the mean width functional, and V_0 corresponds to the Euler characteristic.

The intrinsic volumes V_i are translation and rotation invariant, homogeneous of degree i, monotone with respect to set inclusion, and continuous with respect to the Hausdorff distance. The intrinsic volumes are *additive functionals*, also called *valuations*, which means that

$$V_i(K \cup L) + V_i(K \cap L) = V_i(K) + V_i(L)$$

whenever $K, L, K \cup L \in \mathcal{K}^d$. Moreover, it is a convenient feature of the intrinsic volumes that for $K \subset \mathbb{R}^d \subset \mathbb{R}^N$ the value $V_i(K)$ is independent of the ambient space, $\mathbb{R}^d$ or $\mathbb{R}^N$, in which it is calculated. In particular, for $L \in G(d, 1)$ the intrinsic volume $V_1(K|L)$ is just the length of $K|L$.

A famous theorem due to Hadwiger (see [104, Sect. 14.4]) states that the intrinsic volumes can be characterized by these properties. If μ is a translation and rotation invariant, continuous valuation on $\mathcal{K}^d$, then

$$\mu = \sum_{i=0}^{d} c_i V_i$$

with some constants $c_0, \dots, c_d \in \mathbb{R}$ depending only on μ. If in addition μ is homogeneous of degree i, then $\mu = c_i V_i$. To give a simple example for an application of Hadwiger's theorem, observe that the mean projection volume

$$\int_{G(d,i)} \ell_{d-i}(K|L^{\perp}) \, \nu_i^d(\mathrm{d}L)$$

of a convex body K to a uniform random $(d-i)$-dimensional subspace defines a translation invariant, rotation invariant, monotone and continuous valuation of degree $d-i$. Hence, up to a constant factor (independent of K), it must be equal to $V_{d-i}(K)$. This yields Kubota's formula

$$V_{d-i}(K) = c_{d,i} \int_{G(d,i)} \ell_{d-i}(K|L^{\perp})\, \nu_i^d(\mathrm{d}L),$$

with certain constants $c_{d,i}$ which can be determined by comparing both sides for $K = B^d$. This formula explains why the intrinsic volumes are often encountered in stereological or tomographic investigations and are also called "Quermassintegrals", which is the German name for an integral average of sections or projections of a body.

Applications to stochastic geometry require an extension of intrinsic volumes to the larger class of polyconvex sets. Requiring such an extension to be additive on $\mathcal{R}^d$ suggests to define the intrinsic volumes of polyconvex sets by an inclusion–exclusion formula. The fact that this is indeed possible can be seen from a result due to Groemer [38], [104, Theorem 14.4.2], which says that any continuous valuation on $\mathcal{K}^d$ has an additive extension to $\mathcal{R}^d$. Volume and surface area essentially preserve their interpretation for the extended functionals and also Kubota's formula remains valid for all intrinsic volumes. On the other hand, continuity with respect to the Hausdorff metric is in general not available on $\mathcal{R}^d$.

2.3.2 Mixed Volumes and Area Measures

The Steiner formula can be extended in different directions. Instead of considering the volume of the Minkowski sum of a convex body and a ball, more generally, the volume of a Minkowski combination of finitely many convex bodies $K_1, \dots, K_k \in \mathcal{K}^d$ can be taken. In this case, $V_d(\lambda_1 K_1 + \ldots + \lambda_k K_k)$ is a homogeneous polynomial in $\lambda_1, \dots, \lambda_k \geq 0$ of degree d, whose coefficients are nonnegative functionals of the convex bodies involved (see [101, Chap. 5.1]), which are called *mixed volumes*. We mention only the special case $k = 2$,

$$V_d(\lambda_1 K_1 + \lambda_2 K_2) = \sum_{i=0}^{d} \binom{d}{i} \lambda_1^i \lambda_2^{d-i} V(K_1[i], K_2[d-i]);$$

the bracket notation $K[i]$ means that K enters with multiplicity i. In particular, for $K, L \in \mathcal{K}^d$ we thus get

$$d \cdot V(K[d-1], L) = \lim_{\varepsilon \to 0+} \frac{V_d(K + \varepsilon L) - V_d(K)}{\varepsilon},$$

which provides an interpretation of the special mixed volume $V(K[d-1], L)$ as a relative surface area of K with respect to L. In particular, $d \cdot V(K[d-1], B^d)$ is the surface area of K. The importance of these mixed functionals is partly due to sharp geometric inequalities satisfied by them. For instance, *Minkowski's inequality* (see [101, Chap. 7.2]) states that

$$V(K[d-1], L)^d \geq V_d(K)^{d-1} V_d(L). \tag{2}$$

If K, L are d-dimensional, then (2) holds with equality if and only if K and L are homothetic. Note that the very special case $L = B^d$ of this inequality is the classical isoperimetric inequality for convex sets.

Although Minkowski's inequality is sharp, it can be strengthened by taking into account that the left side is strictly larger than the right side if K and L are not homothetic. Quantitative improvements of (2) which introduce an additional factor $(1 + f(d(K, L))$ on the right-hand side, with a nonnegative function f and a suitable distance $d(K, L)$, are extremely useful and are known as *geometric stability results*.

A second extension is obtained by localizing the parallel sets involved in the Steiner formula. For a given convex body K, this leads to a sequence of Borel measures $S_j(K, \cdot)$, $j = 0, \ldots, d-1$, on S^{d-1}, the *area measures* of the convex body K. The top order area measure $S_{d-1}(K, \cdot)$ can be characterized via the identity

$$d \cdot V(K[d-1], L) = \int_{S^{d-1}} h(L, u)\, S_{d-1}(K, \mathrm{d}u),$$

which holds for all convex bodies $K, L \in \mathcal{K}^d$, and where

$$h(L, u) := \max\{\langle x, u\rangle : x \in L\}, \qquad u \in \mathbb{R}^d,$$

defines the *support function* of L. Moreover, for any Borel set $\omega \subset S^{d-1}$ we have

$$S_{d-1}(K, \omega) = \mathcal{H}^{d-1}(\{x \in \partial K : \langle x, u\rangle = h(K, u) \text{ for some } u \in \omega\}),$$

where $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure. Further extensions and background information are provided in [101] and summarized in [104].

3 Basic Models in Stochastic Geometry

3.1 The Boolean Model

The Boolean model, which is also called Poisson grain model [41], is a basic benchmark model in spatial stochastics. Let $\xi_t = \sum_{i=1}^{\infty} \delta_{x_i}$ denote a stationary Poisson point process in $\mathbb{R}^d$ with intensity $t > 0$. By $\mathcal{K}_0^d$ we denote the set of all convex bodies $K \in \mathcal{K}^d$ for which the origin is the center of the circumball. Let $\mathbb{Q}$ denote a probability distribution on $\mathcal{K}_0^d$, and let $Z_1, Z_2, \ldots$ be an i.i.d. sequence of random convex bodies (particles) which are also independent of ξ_t. If we assume that

$$\int_{\mathcal{K}_0^d} V_j(K)\, \mathbb{Q}(\mathrm{d}K) < \infty \tag{3}$$

for $j = 1, \ldots, d$, then

$$Z = \bigcup_{i=1}^{\infty} (Z_i + x_i)$$

is a stationary random closed set, the Boolean model with grain (or shape) distribution $\mathbb{Q}$ and intensity $t > 0$. Alternatively, one can start from a stationary point process (particle process) η_t on $\mathcal{K}^d$. Then the intensity measure $\mu_t = \mathbb{E}\eta_t$ of η_t is a translation invariant measure on $\mathcal{K}^d$ which can be decomposed in the form

$$\mu_t(\cdot) = t \int_{\mathcal{K}_0^d} \int_{\mathbb{R}^d} \mathbb{1}\{K + x \in \cdot\}\, \ell_d(\mathrm{d}x)\, \mathbb{Q}(\mathrm{d}K).$$

The Poisson particle process η_t is locally finite if and only if its intensity measure μ_t is locally finite, which is equivalent to (3). We obtain again the Boolean model by taking the union of the particles of η_t, that is,

$$Z = Z(\eta_t) = \bigcup_{K \in \eta_t} K.$$

In order to explore a Boolean model Z, which is observed in a window $W \in \mathcal{K}^d$, it is common to consider the values of suitable functionals of the intersection $Z \cap W$ as the information which is available. Due to the convenient properties and the immediate interpretation of the intrinsic volumes V_i, $i \in \{0, \ldots, d\}$, for convex bodies, it is particularly natural to study the random variables $V_i(Z \cap W)$, $i \in \{0, \ldots, d\}$, or to investigate random vectors composed of these random elements. From a practical viewpoint, one aims at retrieving information about the underlying particle process, that is, its intensity and its shape distribution, from such observations.

3.1.1 Mean Values

Let Z_0 be a random convex body having the same distribution as Z_i, $i \in \mathbb{N}$, which is called the *typical grain*. Formulas relating the mean values $\mathbb{E}V_i(Z \cap W)$ to the mean values of the typical grain $v_j = \mathbb{E}V_j(Z_0)$, $j \in \{0, \ldots, d\}$, have been studied for a long time. Particular examples of such relations are

$$\mathbb{E}V_d(Z \cap W) = V_d(W)\left(1 - e^{-tv_d}\right),$$
$$\mathbb{E}V_{d-1}(Z \cap W) = V_d(W)tv_{d-1}e^{-tv_d} + V_{d-1}(W)\left(1 - e^{-tv_d}\right).$$

If $r(W)$ denotes the radius of the inball of W, we deduce from these relations that

$$\lim_{r(W)\to\infty} \frac{\mathbb{E}V_d(Z \cap W)}{V_d(W)} = 1 - e^{-tv_d},$$
$$\lim_{r(W)\to\infty} \frac{\mathbb{E}V_{d-1}(Z \cap W)}{V_d(W)} = tv_{d-1}e^{-tv_d},$$

where the first limit is redundant and equal to $p = \mathbb{P}(o \in Z) = \mathbb{E}V_d(Z \cap W)/V_d(W)$, the volume fraction of the stationary random closed set Z. For the other intrinsic volumes V_i, $i \in \{0, \dots, d-2\}$, the mean values $\mathbb{E}V_i(Z \cap W)$ of the Boolean model Z can still be expressed in terms of the intensity and mean values of the typical grain, but the relations are more complicated and in general they involve mixed functionals of translative integral geometry. The formulas simplify again if Z is additionally assumed to be isotropic (if Z_0 is isotropic). For a stationary and isotropic Boolean model, all mean values $\mathbb{E}V_i(Z \cap W)$ can be expressed in terms of the volume fraction p and a polynomial function of $tv_i, \dots, tv_d$. Moreover, the limits

$$\delta_i := \lim_{r(W)\to\infty} \frac{\mathbb{E}V_i(Z \cap W)}{V_d(W)}$$

exist and are called the densities of the intrinsic volumes for the Boolean model. The system of equations which relates these densities to the (intensity weighted) mean values $tv_0, \dots, tv_d$ can be used to express the latter in terms of the densities $\delta_0, \dots, \delta_d$ of the Boolean model.

3.1.2 Covariances

While such first order results (involving mean values) have been studied for quite some time (see [104] for a detailed description), variances and covariances of arbitrary intrinsic volumes (or of more general functionals) of Boolean models have been out of reach until recently. In [58], second order information for functionals of the Boolean model is derived systematically under optimal moment assumptions. To indicate some of these results, we define for $i, j \in \{0, \dots, d\}$

$$\sigma_{i,j} = \lim_{r(W)\to\infty} \frac{\operatorname{Cov}\left(V_i(Z \cap W), V_j(Z \cap W)\right)}{V_d(W)} \tag{4}$$

as the asymptotic covariances of the stationary Boolean model Z, provided the limit exists. The following results are proved in [58] and ensure the existence of the limit under minimal assumptions. Note that condition (3) is equivalent to $\mathbb{E}V_i(Z_0) < \infty$ for $i = 1, \dots, d$.

Theorem 1 *Assume that* $\mathbb{E}V_i(Z_0)^2 < \infty$ *for* $i \in \{1, \dots, d\}$.

(1) *Then* $\sigma_{i,j}$ *is finite and independent of* W *for all* $i, j \in \{0, \dots, d\}$. *Moreover,* $\sigma_{i,j}$ *can be expressed as an infinite series involving the intensity* t *and integrations with respect to the grain distribution* $\mathbb{Q}$ *and the intensity measure* μ *of* η_t.
(2) *The asymptotic covariance matrix is positive definite if* Z_0 *has nonempty interior with positive probability.*
(3) *If even* $\mathbb{E}V_i(Z_0)^3 < \infty$ *for* $i \in \{0, \dots, d\}$, *then the rate of convergence in* (4) *is of the (optimal) order* $1/r(W)$.

A more general result is obtained in [58], which applies to arbitrary translation invariant, additive functionals which are locally bounded and measurable (geometric functionals). Further examples of such functionals are mixed volumes and certain integrals of area measures. The basic ingredients in the proof are the Fock space representation of Poisson functionals as developed in [73] (see also the contribution by Günter Last in this volume) and new integral geometric bounds for geometric functionals.

For an isotropic Boolean model, the infinite series representation for $\sigma_{i,j}$ can be reduced to an integration with respect to finitely many curvature based moment measures of the typical grain Z_0. As a basic example, which does not require Z to be isotropic, we mention (assuming a full-dimensional typical grain Z_0) that

$$\begin{aligned}\sigma_{d-1,d} = &-e^{-2tv_d} t v_{d-1} \int \left(e^{tC_d(x)} - 1\right) \ell_d(\mathrm{d}x) \\ &+ e^{-2tv_d} t \int e^{tC_d(x-y)} M_{d-1,d}(\mathrm{d}(x, y)),\end{aligned}$$

where $C_d(x) = \mathbb{E}[V_d(Z_0 \cap (Z_0 + x))]$, for $x \in \mathbb{R}^d$, defines the mean covariogram of the typical grain and

$$M_{d-1,d}(\cdot) := \frac{1}{2}\mathbb{E} \int_{Z_0} \int_{\partial Z_0} \mathbb{1}\{(x, y) \in \cdot\} \mathcal{H}^{d-1}(\mathrm{d}x)\, \ell_d(\mathrm{d}y)$$

is a mixed moment measure of the typical grain. A formula for the asymptotic covariance $\sigma_{d-1,d-1}$ is already contained in [42]. For a stationary and isotropic Boolean model in the plane $\mathbb{R}^2$, explicit formulas are provided in [58] for all covariances involving the Euler characteristic $\sigma_{0,0}, \sigma_{0,1}, \sigma_{0,2}$. Moreover, again in general dimensions and for a stationary Boolean model whose typical grain is a deterministic ball, some of these formulas can be specified even further and used to plot the covariances as a function of the intensity. It is an interesting task to interpret these plots and to determine rigorously the analytic properties (e.g., zeros, extremal values) or the asymptotic behavior of the covariances and correlation functions for increasing intensity.

In addition, in [58] univariate and multivariate central limit theorems, including rates of convergence, are derived from general new results on the normal approximation of Poisson functionals via the Malliavin–Stein method [81, 82]. For these we refer to the survey [17], in this volume. Again these results are established for quite general geometric functionals, employing also tools from integral geometry. Some of these results do not require stationarity of the Boolean model or translation invariance of the functionals.

3.2 Random Geometric Graphs

Random graphs play an important role in graph theory since Renyi introduced his famous random graph model. Since then several models of random graphs have been investigated. The use of random graphs as a natural model for telecommunication networks (see, e.g., Zuyev's survey in [115]) gave rise to additional investigations. Here we concentrate on random graphs with a geometric construction rule.

The most natural and best investigated graph is the so-called *Gilbert graph*. Let η_t be a Poisson point process on $\mathbb{R}^d$ with an intensity measure of the form $\mu_t(\cdot) = t\ell_d(\cdot \cap W)$, where $W \subset \mathbb{R}^d$ is a compact convex set with $\ell_d(W) = 1$. Let $(\delta_t : t > 0)$ be a sequence of positive real numbers such that $\delta_t \to 0$ as $t \to \infty$. The Gilbert graph, or random geometric graph, is obtained by taking the points of η_t as vertices and by connecting two distinct points $x, y \in \eta_t$ by an edge if and only if $\|x-y\| \le \delta_t$. There is a vast literature on the Gilbert graph and one should have a look at the seminal book [83] by Penrose or check the recent paper by Reitzner et al. [93] for further references. For natural generalizations one replaces the role of the norm by a suitable symmetric function $G : \mathbb{R}^d \to [0,1]$, where two points of η_t are connected with probability $G(y-x)$. An important particular case is when G is the indicator function of a symmetric set. Recent developments in this direction are due to Bourguin and Peccati [16], and Lachièze-Rey and Peccati [66, 67].

Denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ the resulting graph where $\mathcal{V} = \eta_t$ are the vertices and $\mathcal{E} \subset \eta^2_{t,\neq}$ are the occurring edges. Objects of interest are clearly the number of edges N_t and, more general, functions of the edge lengths

$$\sum_{(x,y)\in\mathcal{E}} g(\|y-x\|).$$

In particular, one is interested in the edge length powers

$$L_t^{(\alpha)} = \frac{1}{2} \sum_{(x,y)\in\eta^2_{t,\neq}} \mathbb{1}\{\|x-y\| \le \delta_t\}\, \|x-y\|^{\alpha}.$$

Clearly $L_t^{(0)} = N_t$. It is well known that for any $\alpha > -d$

$$\mathbb{E}L_t^{(\alpha)} = \frac{d\kappa_d}{2(\alpha + d)} t^2 \delta_t^{\alpha+d} V_d(W)(1 + O(\delta_t)) .$$

This especially shows that the number of edges of the Gilbert graph is of order $t^2\delta_t^d$, whereas its total edge length is of order $t^2\delta_t^{d+1}$. The asymptotic variance is given by

$$\mathrm{Var}L_t^{(\alpha)} = \left(\frac{d\,\kappa_d}{2\,(2\alpha + d)}\, t^2\, \delta_t^{2\alpha+d} + \frac{d^2\,\kappa_d^2}{(\alpha + d)^2}\, t^3\, \delta_t^{2\alpha+2d}\right) V_d(W)(1 + O(\delta_t)),$$

and the asymptotic covariance matrix is computed in [93].

Many investigations benefit from the fact that these functions are Poisson U-statistics of order 2, and thus are perfectly suited to apply the Wiener–Itô chaos expansion, Malliavin calculus and Stein's method. We refer to [91] and [69] (in this volume) for more details. There limit theorems are stated and more recent developments are pointed out.

Questions of interest not mentioned in the current notes concern for instance percolation problems. For recent developments in this context, we refer, e.g., to the recent book by Haenggi [40].

3.2.1 Random Simplicial Complexes

A very recent line of research is based on the use of random geometric graphs for constructing random simplicial complexes. For instance, given the Gilbert graph of a Poisson point process η_t, we construct the Vietoris–Rips complex $R(\delta_t)$ by calling $F = \{x_{i_1}, \dots, x_{i_{k+1}}\}$ a k-face of $R(\delta_t)$ if all pairs of points in F are connected by an edge in the Gilbert graph. This results in a random simplicial complex, and it is particularly interesting to investigate its combinatorial and topological structure.

For example, counting the number $N_t^{(k)}$ of k-faces is equivalent to a particular subgraph counting. By definition this is a U-statistic given by

$$N_t^{(k)} = N_t^{(k)}(W, \delta_t) = \frac{1}{(k+1)!} \sum_{(x_1,\dots,x_{k+1}) \in \eta_{t,\neq}^{k+1}} \mathbb{1}\{\|x_i - x_j\| \le \delta_t,\ \forall 1 \le i, j \le k + 1\}.$$

Using the Slivnyak–Mecke theorem (see [104, Sect. 3.2]), the expectation of $N_t^{(k)}$ can be computed. Central limit theorems and a concentration inequality follow from results for local U-statistics. A particularly tempting problem is the asymptotic behavior of the Betti-numbers of this random simplicial complex. We refer to [29, 60, 62, 69] and to the recent survey article by Kahle [61] for further information.

3.3 Poisson Processes on Grassmannians

Let η_t be a Poisson process on the space $A(d, i)$ of affine i-flats with a σ-finite intensity measure $\mu_t = t\mu_1$, $t > 0$. Assume in particular that μ_t is absolutely continuous with respect to the Haar measure μ_i^d on $A(d, i)$. This implies that two subspaces $L_1, L_2 \in \eta^2_{t,\neq}$ are almost surely in general position. If $2i < d$ the intersection $L_1 \cap L_2$ is almost surely empty and of interest is the linear hull of the subspace parallel to L_1 and L_2, which is of dimension $2i$ with probability one. If $2i \geq d$, then the dimension of the linear hull of the subspace parallel to L_1 and L_2 is d and of interest is the intersection $L_1 \cap L_2$, which is an affine subspace of dimension $2i - d$ with probability one.

Crucial in all the following results mentioned for both cases is the fact that the functionals of interest are Poisson U-statistics and thus admit a finite chaos expansion. This makes it particularly tempting to use methods from the Malliavin calculus for proving distributional results.

3.3.1 Intersection Processes of Poisson Flat Processes

Starting from a stationary process η_t of i-flats in $\mathbb{R}^d$ with $d/2 \leq i \leq d-1$, we obtain for given $k \leq d/(d-i)$ a stationary process $\eta_t^{(k)}$ of $[ki - (k-1)d]$-flats by taking the intersection of any k flats from η_t whose intersection is of the correct dimension. If η_t is Poisson, then the intensity $t^{(k)}$ and the directional distribution $\sigma^{(k)}$ of this k-fold intersection process $\eta_t^{(k)}$ of η_t can be related to the intensity t and the directional distribution σ of η_t by

$$t^{(k)}\sigma^{(k)}(\cdot) = \frac{t^k}{k!} \int\limits_{A(d,i)} \dots \int\limits_{A(d,i)} \mathbb{1}\{L_1 \cap \dots \cap L_k \in \cdot\}[L_1, \dots, L_k]\, \sigma(\mathrm{d}L_k) \dots \sigma(\mathrm{d}L_1),$$

where the subspace determinant $[L_1, \dots, L_k]$ is defined as the $k(d-i)$-dimensional volume of the parallelepiped spanned by orthonormal bases of $L_1^\perp, \dots, L_k^\perp$. Natural questions which arise at this point are the following:

- For which choice of σ will $t^{(k)}$ be maximal if t is fixed?
- Are t and σ uniquely determined by the intersectional data $t^{(k)}$ and $\sigma^{(k)}$?
- If uniqueness holds, is there a stability result as well? That is, are $t\sigma$ and $\hat{t}\hat{\sigma}$ close to each other (in a quantitative sense) if $t^{(k)}\sigma^{(k)}$ and $\hat{t}^{(k)}\hat{\sigma}^{(k)}$ are close?

For further information on this topic, see Sect. 4.4 in [104].

Since in applications the intersection process can only be observed in a convex window W, one is in particular interested in the sum of their j-th intrinsic volumes given by

$$\Phi_t = \frac{1}{k!} \sum_{(L_1,\ldots,L_k)\in\eta^k_{t,\neq}} V_j(L_1 \cap \ldots \cap L_k \cap W)$$

for $j = 0,\ldots,d-k(d-i)$. The fact that the summands in the definition of Φ_t are bounded and have a bounded support ensures that the sum exists.

The expectation of Φ_t can be calculated using the Slivnyak–Mecke theorem, which yields

$$\mathbb{E}\Phi_t = \frac{1}{k!} t^k \int \ldots \int V_j(L_1 \cap \ldots \cap L_k \cap W)\, \mu_1(\mathrm{d}L_1)\ldots\mu_1(\mathrm{d}L_k).$$

If μ_t is also translation invariant this leads to the question to determine certain chord power integrals of the observation window W or more general integrals involving powers of the intrinsic volumes of intersections $L \cap W$ where L is an affine subspace.

Recent contributions deal with variances and covariances, multivariate central limit theorems [74] (see also [69]), and the distribution of the m-smallest intersection [108]. For further detailed investigations we refer to the recent contribution by Hug et al. [59].

3.3.2 Proximity of Poisson Flat Processes

A different situation arises if we consider a stationary process of i-flats in $\mathbb{R}^d$ with $1 \le i < d/2$. In this case, generically we expect that any two different i-flats $L_1, L_2 \in \eta_t$ are disjoint. A natural way to investigate the geometric situation in this setting is to study the distances between disjoint pairs of i-dimensional flats, or more generally to consider the proximity functional.

We associate with such a pair $(L_1, L_2) \in \eta^2_{t,\neq}$ (in general position) a unique pair of points $x_1 \in L_1$ and $x_2 \in L_2$ such that $\|x_1 - x_2\|$ equals the distance between L_1 and L_2. This gives rise to a process of triples $(m(L_1,L_2), d(L_1,L_2), L(L_1,L_2))$, where $m(L_1,L_2) := (x_1 + y_2)/2$ is the midpoint, $d(L_1,L_2) := \|x_1 - x_2\|$ is the distance, and $L(L_1,L_2) \in G(d,1)$ is the subspace spanned by the vector $x_1 - x_2$.

The stationary process of midpoints and its intensity have been studied in [97] for a Poisson process (see also Sect. 4.4 in [104]), and more recently in [109]. Assume that η_t is a Poisson process on the space $A(d,i)$, $i < \frac{d}{2}$, with intensity measure $\mu_t = t\mu_1$. The midpoints $m(L_1,L_2) = \frac{1}{2}(x_1 + x_2)$ form a point process of infinite intensity, hence we restrict it to the point process

$$\{m(L_1,L_2) :\ d(L_1,L_2) \le \delta,\ L_1, L_2 \in \eta^2_{t,\neq}\}$$

and are interested in the number of midpoints in W, that is,

$$\Pi_t = \Pi_t(W,\delta) = \frac{1}{2} \sum_{(L_1,L_2)\in\eta^2_{t,\neq}} \mathbb{1}\{d(L_1,L_2) \le \delta,\ m(L_1,L_2) \in W\}.$$

The Slivnyak–Mecke formula shows that $\mathbb{E}\Pi_t$ is of order $t^2\delta^{d-2i}$. Schulte and Thäle [109] proved convergence of the suitably normalized random variable Π_t to a normally distributed variable with error term of order $t^{-\frac{d-i}{2}}$. Moreover, they showed that after suitable rescaling the ordered distances asymptotically form an inhomogeneous Poisson point process on the positive real axis. In [69], the authors add to this a concentration inequality around the median m_t of Π_t which shows that the tails of the distribution are bounded by

$$\exp\left(-\frac{1}{4}\frac{u}{\sqrt{u+m_t}}\right)$$

for $\frac{u}{\sqrt{u+m_t}} \ge e^2 \sup_{L_0\in[W]} \mu_t(\{L:\ d(L_0,L) \le \delta\})$.

For the process of triples $(m(L_1,L_2), d(L_1,L_2), L(L_1,L_2))$ a more detailed analysis has been carried out in [59], which also emphasizes the duality of concepts and results as compared to the intersection process (of order $k = 2$) described before. While the proximity process provides a "dual counterpart" to the intersection process of order two, no satisfactory analogue for intersection processes of higher order is known so far.

3.4 Random Mosaics

Another widely used model of stochastic geometry is that of a random mosaic (tessellation). A deterministic mosaic of Euclidean space $\mathbb{R}^d$ is a family of countably many d-dimensional convex bodies $C_i \subset \mathbb{R}^d$, $i \in \mathbb{N}$, with mutually disjoint interiors, whose union is the whole space and with the property that each compact set intersects only finitely many of the sets. The individual sets of the family, which necessarily are polytopes, are called the cells of the tessellation. It is clear that this concept can be extended in various directions, for instance by dropping the convexity assumption on the cells or by allowing local accumulations of cells, which leads to a more general partitioning of space.

Formally, a random mosaic (tessellation) X in $\mathbb{R}^d$ is defined as a simple particle process such that for each realization the collection of all particles constitutes a mosaic. In addition to the cells of the mosaic, the collection of k-dimensional faces of the cells, for each $k \in \{0,\dots,d\}$, provides an interesting geometric object which combines features of a particle process, a random closed set (considering for instance the union set), or a random geometric graph. For example, coloring the cells of the tessellation black or white, independently of each other and independently of X, one can ask for the probability of an infinite black connected component or study the asymptotic behavior of mean values and variances of functionals of the intersection sets $Z_B \cap W$, where Z_B denotes the union of the black cells and W is an increasing observation window. For an introduction to such percolation models we refer the reader to [12, 13, 72, 77]. A first systematic investigation of central limit theorems in more general continuous percolation models related to stationary random tessellations is carried out in [78].

3.4.1 Typical Cells and Faces

In the following, we always consider stationary random tessellations X in $\mathbb{R}^d$. By stationarity, the intensity measure $\mathbb{E}X$ of X, which we always assume to be locally finite and nonzero, is translation invariant. Let $c : \mathcal{K}^d \to \mathbb{R}^d$ denote a center function. By this we mean a measurable function which is translation covariant, that is, $c(K+x) = c(K)+x$ for all $K \in \mathcal{K}^d$ and $x \in \mathbb{R}^d$. W.l.o.g. we take $c(K)$ to be the center of the circumball, and define $\mathcal{K}_0^d := \{K \in \mathcal{K}^d : c(K) = o\}$ as in Sect. 2.3. Then

$$\mathbb{E}X = t \int_{\mathcal{K}_0^d} \int_{\mathbb{R}^d} \mathbb{1}\{C + x \in \cdot\}\, \ell_d(\mathrm{d}x)\, \mathbb{Q}(\mathrm{d}C),$$

where $t > 0$ and $\mathbb{Q}$ is a probability measure on $\mathcal{K}_0^d$ which is concentrated on convex polytopes. A random polytope Z with distribution $\mathbb{Q}$ is called a *typical cell* of X. This terminology can be justified by Palm theory or in a "statistical sense." In addition to such a "mean cell" we also consider the cell containing a fixed point in its interior. Because of stationarity, we may choose the origin and hence the zero cell Z_0 of a given stationary tessellation. Applying the same kind of reasoning to the stationary process $X^{(k)}$ of k-faces of X, we are led to the intensity $t^{(k)}$ and the distribution $\mathbb{Q}^{(k)}$ of the typical k-face $Z^{(k)}$ of X which are determined by

$$t^{(k)}\mathbb{Q}^{(k)}(\cdot) = \mathbb{E}\left[\sum_{F \in X^{(k)}} \mathbb{1}\{c(F) \in B\}\mathbb{1}\{F - c(F) \in \cdot\}\right],$$

where $B \subset \mathbb{R}^d$ is a Borel set with $\ell_d(B) = 1$ and

$$t^{(k)} = \mathbb{E}\left[\sum_{F \in X^{(k)}} \mathbb{1}\{c(F) \in B\}\right].$$

Let M_k denote a random measure concentrated on the union of the k-faces of X which is given by

$$M_k(\cdot) = \sum_{F \in X^{(k)}} \mathcal{H}^k(\cdot \cap F).$$

Then the distribution of the k-volume weighted typical k-face $Z_0^{(k)}$ is defined by

$$\frac{1}{\mathbb{E}M_k(B)} \mathbb{E} \int_B \mathbb{1}\left\{F_k(X^{(k)} - x) \in \cdot\right\} M_k(\mathrm{d}x),$$

where again $B \subset \mathbb{R}^d$ is a Borel set with $\ell_d(B) = 1$ and $F_k(X^{(k)} - x)$ is the $\mathbb{P}$-a.s. unique k-face of $X^{(k)} - x$ containing o if x is in the support of M_k. Then, for any nonnegative, measurable function h on convex polytopes, we obtain

$$\mathbb{E}h\left(Z_0^{(k)} - c(Z_0^{(k)})\right) = \frac{\mathbb{E}[h(Z^{(k)}V_k(Z^{(k)})]}{\mathbb{E}[V_k(Z^{(k)})]}, \tag{5}$$

which also explains why $Z_0^{(k)}$ is called the volume weighted typical k-face of X. This relation between the two types of typical faces is implied by Neveu's exchange formula. In the particular case $k = d$ we have $Z_0^{(d)} = Z_0$. Here we followed the presentation in [7, 8, 98, 99].

For general stationary random mosaics it is apparently difficult to establish distributional results. More is known about various mean values and intensities. For instance,

$$\sum_{i=0}^{d} (-1)^i t^{(i)} = 0 \tag{6}$$

is an Euler type relation for the intensities, which points to an underlying general geometric fact (Gram's relation). If Z_k denotes the union of the k-faces of X (its k-skeleton), then the specific Euler characteristic

$$\bar{\chi}_k := \lim_{r\to\infty} \frac{1}{r^d} \mathbb{E}\chi(Z_k \cap r[0,1]^d)$$

exists and satisfies

$$\bar{\chi}_k = \sum_{i=0}^{k} (-1)^i t^{(i)}.$$

Mean value relations for the mean number of j-faces contained in (or containing) a typical k-faces if $j < k$ (respectively, $j \geq k$) or relations for the mean intrinsic volumes of the typical k-faces $t^{(k)}\mathbb{E}V_j(Z^{(k)})$ are also known (see [104, Sect. 10.1] for this and related results). More generally, asymptotic mean values and second order properties for functionals of certain colored random mosaics have been investigated in [78].

A different setting is considered in [43]. The starting point is a general stationary ergodic random tessellation in $\mathbb{R}^d$. With each cell a random inner structure is associated (for instance, a point pattern, fiber system, or random tessellation) independently of the given mosaic and of each other. Formally, this inner structure is generated by a stationary random vector measure J_0. In this framework, with respect to an expanding observation window strong laws of large numbers, asymptotic covariances and multivariate central limit theorems are obtained for a normalized functional, which provides an unbiased estimator for the intensity vector of J_0. Applications to communication networks are then discussed in dimension two under more specific model assumptions involving Poisson–Voronoi and Poisson line tessellations as the frame tessellation as well as the tessellations used for the nesting sequence.

3.4.2 Poisson Hyperplane Mosaics

A hyperplane process η_t in $\mathbb{R}^d$ with intensity $t > 0$ naturally divides $\mathbb{R}^d$ into convex polytopes, and the resulting mosaic is called hyperplane mosaic. In the following, we assume that all required intensities are finite (and positive). Let X be the stationary hyperplane mosaic induced by η_t. Let

$$\frac{d_j^{(k)}}{t^{(k)}} = \int V_j(K)\,\mathbb{Q}^{(k)}(\mathrm{d}K) = \mathbb{E}V_j(Z^{(k)})$$

denote the mean j-th intrinsic volume of the typical k-face $Z^{(k)}$ of the mosaic X, where $t^{(k)}$ is again the intensity of the process of k-faces. We call $d_j^{(k)}$ the specific j-th intrinsic volume of the k-face process $X^{(k)}$. If $n_{k,j}$, for $0 \leq j \leq k \leq d$, denotes the mean number of j-faces of the typical k-face, then the relations

$$d_j^{(k)} = \binom{d-j}{d-k} d^{(j)}, \qquad t^{(k)} = \binom{d}{k} t^{(0)}, \qquad n_{k,j} = 2^{k-j}\binom{k}{j}$$

complement the Euler relation (6) valid for any random tessellation (see [104, Theorem 10.3.1]). In the derivation of these facts the property is used that each j-face of X lies in precisely $\binom{d-j}{d-k}$ flats of the $(d-k)$-fold intersection process $\eta_{t,(d-k)}$ and therefore in $2^{k-j}\binom{d-j}{d-k}$ faces of dimension k of X. Further results can be obtained, for instance, if the underlying stationary hyperplane process η_t is Poisson. To prepare this, we observe that the intensity measure of η_t is of the form

$$t \int_{S^{d-1}} \int_{\mathbb{R}} \mathbb{1}\{u^{\perp} + xu \in \cdot\} \, \ell_1(\mathrm{d}x) \, \sigma(\mathrm{d}u), \tag{7}$$

where $t > 0$ and σ is an even probability measure on the unit sphere. Since for $u \in \mathbb{R}^d$ the left-hand side of

$$\frac{t}{2} \int_{S^{d-1}} |\langle u, v\rangle| \, \sigma(\mathrm{d}v) =: h(\Pi_X, u)$$

is a positively homogeneous convex function (of degree 1), it is the support function of a uniquely defined convex body $\Pi_X \in \mathcal{K}^d$, which is called the *associated zonoid* of X. This zonoid can be used to express basic quantities of the mosaic X. For instance, we have

$$d_j^{(k)} = \binom{d-j}{d-k} V_{d-j}(\Pi_X), \qquad t^{(k)} = \binom{d}{k} V_d(\Pi_X)$$

(see [104, Theorem 10.3.3]). If X (or η_t) is isotropic, then Π_X is a ball and these relations are directly expressed in terms of constants and the intensity t.

In [102], Schneider found an explicit formula for the covariances of the total face contents of the typical k-face of a stationary Poisson hyperplane mosaic. Let $L_i(P)$ be the total i-face contents of a polytope $P \subset \mathbb{R}^d$, that is,

$$L_i(P) = \sum_{F \in \mathcal{F}_i(P)} \mathcal{H}^i(F).$$

The main result is a general new formula for the second moments $\mathbb{E}(L_r L_s)(Z^{(k)})$, which is obtained by an application of the Slivnyak–Mecke formula and clever geometric dissection arguments (refining ideas of R. Miles) in combination with the mean values

$$\mathbb{E}L_r(Z^{(k)}) = \frac{2^{k-r}\binom{k}{r}}{t\binom{d}{r}} V_{d-r}(\Pi_X),$$

which follow from [100]. As a consequence of these formulas and deep geometric inequalities, namely the *Blaschke–Santaló inequality* and the *Mahler inequality for zonoids*, he deduced that the variance $\mathrm{Var}(f_0(Z^{(k)}))$ is maximal if and only if X is isotropic and minimal if and only if X is a parallel process (involving d fixed directions only). A similar result is obtained for the variance of the volume of the typical cell. In the isotropic case, explicit formulas for these variances and, more generally, for the covariances of the face contents are obtained.

In addition to the typical cell $Z = Z^{(d)}$ of a stationary hyperplane tessellation, we consider the almost surely unique cell $Z_0 = Z_0^{(d)}$ containing the origin (the zero cell). One relation between these two random polytopes is given in (5). Another one describes the distribution of the typical cell (where here the highest vertex in a certain admissible direction is chosen as the center function) as the intersection of Z_0 with a random cone $T(H_1, \dots, H_d)$ generated by d independent random hyperplanes sampled according to a distribution determined by the direction distribution σ of η_t. From this description, one can deduce that up to a random translation, Z is contained in Z_0 (see Theorem 10.4.7 and Corollary 10.4.1 in [104]).

For the zero cell, mean values of some functionals are explicitly known. For instance,

$$\mathbb{E}L_r(Z_0) = 2^{-d} d! V_{d-r}(\Pi_X) V_d(\Pi_X^\circ),$$

where Π_X° is the polar body of the associated zonoid of X. Choosing $r = 0$, we get the mean number of vertices, and the choice $r = d$ gives the mean volume of Z_0. It follows, for instance, that

$$2^d \leq \mathbb{E}f_0(Z_0) \leq d! 2^{-d} \kappa_d^2$$

with equality on the left side if X is a parallel process, and with equality on the right side if X is isotropic. A related stability result has been established in [14].

3.4.3 Distributional Results

One of the very few distributional results which are known for hyperplane processes is the following. It involves the inradius $r(K)$ of a convex body K, which is defined as the maximal radius of a ball contained in K. We call a hyperplane process nondegenerate if its directional distribution is not concentrated on any great subsphere.

Theorem 2 *Let Z be the typical cell of a stationary mosaic generated by a (nondegenerate) stationary Poisson hyperplane process η_t with intensity $t > 0$. Then*

$$\mathbb{P}(r(Z) \leq a) = 1 - \exp(-2ta), \qquad a \geq 0.$$

Clearly $r(Z) \geq a$ if and only if a ball of radius a is contained in Z. An extension covering more general inclusion probabilities (for homothetic copies of an arbitrary convex body) and typical k-faces has been established in [54, Sect. 4, (9)].

In order to study distributional properties of lower-dimensional typical faces, Schneider [98] showed that for $k \in \{1, \dots, d-1\}$ the distribution of the volume-weighted typical k-face can be described as the intersection of the zero cell with a random k-dimensional linear subspace. To state this result, let η_t denote a stationary Poisson hyperplane process in $\mathbb{R}^d$ with intensity measure as given in (7). Further, let

$t^{(d-k)}$ denote the intensity and $\sigma^{(d-k)}$ the directional distribution (a measure on the Borel sets of $G(d,k)$) of the intersection process $\eta_{t,(d-k)}$ of order $d-k$ of η_t. Both quantities are determined by the relation

$$t^{(d-k)}\sigma^{(d-k)}(\cdot) = \frac{t^{d-k}}{(d-k)!}\int\limits_{(S^{d-1})^{d-k}} \mathbb{1}\{u_1^\perp \cap \ldots \cap u_{d-k}^\perp \in \cdot\}$$
$$[u_1,\ldots,u_{d-k}]\,\sigma^{d-k}(\mathrm{d}(u_1,\ldots,u_{d-k})),$$

where $[u_1,\ldots,u_{d-k}]$ denotes the $(d-k)$-volume of the parallelepiped spanned by $u_1,\ldots,u_{d-k}$.

The next theorem summarizes results from [98, Theorem 1] and from [54, Theorem 1].

Theorem 3 *Let X denote the stationary hyperplane mosaic generated by a stationary Poisson hyperplane process η_t. Then the distribution of the volume-weighted typical k-face of X is given by*

$$\mathbb{P}(Z_0^{(k)} \in \cdot) = \int\limits_{G(d,k)} \mathbb{P}(Z_0 \cap L \in \cdot)\,\sigma^{(d-k)}(\mathrm{d}L).$$

The distribution of the typical k-face equals

$$\mathbb{P}(Z^{(k)} \in \cdot) = \int\limits_{G(d,k)} \mathbb{P}(Z(X \cap L) \in \cdot)\,R_k(\mathrm{d}L),$$

hence it is described in terms of the typical cells of the induced mosaics $X \cap L$ in k-dimensional subspaces sampled according to the directional distribution

$$R_k(\cdot) = \frac{V_{d-k}(\Pi_X)}{\binom{d}{k}V_d(\Pi_X)}\int\limits_{G(d,k)} \mathbb{1}\{L \in \cdot\}V_k(\Pi_X|L)\,\sigma^{(d-k)}(\mathrm{d}L)$$

of the typical k-face of X.

These results turned out to be crucial for extending various results for typical (volume-weighted) faces, which had been obtained before for the typical cell (the zero cell).

3.4.4 Large Cells: Kendall's Problem

Next we turn to Kendall's problem on the asymptotic shape of the large cells of a stationary but not necessarily isotropic Poisson hyperplane tessellation. The original problem (Kendall's conjecture) concerned a stationary isotropic Poisson line tessellation in the plane and suggested that the conditional law for the shape of the zero cell Z_0, given its area $V_2(Z_0) \to \infty$, converges weakly to the degenerate law concentrated at the circular shape. Miles [75] provided some heuristic ideas for the proof of such a result and suggested also various modifications. The conjecture was strongly supported by Goldman [34], a first solution came from Kovalenko [64, 65]. Still the approaches of these papers were essentially restricted to the Euclidean plane and made essential use of the isotropy assumption.

The contribution [56] marks the starting point for a sequence of investigations which provide a resolution of Kendall's problem in a substantially generalized form. To describe the result in some more detail, let η_t be a (nondegenerate) stationary Poisson hyperplane process in $\mathbb{R}^d$ with intensity $t > 0$ and directional distribution σ. In order to find a potential asymptotic shape for the zero cell Z_0 of the induced Poisson hyperplane tessellation, we first have to exhibit a candidate for such a shape (if it exists), then we have to clarify what we mean by saying that two shapes are close and finally it remains to determine a quantity which should be used instead of the "area" of Z_0 to measure the size of the zero cell.

Clearly, a natural candidate for a size functional is the volume V_d. The answer to the first question is less obvious, but is based on a strategy that has repeatedly been used in the literature with great success (see [104, Sect. 4.6] for various examples and references). The main idea is to describe the direction distribution σ in geometric terms. This allows one to apply geometric inequalities such as Minkowski's inequality (2) and its stability improvement, which then can be reinterpreted again in probabilistic terms. Instead of the associated zonoid, for the present problem the Blaschke body associated with η_t, alternatively the direction body B of η_t, turns out to be the right tool. This auxiliary body B is characterized as the unique centered (that is, $B = -B$) d-dimensional convex body $B \in \mathcal{K}^d$ such that the area measure of B satisfies $S_{d-1}(B, \cdot) = \sigma$. The existence and uniqueness of B, for given σ, is a deep result from convex geometry which in its original form is also due to Minkowski (see [101]). Finally, we say that the shape of $K \in \mathcal{K}^d$ is close to the shape of B if

$$r_B(K) = \inf\{s/r - 1 : rB + z \subset K \subset sB + z, z \in \mathbb{R}^d, r, s > 0\}$$

is small. In particular, $r_B(K) = 0$ if and only if K and B are homothetic. Let $\mathcal{K}^d_{(o)}$ denote the set of all $K \in \mathcal{K}^d$ with $o \in K$. For any such K we introduce the constant

$$\tau = \min\{t^{-1}\mathbb{E}\eta_t([K]) : K \in \mathcal{K}^d_{(o)}, V_d(K) = 1\}$$

of isoperimetric type, which can also be expressed in the form

$$e^{-\tau t} = \max\{\mathbb{P}(K \subset Z_0) : K \in \mathcal{K}^d_{(o)}, V_d(K) = 1\}.$$

The following theorem summarizes Theorems 1 and 2 in [56] and a special case of Theorem 2 in [51]. The latter provides a far reaching generalization of a result in [34] on the asymptotic distribution of the area of the zero cell of an isotropic stationary Poisson line tessellation in the plane.

Theorem 4 *Under the preceding assumptions, there is a positive constant c_0, depending only on B, such that for every $\epsilon \in (0,1)$ and for every interval $I = [a,b)$ with $a^{1/d}t \geq 1$,*

$$\mathbb{P}(r_B(Z_0) \geq \epsilon \mid V_d(Z_0) \in I) \leq c \exp\left(-c_0 \epsilon^{d+1} a^{1/d} t\right),$$

where c is a constant depending on B and ϵ. Moreover,

$$\lim_{a\to\infty} a^{-1/d} \ln \mathbb{P}(V_d(Z_0) \geq a) = -\tau t.$$

The same result holds for the typical cell Z.

If the size of Z_0 is measured by some other intrinsic volume $V_i(Z_0)$, for $i \in \{2,\ldots,d-1\}$, a similar result is true if η_t is also isotropic (see [57, Theorem 2]). No such result can be expected for the mean width functional V_1. In fact, no limit shape may exist if size is measured by the mean width, which is proved in [51, Theorem 4] for directional distributions with finite support. Most likely a limit shape does not exist if size is measured by the mean width, but for arbitrary σ or in case of the typical cell this is still an open question. Crucial ingredients in the proofs of the results described so far are geometric stability results, which refine geometric inequalities and the discussion of the equality cases for these inequalities.

3.4.5 A General Framework

The results described so far suggest the general question which size functionals indeed lead to asymptotic or limit shapes and how these asymptotic or limit shapes are determined. A general axiomatic framework for analyzing these questions is developed in [51]. The main object of investigation is a Poisson hyperplane process η_t in $\mathbb{R}^d$ (and its induced tessellation) with intensity measure of the form

$$\mathbb{E}\eta_t = t \int_{S^{d-1}} \int_0^\infty \mathbb{1}\{H(u,x) \in \cdot\} x^{r-1} \, \ell_1(\mathrm{d}x)\, \sigma(\mathrm{d}u), \tag{8}$$

where $t > 0$, $r \geq 1$, and σ is an even nondegenerate (that is, not concentrated on any great subsphere) probability measure on the Borel sets of the unit sphere. The case $r = 1$ corresponds to the stationary case. We refer to t as the intensity, r as the distance exponent, and σ as the directional distribution of η_t. Let

$$\Phi(K) := t^{-1}\mathbb{E}\eta_t([K]) = \frac{1}{r} \int_{S^{d-1}} h(K,u)^r \, \sigma(\mathrm{d}u), \qquad K \in \mathcal{K}^d_{(o)},$$

which is called the hitting or parameter functional of η_t, since $t\Phi(K)$ is the mean number of hyperplanes of η_t hitting K. Moreover, we have

$$\mathbb{P}(\eta_t([K]) = n) = \frac{[\Phi(K)t]^n}{n!} \exp(-\Phi(K)t), \qquad n \in \mathbb{N}_0,$$

by the Poisson assumption on η_t.

In Theorem 4 we used the volume functional to bound the size of the zero cell. Many other functionals are conceivable such as the (centered) inradius, the diameter, the width in a given direction, or the largest distance to a vertex of Z_0. It was realized in [51] that in fact any functional Σ on $\mathcal{K}^d_{(o)}$ which satisfies some natural axioms (continuity, homogeneity of a fixed degree $k > 0$ and monotonicity under set inclusion) qualifies as a size functional. From this it already follows that a general sharp inequality of isoperimetric type is satisfied, that is,

$$\Phi(K) \geq \tau \Sigma(K)^{r/k}, \qquad K \in \mathcal{K}^d_{(o)}, \tag{9}$$

with a positive constant $\tau > 0$. The convex bodies K for which equality is attained are called extremal. Among the bodies of size $\Sigma(K) = 1$ these are precisely the bodies for which

$$\mathbb{P}(K \subset Z_0) \leq e^{-\tau t}$$

holds with equality (thus maximizing the inclusion probability). The final ingredient required in this general setting, if Φ, Σ are given, is a deviation functional ϑ on $\{K \in \mathcal{K}^d_{(o)} : \Sigma(K) > 0\}$, which should be continuous, nonnegative, homogeneous of degree zero, and satisfy $\vartheta(K) = 0$ for some K with $\Sigma(K) > 0$ if and only if K is extremal. Then exponential bounds of the form

$$\mathbb{P}(\vartheta(Z_0) \geq \epsilon \mid \Sigma(Z_0) \in [a, b]) \leq c \exp\left(-c_0 f(\epsilon) a^{r/k} t\right) \tag{10}$$

with a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ which is positive on $(0, \infty)$, with $f(0) = 0$, and which satisfies

$$\Phi(K) \geq (1 + f(\epsilon))\tau \Sigma(K)^{r/k} \quad \text{if } \vartheta(K) \geq \epsilon,$$

are established in [51]. Thus if we know that K has positive distance $\vartheta(K)$ from an extremal body, we can again use this information to obtain an improved version of a very general inequality of isoperimetric type. As mentioned before, results of this form are known as stability results. Note that for the choice $\Sigma = \Phi$, the inequality (9) becomes a tautological identity and all $K \in \mathcal{K}^d_{(o)}$ with $K \neq \{o\}$ are extremal. Hence, in this case ϑ is identically zero and (10) holds trivially.

Moreover, for the asymptotic distributions of size functionals it is shown that

$$\lim_{a \to \infty} a^{-r/k} \ln \mathbb{P}(\Sigma(Z_0) \geq a) = -\tau t,$$

thus providing a far reaching extension of the result for the volume functional [51]. The paper [51] contains also a detailed discussion of various specific choices of parameters and functionals which naturally occur in this context and which exhibit a rich variety of phenomena. In the next subsection we point out how this setting extends to Poisson–Voronoi tessellations. In the case of stationary and isotropic Poisson hyperplane tessellations, a similar general investigation is carried out in [52]. Extensions to lower-dimensional faces in Poisson hyperplane mosaics, which are based on the above-mentioned distributional results for k-faces, are considered in [53, 54].

Much less is known about the shape of small cells, although this has also been asked for by Miles [75]. For parallel mosaics in the plane, some work has been done in [10]. Recently, limit theorems for extremes of stationary random tessellations have been explored in [22, 27], but the topic has not been exhaustively investigated so far. In the survey [21], Calka discusses some generalizations of distributional results for the largest centered inball (centered inradius) R_M, the smallest centered circumball (centered circumradius) and their joint distribution, for an isotropic Poisson hyperplane process with distance exponent $r \geq 1$. These radii are related to covering probabilities of the unit sphere by random caps. The two-dimensional situation had already been considered in [20]. In particular, Calka points out that after a geometric inversion at the unit sphere and by results available for convex hulls of Poisson point processes in the unit ball (see [23, 24]), the asymptotic behavior of $\mathbb{P}(R_M \geq t + t^{\delta} \mid R_m = t)$ can be determined for a suitable choice of δ as $t \to \infty$. In addition, L^1-convergence, a central limit theorem, and a moderate deviation result are available for the number of facets and the volume of Z_0.

3.4.6 Random Polyhedra

The techniques developed for the solution of Kendall's problem turned out to be useful also for the investigation of approximation properties of random polyhedra derived from a stationary Poisson hyperplane process η_t with intensity $t > 0$ and directional distribution σ. Here the basic idea is to replace the zero cell by the K-cell Z_t^K defined as the intersection of all half-spaces H^- bounded by hyperplanes $H \in \eta_t$ for which $K \subset H^-$. Let d_H denote the Hausdorff distance of compact sets in $\mathbb{R}^d$, and let K^y be the convex hull of K and $\{y\}$. If the support of the area measure $S_{d-1}(K, \cdot)$ is contained in the support of σ, then

$$\mathbb{P}(d_H(K, Z_t^K) > \epsilon) \leq c_1(\epsilon) \exp\left(-c_2 t \mu(K, \sigma, \epsilon)\right),$$

where $c_1(\varepsilon), c_2$ are constants and

$$\mu(K, \sigma, \epsilon) = \min_{y \in \partial(K + \epsilon B^d)} \int_{S^{d-1}} [h(K^y, u) - h(K, u)]\, \sigma(\mathrm{d}u) > 0;$$

see [55, Theorem1]. Using this bound as a starting point, under various assumptions on the relation between the body K to be approximated and the directional distribution σ of the approximating hyperplane process, almost sure convergence $d_H(K, Z_t^K) \to 0$ is shown as the intensity $t \to \infty$, including bounds for the speed of convergence. It would be interesting to consider the rescaled sequence

$$\left(\frac{t}{\log t}\right)^{\frac{2}{d+1}} d_H(K, Z_t^K)$$

and to obtain further geometric information about the limit, for instance, if σ is bounded from above and from below by a multiple of spherical Lebesgue measure.

3.4.7 Poisson–Voronoi and Delaunay Mosaics

Perhaps the most common and best known tessellation in Euclidean space is the Voronoi tessellation. A Voronoi tessellation arises from a locally finite set $\eta_t \subset \mathbb{R}^d$ (deterministic or random) of points by associating with each point $x \in \eta_t$ the cell

$$v_{\eta_t}(x) := \{z \in \mathbb{R}^d : \|z - x\| \leqslant \|z - y\| \text{ for all } y \in \eta_t\}$$

with nucleus (center) x. One reason for the omnipresence of Voronoi tessellations is that they are related to a natural growth process starting simultaneously at all nuclei at the same time. If η_t is a stationary Poisson process with intensity $t > 0$, then the collection of all cells $v_{\eta_t}(x)$, $x \in \eta_t$, is a random tessellation X of $\mathbb{R}^d$ which is called Poisson–Voronoi tessellation. The distribution of the typical cell of X is naturally defined by

$$\mathbb{Q}(\cdot) := \frac{1}{t}\mathbb{E}\int_B \mathbb{1}\{v_{\eta_t}(x) - x \in \cdot\}\, \eta_t(\mathrm{d}x), \tag{11}$$

where $B \subset \mathbb{R}^d$ is an arbitrary Borel set with volume 1. A random polytope Z with distribution $\mathbb{Q}$ is called typical cell of X. An application of the Slivnyak–Mecke theorem shows that the typical cell Z is equal in distribution to $v_{\eta_t+\delta_o}(o)$, hence Z is stochastically equivalent to the zero cell of a Poisson hyperplane tessellation with generating Poisson hyperplane process given by $Y = \sum_{x\in\eta_t} \delta_{H(x)}$, where $H(x)$ is the mid-hyperplane of o and x. It is easy to check that Y is isotropic but nonstationary with intensity measure

$$\mathbb{E}Y(\cdot) = 2^d t \int_{S^{d-1}} \int_0^\infty \mathbb{1}\{H(u,x) \in \cdot\}\, x^{d-1}\, \ell_1(\mathrm{d}x)\, \mathcal{H}^{d-1}(\mathrm{d}u), \tag{12}$$

where $H(u,x) := u^{\perp} + xu$ is the hyperplane normal to u and passing through xu. Hence, Y perfectly fits into the framework of the parametric class of Poisson hyperplane processes discussed before. This also leads to the following analogue (see [57]) of Theorem 4. To state it, let $\vartheta(K)$, for a convex body K containing the origin in its interior, be defined by $\vartheta(K) := (R_o - r_o)/(R_o + r_o)$, where R_o is the radius of the smallest ball with center o containing K and r_o is the radius of the largest ball contained in K and center o.

Theorem 5 *Let X be a Poisson–Voronoi tessellation as described above with typical cell Z. Let $k \in \{1,\ldots,d\}$. There is a constant c_d, depending only on the dimension, such that the following is true. If $\varepsilon \in (0,1)$ and $I = [a,b)$ ($b = \infty$ permitted) with $a^{d/k}t \geq 1$, then*

$$\mathbb{P}\left(\vartheta(Z) \geq \varepsilon \mid V_k(Z) \in I\right) \leq c_{d,\varepsilon} \exp\left(-c_d \varepsilon^{(d+3)/2} a^{d/k} t\right),$$

where $c_{d,\varepsilon}$ is a constant depending on d and ε.

It should be noted that conditioning on the mean width V_1 is not excluded here. Moreover, asymptotic distributions of the intrinsic volumes of the typical cell can be determined as well. Although in retrospect this follows from the general results in [51], specific geometric stability results have to be established.

The shape of large typical k-faces in Poisson–Voronoi tessellations, with respect to the generalized nucleus as center function, has been explored in [53]. Here large typical faces are assumed to have a large centered inradius. A corresponding analysis for large k-volume seems to be difficult. In this context, the joint distribution of the typical k-face and the typical k-co-radius is described explicitly and related to a Poisson process of k-dimensional halfspaces with explicitly given intensity measure.

The distributional results obtained in [53] complement fairly general distributional properties of stationary Poisson–Voronoi tessellations that have been established by Baumstark and Last [7]. In particular, they describe the joint distribution of the $d-k+1$ neighbors of the k-dimensional face containing a typical point (i.e., a point chosen uniformly) on the k-faces of the tessellation. Thus they generalize in particular the classical result about the distribution of the typical cell of the Poisson–Delaunay tessellation, which is dual to the given Poisson–Voronoi tessellation. The combinatorial nature of this duality and its consequences are nicely described in [104, Sect. 10.2]. Kendall's problem for the typical cell in Poisson–Delaunay tessellations is explored in [50] (see also [48]).

3.4.8 High-Dimensional Mosaics and Polytopes

Despite significant progress, precise and explicit information about mean values or even variances and higher moments in stochastic geometry is rather rare. This is one reason why often asymptotic regimes are considered, where the number of points, the intensity of a point process, or the size of an observation window is growing to infinity. On the other hand, high-dimensional spaces are a central and challenging topic which has been explored for quite some time, motivated by intrinsic interest and applications.

Let X be a Poisson–Voronoi tessellation generated by a stationary Poisson point process with intensity t in $\mathbb{R}^d$. As before, let Z denote its typical cell. By definition (11), Z contains the origin in its interior. It is not hard to show that $t^{-k} \leq \mathbb{E}[V_d(Z)^k] \leq k!t^{-k}$, in particular, $\mathbb{E}[V_d(Z)] = 1/t$. These bounds are independent of the dimension d. Using a much finer analysis, Alishahi and Sharifitabar [1] showed that

$$\frac{c}{t^2\sqrt{d}}\left(\frac{4}{3\sqrt{3}}\right)^d \leq \operatorname{Var}(V_d(Z)) \leq \frac{C}{t^2\sqrt{d}}\left(\frac{4}{3\sqrt{3}}\right)^d,$$

where $c, C > 0$ are absolute constants. In a sense, this suggests that $V_d(Z)$ gets increasingly deterministic. On the other hand, if $B^d(u)$ is a ball of volume u centered at the origin, then

$$V_d(Z \cap B^d(u)) \to t^{-1}\left(1 - e^{-tu}\right), \qquad d \to \infty,$$

in L^2 and in distribution. The paper [1] was the starting point for a more general high-dimensional investigation of the volume of the zero cell Z_0 in a parametric class of isotropic but not necessarily stationary Poisson hyperplane tessellations. This parametric class is characterized by the intensity measure of the underlying Poisson hyperplane process which is of the form (8) but with σ being the normalized spherical Lebesgue measure. That the case of the typical cell of a Poisson–Voronoi tessellation is included in this model can be seen from (12) by choosing the distance exponent $r = d$ and by adjusting the intensities. Depending on the intensity t, the distance parameter r, and the dimension d, explicit formulas for the second moment $\mathbb{E}(V_d(Z_0)^2)$ and the variance $\operatorname{Var}(V_d(Z_0))$ as well as sharp bounds for these characteristics were derived in [45]. Depending on the tuning of these parameters, the asymptotic behavior of $V_d(Z_0)$ can differ dramatically.

To describe an interesting consequence of such variance bounds, we define by $\overline{Z} := V_d(Z)^{-1/d}Z$ the volume normalized typical cell of a Poisson–Voronoi tessellation with intensity t (as above). Let $L \subset \mathbb{R}^d$ be a co-dimension one linear subspace. Then there is an absolute constant $c > 0$ such that

$$\mathbb{P}\left(V_{d-1}\left(\overline{Z} \cap L\right) \geq \sqrt{e}/2\right) \geq 1 - c \cdot \frac{1}{\sqrt{d}}\left(\frac{4}{3\sqrt{3}}\right)^d.$$

This is a very special case of Theorem 3.17 in [46]. It can be paraphrased by saying that with overwhelming probability the *hyperplane conjecture*, a major problem in the asymptotic theory of Banach spaces, is true for this class of random polytopes, see Milman and Pajor [76].

In [46] also the high-dimensional limits of the mean number of faces and an isoperimetric ratio of a mean volume and a mean surface area are studied for the zero cell of a parametric class of random tessellations (as an example of a random polytope). As a particular instance of such a result, we mention that

$$\lim_{d\to\infty} d^{-1/2} \sqrt[d]{\mathbb{E} f_\ell(Z_0)} = \sqrt{2\pi b},$$

where $r = bd$ (with b fixed) increases proportional to the dimension d and ℓ is fixed. It is remarkable that this limit is independent of ℓ. At the basis of this and other results are identities connecting the f-vector of Z_0 to certain dual intrinsic volumes of projections of Z_0 to a deterministic subspace.

3.4.9 Poisson–Voronoi Approximation

Let A be a Borel set in $\mathbb{R}^d$ and let η_t be a Poisson point process in $\mathbb{R}^d$. Assume that we observe η_t and the only information about A at our disposal is which points of η_t lie in A, i.e., we have the partition of the process η_t into $\eta_t \cap A$ and $\eta_t \setminus A$. We try to reconstruct the set A just by the information contained in these two point sets. For that aim we approximate A by the set A_{η_t} of all points in $\mathbb{R}^d$ which are closer to $\eta_t \cap A$ than to $\eta_t \setminus A$.

Applications of the Poisson–Voronoi approximation include nonparametric statistics (see Einmahl and Khmaladze [32, Sect. 3]), image analysis (reconstructing an image from its intersection with a Poisson point process, see [63]), quantization problems (see, e.g., Chap. 9 in the book of Graf and Luschgy [35]), and numerical integration (approximation of the volume of a set A using its intersection with a point process $\eta_t \cap A$).

More formally, let η_t be a homogeneous Poisson point process of intensity $t > 0$, and denote by $v_{\eta_t}(x)$ the Voronoi cell generated by η_t with center $x \in \eta_t$. Then the set A_{η_t} is just the union of the Poisson–Voronoi cells with center lying in A, i.e.,

$$A_{\eta_t} = \bigcup_{x \in \eta_t \cap A} v_{\eta_t}(x).$$

We call this set the *Poisson–Voronoi approximation* of the set A. It was first introduced by Khmaladze and Toronjadze in [63]. They proposed A_{η_t} to be an estimator for A when t is large. In particular, they conjectured that for arbitrary bounded Borel sets $A \subset \mathbb{R}^d$, $d \geqslant 1$,

$$V_d(A_{\eta_t}) \to V_d(A), \quad t \to \infty,$$

$$V_d(A \triangle A_{\eta_t}) \to 0, \quad t \to \infty, \tag{13}$$

almost surely, where $\triangle$ is the operation of the symmetric difference of sets. In full generality this was proved by Penrose [84].

It can be easily shown that for any Borel set $A \subset \mathbb{R}^d$ we have

$$\mathbb{E}V_d(A_{\eta_t}) = V_d(A),$$

since η_t is a stationary point process. Thus $V_d(A_{\eta_t})$ is an unbiased estimator for the volume of A. Relation (13) suggests that

$$\mathbb{E}V_d(A\triangle A_{\eta_t}) \to 0, \quad t \to \infty, \tag{14}$$

although this is not a direct corollary. The more interesting problems are to find exact asymptotic of $\mathbb{E}V_d(A\triangle A_{\eta_t})$, $\mathrm{Var}V_d(A_{\eta_t})$, and $\mathrm{Var}V_d(A\triangle A_{\eta_t})$.

Very general results in this direction are provided by Reitzner et al. [92]. Their results for Borel sets with finite volume $V_d(A)$ depend on the perimeter Per(A) of the set A in the sense of variational calculus. If A is a compact set with Lipschitz boundary (e.g., a convex body), then $\mathrm{Per}(A)$ equals the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}(\partial A)$ of the boundary ∂A of A. In the general case $\mathrm{Per}(A) \leqslant \mathcal{H}^{d-1}(\partial A)$ holds.

If $A \subset \mathbb{R}^d$ is a Borel set with $V_d(A) < \infty$ and $\mathrm{Per}(A) < \infty$, then

$$\mathbb{E}V_d(A\triangle A_{\eta_t}) = c_d \cdot \mathrm{Per}(A) \cdot t^{-1/d}(1 + o(1)), \quad t \to \infty, \tag{15}$$

where $c_d = 2d^{-2}\Gamma(1/d)\kappa_{d-1}\kappa_d^{-1-1/d}$.

The asymptotic order of the variances of A_{η_t} and $A\triangle A_{\eta_t}$ as $t \to \infty$ was first studied in [44] for convex sets and then extended in [92] to arbitrary Borel sets, where also sharp upper bounds in terms of the perimeter are given. A very general result in this direction is due to Yukich [114]. If $A \subset \mathbb{R}^d$ is a Borel set with $V_d(A) < \infty$ and finite $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}(\partial A)$ of the boundary of A, then

$$\mathrm{Var}V_d(A_{\eta_t}) = C_1(A)t^{-1-1/d}(1 + o(1)),$$

and

$$\mathrm{Var}V_d(A\triangle A_{\eta_t}) = C_2(A)t^{-1-1/d}(1 + o(1)), \quad t \to \infty,$$

with explicitly given constants $C_i(A)$.

A breakthrough was achieved by Schulte [107] for *convex* sets A and, more generally, by Yukich [114] for sets with a boundary of finite $(d-1)$-dimensional Hausdorff measure. They proved central limit theorems for $V_d(A_{\eta_t})$ and $V_d(A\triangle A_{\eta_t})$.

Recently, Lachièze-Rey and Peccati [68] proved bounds for the variance, higher moments, and central limit theorems for a huge class of sets, including fractals.

Another interesting open problem is to measure the quality of approximation of a convex set K by K_{η_t} in terms of the Hausdorff distance between both sets. First estimates for the Hausdorff distance are due to Calka and Chenavier [22], very recently Lachièze-Rey and Vega [70] proved precise results on the Hausdorff distance even for irregular sets.

Since $A_{\eta_t} \to A$ in the sense described above, it is of interest to compare the boundary ∂A to the boundary of the Poisson–Voronoi approximation ∂A_{η_t}. This has been explored recently by Yukich [114] who showed that $\mathcal{H}^{d-1}(\partial A_{\eta_t})$—scaled by a suitable factor independent of A—is an unbiased estimator for $\mathcal{H}^{d-1}(\partial A)$, and he also obtained variance asymptotics. We also mention a very recent deep contribution due to Thäle and Yukich [111] who investigate a large number of functionals of A_{η_t}.

3.5 Random Polytopes

The investigation of random polytopes started 150 years ago when Sylvester stated in 1864 his four-point-problem in the Educational Times. Choose n points independently according to some probability measure in $\mathbb{R}^d$. Denote the convex hull of these points by $\text{conv}\{X_1, \ldots, X_n\}$. Sylvester asked for the distribution function of the number of vertices of $\text{conv}\{X_1, \ldots, X_4\}$ in the case $d = 2$.

Random polytopes are linked to other fields and have important applications. We mention the connection to functional analysis: Milman and Pajor [76] showed that the expected volume of a random simplex is closely connected to the so-called isotropic constant of a convex set which is a fundamental quantity in the local theory of Banach spaces.

In this section we will concentrate on recent contributions and refer to the surveys by Hug [49], Reitzner [90], and Schneider [103] for additional information. Let η_t be a Poisson point process with intensity measure of the form $\mu_t = t\mu_1$, $t > 0$, where μ_1 is an absolutely continuous probability measure on $\mathbb{R}^d$. Then the Poisson polytope is defined as $\Pi_t = \text{conv}(\eta_t)$.

There are only few results for given t and general probability measures μ_1. In analogy to Efron [31], it immediately follows from the Slivnyak–Mecke theorem that $\mathbb{E}f_0(\Pi_t) = t - \mathbb{E}\mu_t(\Pi_t))$, connecting the probability content $\mathbb{E}\mu_t(\Pi_t)$ and the expected number of vertices $\mathbb{E}f_0(\Pi_t)$. Identities for higher moments have been given by Beermann and Reitzner [9] who extended this further to an identity between the generating function $g_{I(\Pi_t)}$ of the number of non-vertices or inner points $I(\Pi_t) = |\eta_t| - f_0(\Pi_t)$ and the moment generating function $h_{\mu_t(\Pi_t)}$ of the μ_t-measure of Π_t. Both functions are entire functions on $\mathbb{C}$ and satisfy

$$g_{I(\Pi_t)}(z+1) = h_{\mu_t(\Pi_t)}(z), \qquad z \in \mathbb{C},$$

thus relating the distributions of the number of vertices and the μ_t-measure of Π_t.

3.5.1 General Inequalities

Assume that $K \subset \mathbb{R}^d$ is a compact convex set and set $\mu_t(\cdot) = tV_d(K \cap \cdot)$. We denote by $\Pi_t^K = \text{conv}[\eta_t]$ the Poisson polytope in K.

In this section we describe some inequalities for Poisson polytopes. Based on the work of Blaschke [11], Dalla and Larman [28], Giannopoulos [33], and Groemer [36, 37] showed that

$$\mathbb{E}V_d(\Pi_t^B) \leq \mathbb{E}V_d(\Pi_t^K) \leq \mathbb{E}V_d(\Pi_t^\triangle) \tag{16}$$

where $\Pi_t^\triangle$, resp. Π_t^B denotes the Poisson polytope where the underlying convex set is a simplex, resp. a ball of the same volume as K. The left inequality is true in arbitrary dimensions, whereas the right inequality is just known in dimension $d = 2$ and open in higher dimensions. To prove this extremal property of the simplex in arbitrary dimensions seems to be very difficult and is still a challenging open problem. A positive solution to this problem would immediately imply a solution to the hyperplane conjecture, see Milman and Pajor [76].

There are some elementary questions concerning the monotonicity of functionals of Π_t^K. First, it is immediate that for all $K \in \mathcal{K}^d$ and $i = 1, \ldots, d$,

$$\mathbb{E}V_i(\Pi_t^K) \leq \mathbb{E}V_i(\Pi_s^K)$$

for $t \leq s$. Second, an analogous inequality for the number of vertices is still widely open. It is only known, see [30], that for $t \leq s$

$$\mathbb{E}f_0(\Pi_t^K) \leq \mathbb{E}f_0(\Pi_s^K)$$

for $d = 2$ (and also for smooth convex sets $K \subset \mathbb{R}^3$ if t is sufficiently large). Thirdly, the very natural implication

$$K \subset L \ \Rightarrow \ \mathbb{E}V_d(\Pi_t^K \mid \eta_t(K) = n) \leq \mathbb{E}V_d(\Pi_t^L \mid \eta_t(L) = n)$$

was asked by Meckes and disproved by Rademacher [85]. He showed that for dimension $d \geq 4$ there are convex sets $K \subset L$ such that for t sufficiently small $\mathbb{E}V_d(\Pi_t^K \mid \eta_t(K) = n) > \mathbb{E}V_d(\Pi_t^L \mid \eta_t(L) = n)$. In addition, Rademacher showed that in the planar case this natural implication is true. The case $d = 3$ is still open.

3.5.2 Asymptotic Behavior of the Expectations

Starting with two famous articles by Rényi and Sulanke [94, 95], the investigations focused on the asymptotic behavior of the expected values as t tends to infinity. Due to work of Wieacker [113], Schneider and Wieacker [106], Bárány [2], and

Reitzner [87], for $i = 1, \dots, d$,

$$V_i(K) - \mathbb{E}V_i(\Pi_t^K) = c_i(K)t^{-\frac{2}{d+1}} + o\left(t^{-\frac{2}{d+1}}\right) \tag{17}$$

if K is sufficiently smooth. Investigations by Schütt [110] and more recently by Böröczky et al. [15] succeeded in weakening the smoothness assumption. Clearly, Efron's identity yields a similar result for the number of vertices.

The corresponding results for polytopes are known only for $i = 1$ and $i = d$. In a long and intricate proof, Bárány and Buchta [3] showed that

$$V_d(K) - \mathbb{E}V_d(\Pi_t^K) = c_d(K)t^{-1}\ln^{d-1} t + O\left(t^{-1}\ln^{d-2} t \ln\ln t\right).$$

For $i = 1$, Buchta [18] and Schneider [96] proved that

$$V_1(K) - \mathbb{E}V_1(\Pi_t^K) = c(K)t^{-\frac{1}{d}} + o(t^{-\frac{1}{d}}).$$

Somehow surprisingly, the cases $2 \le i \le d-1$ are still open.

Due to Efron's identity, the results concerning $\mathbb{E}V_d(\Pi_t^K)$ can be used to determine the expected number of vertices of Π_t^K. In [89], Reitzner generalized these results for $\mathbb{E}f_0(\Pi_t^K)$ to arbitrary face numbers $\mathbb{E}f_\ell(\Pi_t^K)$, $\ell \in \{0, \dots, d-1\}$.

3.5.3 Variances

In the last years several estimates have been obtained from which the order of the variances can be deduced, see Reitzner [86, 88, 89], Vu [112], Bárány and Reitzner [5], and Bárány et al. [6]. The results can be summarized by saying that there are constants $\underline{c}(K), \overline{c}(K) > 0$ such that

$$\underline{c}(K)t^{-1}\mathbb{E}V_i(\Pi_t^K) \le \operatorname{Var}V_i(\Pi_t^K) \le \overline{c}(K)t^{-1}\mathbb{E}V_i(\Pi_t^K)$$

and

$$\underline{c}(K)t^{-1}\mathbb{E}f_\ell(\Pi_t^K) \le \operatorname{Var}f_\ell(\Pi_t^K) \le \overline{c}(K)t^{-1}\mathbb{E}f_\ell(\Pi_t^K)$$

if K is smooth or a polytope. It is conjectured that these inequalities hold for general convex bodies. That the lower bound holds in general has been proved in Bárány and Reitzner [5], but the general upper bounds are missing.

A breakthrough are recent results by Calka et al. [26] and Calka and Yukich [25] who succeeded in giving the precise asymptotics of these variances,

$$\operatorname{Var}V_i(\Pi_t^K) = c_{d,i}(K)\, t^{-\frac{d+3}{d+1}} + o(t^{-\frac{d+3}{d+1}})$$

for $i = 1, d$, and

$$\mathrm{Var} f_\ell(\Pi_t^K) = \bar{c}_{d,\ell}(K)\, t^{\frac{d-1}{d+1}} + o(t^{\frac{d-1}{d+1}})$$

if K is a smooth convex body. The dependence of $\bar{c}_{d,\ell}(K)$ on K is known explicitly.

3.5.4 Limit Theorems

First CLTs have been proved by Groeneboom [39], Cabo and Groeneboom [19], and Hsing [47] but only in the planar case. In recent years, methods have been developed to prove CLTs for the random variables $V_d(\Pi_t^K)$ and $f_\ell(\Pi_t^K)$ in arbitrary dimensions. The main ingredients are Stein's method and some kind of localization arguments. For smooth convex sets this was achieved in Reitzner [88], and for polytopes in a paper by Bárány and Reitzner [4]. The results state that there is a constant $c(K)$ and a function $\varepsilon(t)$, tending to zero as $t \to \infty$, such that

$$\left| \mathbb{P}\left(\frac{V_d(\Pi_t^K) - \mathbb{E} V_d(\Pi_t^K)}{\sqrt{\mathrm{Var} V_d(\Pi_t^K)}} \le x \right) - \Phi(x) \right| \le c(K)\, \varepsilon(t)$$

and

$$\left| \mathbb{P}\left(\frac{f_\ell(\Pi_t^K) - \mathbb{E} f_\ell(\Pi_t^K)}{\sqrt{\mathrm{Var} f_\ell(\Pi_t^K)}} \le x \right) - \Phi(x) \right| \le c(K)\, \varepsilon(t).$$

A surprising recent result is due to Pardon [79, 80] who proved in the Euclidean plane a CLT for the volume of Π_t^K for *all* convex bodies K without any restriction on the boundary structure of K. A similar general result in higher dimensions seems to be out of reach at the moment.

References

1. Alishahi, K., Sharifitabar, M.: Volume degeneracy of the typical cell and the chord length distribution for Poisson–Voronoi tessellations in high dimensions. Adv. Appl. Probab. **40**, 919–938 (2008)
2. Bárány, I.: Random polytopes in smooth convex bodies. Mathematika **39**, 81–92 (1992)
3. Bárány, I., Buchta, C.: Random polytopes in a convex polytope, independence of shape, and concentration of vertices. Math. Ann. **297**, 467–497 (1993)
4. Bárány, I., Reitzner, M.: Poisson polytopes. Ann. Probab. **38**, 1507–1531 (2008)
5. Bárány, I., Reitzner, M.: On the variance of random polytopes. Adv. Math. **225**, 1986–2001 (2010)
6. Bárány, I., Fodor, F., Vígh, V.: Intrinsic volumes of inscribed random polytopes in smooth convex bodies. Adv. Appl. Probab. **42**, 605–619 (2010)

7. Baumstark, V., Last, G.: Some distributional results for Poisson–Voronoi tessellations. Adv. Appl. Probab. **39**, 16–40 (2007)
8. Baumstark, V., Last, G.: Gamma distributions for stationary Poisson flat processes. Adv. Appl. Probab. **41**, 911–939 (2009)
9. Beermann, M., Reitzner, M.: Beyond the Efron-Buchta identities: distributional results for Poisson polytopes. Discrete Comput. Geom. **53**, 226–244 (2015)
10. Beermann, M., Redenbach, C., Thäle, C.: Asymptotic shape of small cells. Math. Nachr. **287**, 737–747 (2014)
11. Blaschke, W.: Über affine Geometrie XI: Lösung des "Vierpunktproblems" von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten. Ber. Verh. Sächs. Ges. Wiss. Leipzig Math.-Phys. Kl. **69**, 436–453 (1917); reprinted in: Burau, W., et al. (eds.) Wilhelm Blaschke. Gesammelte Werke. Konvexgeometrie, vol. 3, pp. 284–301. Essen, Thales (1985)
12. Bollobás, B., Riordan, O.: The critical probability for random Voronoi percolation in the plane is 1/2. Probab. Theory Relat. Fields **136**, 417–468 (2006)
13. Bollobás, B., Riordan, O.: Percolation. Cambridge University Press, New York (2006)
14. Böröczky, K., Hug, D.: Stability of the reverse Blaschke–Santaló inequality for zonoids and applications. Adv. Appl. Math. **44**, 309–328 (2010)
15. Böröczky, K., Hoffmann, L.M., Hug, D.: Expectation of intrinsic volumes of random polytopes. Period. Math. Hung. **57**, 143–164 (2008)
16. Bourguin, S., Peccati, G.: Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering. Electron. J. Probab. **19**, 42 (2014)
17. Bourguin, S., Peccati, G.: The Malliavin-Stein method on the Poisson space. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 185–228. Springer, Cham (2016)
18. Buchta, C.: Stochastische Approximation Konvexer Polygone. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **67**, 283–304 (1984)
19. Cabo, A.J., Groeneboom, P.: Limit theorems for functionals of convex hulls. Probab. Theory Relat. Fields **100**, 31–55 (1994)
20. Calka, P.: The distributions of the smallest disks containing the Poisson Voronoi typical cell and the Crofton cell in the plane. Adv. Appl. Probab. **34**, 702–717 (2002)
21. Calka, P.: Tessellations. In: Kendall, W.S., Molchanov, I. (eds.) New Perspectives in Stochastic Geometry, pp. 145–169. Oxford University Press, London (2010)
22. Calka, P., Chenavier, N.: Extreme values for characteristic radii of a Poisson–Voronoi tessellation. Extremes **17**, 359–385 (2014)
23. Calka, P., Schreiber, T.: Limit theorems for the typical Poisson–Voronoi cell and the Crofton cell with a large inradius. Ann. Probab. **33**, 1625–1642 (2005)
24. Calka, P., Schreiber, T.: Large deviation probabilities for the number of vertices of random polytopes in the ball. Adv. Appl. Probab. **38**, 47–58 (2006)
25. Calka, P., Yukich, J.E.: Variance asymptotics for random polytopes in smooth convex bodies. Probab. Theory Relat. Fields **158**, 435–463 (2014)
26. Calka, P., Schreiber, T., Yukich, J.E.: Brownian limits, local limits and variance asymptotics for convex hulls in the ball. Ann. Probab. **41**, 50–108 (2013)
27. Chenavier, N.: A general study of extremes of stationary tessellations with applications. Stoch. Process. Appl. **124**, 2917–2953 (2014)
28. Dalla, L., Larman, D.G.: Volumes of a random polytope in a convex set. In: Gritzmann, P., Sturmfels, B. (eds.) Applied Geometry and Discrete Mathematics. The Victor Klee Festschrift. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 4, pp. 175–180. American Mathematical Society, Providence (1991)
29. Decreusefond, L., Ferraz, E., Randriam, H., Vergne, A.: Simplicial homology of random configurations. Adv. Appl. Probab. **46**, 1–20 (2014)
30. Devillers, O., Glisse, M., Goaoc, X., Moroz, G., Reitzner, M.: The monotonicity of f-vectors of random polytopes. Electron. Commun. Probab. **18**, 1–8 (2013)
31. Efron, B.: The convex hull of a random set of points. Biometrika **52**, 331–343 (1965)

32. Einmahl, J.H.J., Khmaladze, E.V.: The two-sample problem in $\mathbb{R}^m$ and measure-valued martingales. In: State of the Art in Probability and Statistics (Leiden, 1999). IMS Lecture Notes Monograph Series, vol. 36, pp. 434–463. Institute of Mathematical Statistics, Beachwood (2001)
33. Giannopoulos, A.: On the mean value of the area of a random polygon in a plane convex body. Mathematika **39**, 279–290 (1992)
34. Goldmann, A.: Sur une conjecture de D.G. Kendall concernant la cellule do Crofton du plan et sur sa contrepartie brownienne. Ann. Probab. **26**, 1727–1750 (1998)
35. Graf, S., Luschgy, H.: Foundations of Quantization for Probability Distributions. Lecture Notes in Mathematics, vol. 1730. Springer, Berlin (2000)
36. Groemer, H.: On some mean values associated with a randomly selected simplex in a convex set. Pac. J. Math. **45**, 525–533 (1973)
37. Groemer, H.: On the mean value of the volume of a random polytope in a convex set. Arch. Math. **25**, 86–90 (1974)
38. Groemer, H.: On the extension of additive functionals on classes of convex sets. Pac. J. Math. **75**, 397–410 (1978)
39. Groeneboom, P.: Limit theorems for convex hulls. Probab. Theory Relat. Fields **79**, 327–368 (1988)
40. Haenggi, M.: Stochastic Geometry for Wireless Networks. Cambridge University Press, Cambridge (2013)
41. Heinrich, L.: Large deviations of the empirical volume fraction for stationary Poisson grain models. Ann. Appl. Probab. **15**, 392–420 (2005)
42. Heinrich, L., Molchanov, I.S.: Central limit theorems for a class of random measures associated with germ-grain models. Adv. Appl. Probab. **31**, 283–314 (1999)
43. Heinrich, L., Schmidt, H., Schmidt, V.: Limit theorems for stationary tessellations with random inner cell structure. Adv. Appl. Probab. **37**, 25–47 (2005)
44. Heveling, M., Reitzner, M.: Poisson–Voronoi Approximation. Ann. Appl. Probab. **19**, 719–736 (2009)
45. Hörrmann, J., Hug, D.: On the volume of the zero cell of a class of isotropic Poisson hyperplane tessellations. Adv. Appl. Probab. **46**, 622–642 (2014)
46. Hörrmann, J., Hug, D., Reitzner, M., Thähle, C.: Poisson polyhedra in high dimensions. Adv. Math. **281**, 1–39 (2015)
47. Hsing, T.: On the asymptotic distribution of the area outside a random convex hull in a disk. Ann. Appl. Probab. **4**, 478–493 (1994)
48. Hug, D.: Random mosaics. In: Baddeley, A., Bárány, I., Schneider, R., Weil, W. (eds.) Stochastic Geometry. Edited by W. Weil: Lecture Notes in Mathematics, vol. 1892, pp. 247–266. Springer, Berlin (2007)
49. Hug, D.: Random polytopes. In: Spodarev, E. (ed.) Stochastic Geometry, Spatial Statistics and Random Fields. Lecture Notes in Mathematics, vol. 2068, pp. 205–238. Springer, Heidelberg (2013)
50. Hug, D., Schneider, R.: Large typical cells in Poisson–Delaunay mosaics. Rev. Roum. Math. Pures Appl. **50**, 657–670 (2005)
51. Hug, D., Schneider, R.: Asymptotic shapes of large cells in random tessellations. Geom. Funct. Anal. **17**, 156–191 (2007)
52. Hug, D., Schneider, R.: Typical cells in Poisson hyperplane tessellations. Discrete Comput. Geom. **38**, 305–319 (2007)
53. Hug, D., Schneider, R.: Large faces in Poisson hyperplane mosaics. Ann. Probab. **38**, 1320–1344 (2010)
54. Hug, D., Schneider, R.: Faces with given directions in anisotropic Poisson hyperplane mosaics. Adv. Appl. Probab. **43**, 308–321 (2011)
55. Hug, D., Schneider, R.: Approximation properties of random polytopes associated with Poisson hyperplane processes. Adv. Appl. Probab. **46**, 919–936 (2014)
56. Hug, D., Reitzner, M., Schneider, R.: The limit shape of the zero cell in a stationary Poisson hyperplane tessellation. Ann. Probab. **32**, 1140–1167 (2004)

57. Hug, D., Reitzner, M., Schneider, R.: Large Poisson–Voronoi cells and Crofton cells. Adv. Appl. Probab. **36**, 667–690 (2004)
58. Hug, D., Last, G., Schulte, M.: Second order properties and central limit theorems for geometric functionals of Boolean models. Ann. Appl. Probab. **26**, 73–135 (2016)
59. Hug, D., Thäle, C., Weil, W.: Intersection and proximity of processes of flats. J. Math. Anal. Appl. **426**, 1–42 (2015)
60. Kahle, M.: Random geometric complexes. Discrete Comput. Geom. **45**, 553–573 (2011)
61. Kahle, M.: Topology of random simplicial complexes: a survey. Algebraic Topology: Applications and New Directions. Contemporary Mathematics, vol. 620, pp. 201–221. American Mathematical Society, Providence (2014)
62. Kahle, M., Meckes, E.: Limit theorems for Betti numbers of random simplicial complexes. Homol. Homot. Appl. **15**, 343–374 (2013)
63. Khmaladze, E., Toronjadze, N.: On the almost sure coverage property of Voronoi tessellation: the $\mathbb{R}^1$ case. Adv. Appl. Probab. **33**, 756–764 (2001)
64. Kovalenko, I.N.: A proof of a conjecture of David Kendall on the shape of random polygons of large area. (Russian) Kibernet. Sistem. Anal. **187**, 3–10. Engl. transl. Cybernet. Syst. Anal. **33**, 461–467 (1997)
65. Kovalenko, I.N.: A simplified proof of a conjecture of D.G. Kendall concerning shapes of random polygons. J. Appl. Math. Stoch. Anal. **12**, 301–310 (1999)
66. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs. Electron. J. Probab. **18**(Article 32), 1–32 (2013)
67. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, II: rescaled kernels, marked processes and geometric U-statistics. Stoch. Process. Appl. **123**, 4186–4218 (2013)
68. Lachièze-Rey, R., Peccati, G.: New Kolmogorov bounds for functionals of binomial point processes. arXiv:1505.04640 (2015)
69. Lachièze-Rey, R., Reitzner, M.: U-statistics in stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 229–253. Springer, Cham (2016)
70. Lachièze-Rey, R., Vega, S.: Boundary density and Voronoi set estimation for irregular sets. arXiv:1501.04724 (2015)
71. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi and Springer Series, vol. 6, pp. 1–36. Springer/Bocconi University Press, Milan (2016)
72. Last, G., Ochsenreither, E.: Percolation on stationary tessellations: models, mean values and second order structure. J. Appl. Probab. **51A**, 311–332 (2014). An earlier and more comprehensive version is available via arXiv:1312.6366
73. Last, G., Penrose, M.: Fock space representation, chaos expansion and covariance inequalities for general Poisson processes. Probab. Theory Relat. Fields **150**, 663–690 (2011)
74. Last, G., Penrose, M., Schulte, M., Thäle, C.: Central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**, 348–364 (2014)
75. Miles, R.E.: A heuristic proof of a conjecture of D.G. Kendall concerning shapes of random polygons. J. Appl. Math. Stoch. Anal. **12**, 301–310 (1995)
76. Milman, V.D., Pajor, A.: Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. In: Geometric Aspects of Functional Analysis (1987–1988). Lecture Notes in Mathematics, vol. 1376, pp. 64–104. Springer, Berlin (1989)
77. Neher, R.A., Mecke, K., Wagner, H.: Topological estimation of percolation thresholds. J. Stat. Mech. Theory Exp. P01011 (2008)
78. Ochsenreither, E.: Geometrische Kenngrößen und zentrale Grenzwertsätze stetiger Perkolationsmodelle. Ph.D. thesis, Karlsruhe (2014)

79. Pardon, J.: Central limit theorems for random polygons in an arbitrary convex set. Ann. Probab. **39**, 881–903 (2001)
80. Pardon, J.: Central limit theorems for uniform model random polygons. J. Theor. Probab. **25**, 823–833 (2012)
81. Peccati, G., Zheng, C.: Multi-dimensional Gaussian fluctuations on the Poisson space. Electron. J. Probab. **15**, 1487–1527 (2010)
82. Peccati, G., Solé, J.L., Taqqu, M.S., Utzet, F.: Stein's method and normal approximation of Poisson functionals. Ann. Prob. **38**, 443–478 (2010)
83. Penrose, M.D.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
84. Penrose, M.D.: Laws of large numbers in stochastic geometry with statistical applications. Bernoulli **13**, 1124–1150 (2007)
85. Rademacher, L.: On the monotonicity of the expected volume of a random simplex. Mathematika **58**, 77–91 (2013)
86. Reitzner, M.: Random polytopes and the Efron–Stein jackknife inequality. Ann. Probab. **31**, 2136–2166 (2003)
87. Reitzner, M.: Stochastical approximation of smooth convex bodies. Mathematika **51**, 11–29 (2004)
88. Reitzner, M.: Central limit theorems for random polytopes. Probab. Theory Relat. Fields **133**, 483–507 (2005)
89. Reitzner, M.: The combinatorial structure of random polytopes. Adv. Math. **191**, 178–208 (2005)
90. Reitzner, M.: Random polytopes. In: Kendall W.S., Molchanov, I. (eds.) New Perspectives in Stochastic Geometry, pp. 45–76. Oxford University Press, Oxford (2010)
91. Reitzner, M., Schulte, M.: Central limit theorems for U-statistics of Poisson point processes. Ann. Prob. **41**(6), 3879–3909 (2013)
92. Reitzner, M., Spodarev, E., Zaporozhets, D.: Set reconstruction by Voronoi cells. Adv. Appl. Probab. **44**, 938–953 (2012)
93. Reitzner, M., Schulte, M., Thäle, C.: Limit theory for the Gilbert graph. arXiv:1312.4861 (2014)
94. Rényi, A., Sulanke, R.: Über die konvexe Hülle von *n* zufällig gewählten Punkten. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **2**, 75–84 (1963)
95. Rényi, A., Sulanke, R.: Über die konvexe Hülle von *n* zufällig gewählten Punkten II. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **3**, 138–147 (1964)
96. Schneider, R.: Approximation of convex bodies by random polytopes. Aequationes Math. **32**, 304–310 (1987)
97. Schneider, R.: A duality for Poisson flats. Adv. Appl. Probab. **31**, 63–68 (1999)
98. Schneider, R.: Weighted faces of Poisson hyperplane tessellations. Adv. Appl. Probab. **41**, 682–694 (2009)
99. Schneider, R.: Vertex numbers of weighted faces in Poisson hyperplane mosaics. Discrete Comput. Geom. **44**, 599–907 (2010)
100. Schneider, R.: Extremal properties of random mosaics. In: Bárány, I., Böröczky, K.J., Fejes Tóth, G., Pach, J. (eds.) Geometry: Intuitive, Discrete, and Convex. A Tribute to László Fejes Tóth. Bolyai Society Mathematical Studies, vol. 24, pp. 301–330. Springer, Berlin (2013)
101. Schneider, R.: Convex Bodies: The Brunn-Minkowski Theory, 2nd expanded edn. Cambridge University Press, New York (2014)
102. Schneider, R.: Second moments related to Poisson hyperplane tessellations. Manuscript (2015)
103. Schneider, R.: Discrete aspects of stochastic geometry. In: Goodman, J.E., O'Rourke, J. (eds.) Handbook of Discrete and Computational Geometry. CRC Press Series on Discrete Mathematics and its Applications, pp. 167–184. CRC, Boca Raton (2015)
104. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Probability and its Applications (New York). Springer, Berlin (2008)
105. Schneider, R., Weil, W.: Classical stochastic geometry. In: Kendall, W.S., Molchanov, I. (eds.) New Perspectives in Stochastic Geometry, pp. 1–42. Oxford University Press, Oxford (2010)

106. Schneider, R., Wieacker, J.A.: Random polytopes in a convex body. Z. Wahrscheinlichkeitstheor. Verwandte Geb. **52**, 69–73 (1980)
107. Schulte, M.: A central limit theorem for the Poisson-Voronoi approximation. Adv. Appl. Math. **49**, 285–306 (2012)
108. Schulte, M., Thäle, C.: The scaling limit of Poisson-driven order statistics with applications in geometric probability. Stoch. Process. Appl. **122**, 4096–4120 (2012)
109. Schulte, M., Thäle, C.: Distances between Poisson k-flats. Methodol. Comput. Appl. Probab. **16**, 311–329 (2013)
110. Schütt, C.: Random polytopes and affine surface area. Math. Nachr. **170**, 227–249 (1994)
111. Thäle, C., Yukich, J.E.: Asymptotic theory for statistics of the Poisson–Voronoi approximation. Bernoulli. J. Math. Anal. Appl. **434**, 1365–1375 (2016)
112. Vu, V.H.: Sharp concentration of random polytopes. Geom. Funct. Anal. **15**, 1284–1318 (2005)
113. Wieacker, J.A.: Einige Probleme der polyedrischen Approximation. Diplomarbeit, Freiburg im Breisgau (1978)
114. Yukich, J.E.: Surface order scaling in stochastic geometry. Ann. Appl. Probab. **25**, 177–210 (2015)
115. Zuyev, S.: Stochastic geometry and telecommunications networks. In: Kendall, W.S., Molchanov, I. (eds.) New Perspectives in Stochastic Geometry, pp. 520–554, Oxford university Press, Oxford (2010)

The Malliavin–Stein Method on the Poisson Space

Solesne Bourguin and Giovanni Peccati

Abstract This chapter provides a detailed and unified discussion of a collection of recently introduced techniques, allowing one to establish limit theorems with explicit rates of convergence, by combining the Stein's and Chen–Stein methods with Malliavin calculus. Some results concerning multiple integrals are discussed in detail.

1 Introduction

The aim of this chapter is to show that the tools of stochastic analysis developed in the previous parts of the book (see [19, 30]) may be combined very naturally with two powerful probabilistic techniques, namely the *Stein's* and *Chen–Stein methods* for probabilistic approximations. Several remarkable applications of the content of the present chapter in a geometric context are presented in [18, 32].

2 The Stein's and Chen–Stein Methods

The Stein's and Chen–Stein methods can be roughly described as collections of techniques, allowing one to use differential operators in order to explicitly assess the distance between probability distributions. In general, these techniques are applied whenever one wants to compare a known "target" distribution (Gaussian, Poisson, Gamma, Binomial distributions among others) with an unknown one—that

S. Bourguin (✉)
Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA 02215, USA
e-mail: solesne.bourguin@gmail.com

G. Peccati
Unité de Recherche en Mathématiques, Université du Luxembourg, 6 rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Luxembourg
e-mail: giovanni.peccati@gmail.com

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_6

is typically not amenable to direct analysis. As it is to be expected, the nature of the method changes mostly according to the structure of the target distribution.

In this section, Stein's original method along with two of its variants are presented, namely the (original) one-dimensional Stein's method for normal approximations, the one-dimensional Chen–Stein method for Poisson approximations, and finally, a multidimensional version of Stein's method for normal approximations. In what follows, we shall assume that every random element is defined on an adequate probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

2.1 Distances Between Distributions

A crucial notion that will be needed throughout this chapter is that of a *distance* between two probability distributions. Recall that a class $\mathscr{H}$ of real-valued functions on $\mathbb{R}^d$ is said to be *separating* if the following implication holds: if F, G are two $\mathbb{R}^d$-valued random elements such that $h(G), h(F) \in L^1(\Omega)$ and $\mathbb{E}[h(G)] = \mathbb{E}[h(F)]$ for every $h \in \mathscr{H}$, then F and G have the same distribution. Also, we shall say (as usual!) that a sequence of $\mathbb{R}^d$-valued random variables $\{F_n : n \geq 1\}$ converges in distribution (or in law) to F if, for every $h : \mathbb{R}^d \to \mathbb{R}$ bounded and continuous,

$$\mathbb{E}[h(F_n)] \to \mathbb{E}[h(F)], \quad \text{as} \quad n \to \infty.$$

Definition 1 (Distance Between Probability Distributions) Given a separating class of real-valued functions $\mathscr{H}$ on $\mathbb{R}^d$, the distance $d_{\mathscr{H}}(F, G)$ between the laws of two $\mathbb{R}^d$-valued random elements F and G—verifying $h(F), h(G) \in L^1(\Omega)$ for every $h \in \mathscr{H}$—is defined as

$$d_{\mathscr{H}}(F, G) = \sup\left\{\left|\mathbb{E}\left[h(F)\right] - \mathbb{E}\left[h(G)\right]\right| : h \in \mathscr{H}\right\}. \tag{1}$$

It is easily checked that the mapping $d_{\mathscr{H}}(\cdot,\cdot)$ verifies the usual axioms of a distance (or metric) on the class of all probability distributions π on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |h(x)| \mathrm{d}\pi(x) < \infty$ for every $h \in \mathscr{H}$. We will now present several specific distances that will be used throughout the chapter. The reader is referred, e.g., to [11, Chap. 11] or [21, Appendix C] (and the references therein) for any unexplained definition or result concerning the topological properties of the class of probability distributions on a metric space.

Definition 2 Fix $d \geq 1$, and write $\mathscr{B}(\mathbb{R}^d)$ to indicate the corresponding Borel σ-field.

1. The *total variation distance* between the laws of two $\mathbb{R}^d$-valued random variables F and G, denoted by $d_{\mathrm{TV}}(F, G)$, is obtained from (1) by taking $\mathscr{H}$ to be the set of all functions $h\colon \mathbb{R}^d \to \mathbb{R}$ of the type $h(x) = \mathbb{1}_B(x)$, where $B \in \mathscr{B}(\mathbb{R}^d)$. This

class of functions will be denoted by $\mathscr{H}_{\mathrm{TV}}$ in the sequel, in such a way that

$$d_{\mathrm{TV}}(F,G) = d_{\mathscr{H}_{\mathrm{TV}}}(F,G) = \sup_{B\in\mathscr{B}(\mathbb{R}^d)} |\mathbb{P}(F\in B) - \mathbb{P}(G\in B)|\,.$$

2. The *Kolmogorov distance* between the laws of two random variables F and G, denoted by $d_K(F,G)$, is obtained from (1) by taking $\mathscr{H}$ to be the class of all functions $h\colon \mathbb{R}^d \to \mathbb{R}$ of the type $h(x_1,\dots,x_d) = \mathbb{1}_{(-\infty,z_1]\times\cdots\times(-\infty,z_d]}(x_1,\dots,x_d)$, where $z_1,\dots,z_d \in \mathbb{R}$. This class of functions will be denoted by $\mathscr{H}_K$ in the sequel. In particular,

$$\begin{aligned} d_K(F,G) &= d_{\mathscr{H}_K}(F,G) \\ &= \sup_{z_1,\dots,z_d\in\mathbb{R}} |\mathbb{P}(F\in(-\infty,z_1]\times\cdots\times(-\infty,z_d]) \\ &\qquad -\mathbb{P}(G\in(-\infty,z_1]\times\cdots\times(-\infty,z_d])|\,. \end{aligned}$$

3. Let F,G be two $\mathbb{R}^d$-valued random elements such that $\mathbb{E}[\|F\|_{\mathbb{R}^d}], \mathbb{E}[\|G\|_{\mathbb{R}^d}] < \infty$. The *Wasserstein distance* between the laws of F and G, denoted by $d_W(F,G)$, is obtained from (1) by taking $\mathscr{H}$ to be the set of all functions $h\colon\mathbb{R}^d\to\mathbb{R}$ such that $\|h\|_{\mathrm{Lip}} \le 1$, where

$$\|h\|_{\mathrm{Lip}} = \sup_{x,y\in\mathbb{R}^d : x\neq y} \frac{|h(x)-h(y)|}{\|x-y\|_{\mathbb{R}^d}}.$$

This class of functions will be denoted by $\mathscr{H}_W$ in the sequel.

4. Let F,G be two $\mathbb{R}^d$-valued random elements such that $\mathbb{E}[\|F\|_{\mathbb{R}^d}], \mathbb{E}[\|G\|_{\mathbb{R}^d}] < \infty$. The distance $d_2(F,G)$ between the laws of F and G is obtained from (1) by taking $\mathscr{H}$ to be the set of all functions $h\colon\mathbb{R}^d\to\mathbb{R}$, such that $h\in\mathscr{C}^1$, $\|h\|_{\mathrm{Lip}}\le 1$ and

$$M_2(h) := \sup_{x,y\in\mathbb{R}^d : x\neq y} \frac{\|\nabla h(x) - \nabla h(y)\|_{\mathbb{R}^d}}{\|x-y\|_{\mathbb{R}^d}} \le 1.$$

This class of functions is denoted by $\mathscr{H}_2$. Note that, if $h\in\mathscr{C}^2$, then $M_2(h) = \sup_{x\in\mathbb{R}^d} \|\mathrm{Hess}\, h(x)\|_{\mathrm{op}}$, where $\mathrm{Hess}\, h(x)$ stands for the Hessian matrix of h evaluated at x, and the operator norm of a $d\times d$ matrix A is defined as $\|A\|_{\mathrm{op}} = \sup\left\{\|Ax\|_{\mathbb{R}^d} : x\in\mathbb{R}^d,\ \|x\|_{\mathbb{R}^d} = 1\right\}$.

5. Let F,G be two $\mathbb{R}^d$-valued random elements such that $\mathbb{E}[\|F\|^2_{\mathbb{R}^d}], \mathbb{E}[\|G\|^2_{\mathbb{R}^d}] < \infty$. The distance $d_3(F,G)$ between the laws of F and G is obtained from (1) by taking $\mathscr{H}$ to be the set of all functions $h\colon\mathbb{R}^d\to\mathbb{R}$, such that h is three times differentiable and all partial derivatives of order 2 and 3 are bounded by 1. This class of functions is denoted by $\mathscr{H}_3$.

Observe that, in the notation introduced above, there is no explicit dependence on the dimension d: indeed, it will always be the case that the exact value of d is clear from the context. As discussed below, the distances d_2 and d_3 are used mainly in a multidimensional setting. We observe the following basic facts:

- The five classes $\mathscr{H}_{\mathrm{TV}}$, $\mathscr{H}_K$, $\mathscr{H}_W$, $\mathscr{H}_2$, $\mathscr{H}_3$ are all separating, and the topologies induced by the corresponding distances are all strictly stronger than the one induced by the convergence in distribution. In particular, if $d(F_n, F) \to 0$ (where d stands for any of the distances $d_{\mathrm{TV}}, d_K, d_W, d_2$ or d_3), then F_n converges in distribution to F (observe that the converse implication might fail).
- One has that $d_{\mathrm{TV}} \geq d_K$ and $d_W \geq d_2$. Moreover, if $d = 1$, $N \sim \mathscr{N}(0,1)$ and F is any random variable in $L^1(\Omega)$, then $d_K(F,N) \leq 2\sqrt{d_W(F,N)}$.
- If $d = 1$ and the mapping $z \mapsto \mathbb{P}(F \leq z)$ is continuous for every $z \in \mathbb{R}$, then, as $n \to \infty$, F_n converges in distribution to F if and only if $d_K(F_n, F) \to 0$.
- The total variation distance also has the following useful equivalent representation:

$$d_{\mathrm{TV}}(F,G) = \frac{1}{2} \sup \{ |\mathbb{E}[h(F)] - \mathbb{E}[h(G)]| : \|h\|_\infty \leq 1 \}.$$

2.2 The One-Dimensional Stein's Method for Normal Approximations

We will say that a random variable N has the standard Gaussian $\mathscr{N}(0,1)$ distribution (in symbols: $N \sim \mathscr{N}(0,1)$) if the law of N is given by the measure $\mathrm{d}\gamma(x) = (2\pi)^{-1/2} e^{-x^2/2} \mathrm{d}x$. More generally, we shall use the symbol $\mathscr{N}(m, \sigma^2)$ to indicate the one-dimensional Gaussian distribution with mean m and variance σ^2, that is, $Y \sim \mathscr{N}(m, \sigma^2)$ if and only if Y has the same distribution as $m + \sigma N$, where $N \sim \mathscr{N}(0,1)$. The starting point of Stein's method is the following result, universally known as "Stein's Lemma," which provides a useful characterization of the measure γ.

Lemma 1 (Stein's Lemma) *Let N be a real-valued random variable. Then, $N \sim \mathscr{N}(0,1)$ if and only if, for every differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ such that $f' \in L^1(\gamma)$, the expectations $\mathbb{E}[Nf(N)]$ and $\mathbb{E}[f'(N)]$ are finite and*

$$\mathbb{E}[Nf(N)] = \mathbb{E}\left[f'(N)\right].$$

A proof of this elementary statement can be found, e.g., in [21, Proof of Lemma 3.1.2]. Now assume that F is some random variable such that the quantity

$$\mathbb{E}\left[Ff(F) - f'(F)\right]$$

is close to zero for a large class of smooth functions f: is it possible to conclude that the law of F is close to the $\mathscr{N}(0,1)$ distribution in some meaningful probabilistic sense? Somehow surprisingly, such a question admits a positive and rigorous answer, that one can formulate by means of the crucial concept of the *Stein's equation* associated with a given test function h.

Definition 3 (Stein's Equations) Let $N \sim \mathcal{N}(0,1)$ and let $h: \mathbb{R} \to \mathbb{R}$ be a Borel function such that $\mathbb{E}\,|h(N)| < \infty$. The *Stein's equation* associated with h is the ordinary differential equation

$$f'(x) - xf(x) = h(x) - \mathbb{E}\,[h(N)]\,, \quad x \in \mathbb{R}. \tag{2}$$

A solution to the equation (2) is a function f that is absolutely continuous (on compact intervals) and such that there exists a version of the derivative f' verifying (2) for every $x \in \mathbb{R}$.

Elementary considerations show that every solution to (2) necessarily has the form

$$\begin{aligned} f(x) &= ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} \{h(t) - \mathbb{E}[h(N)]\} e^{-t^2/2} \mathrm{d}t \\ &= ce^{x^2/2} - e^{x^2/2} \int_{x}^{\infty} \{h(t) - \mathbb{E}[h(N)]\} e^{-t^2/2} \mathrm{d}t, \end{aligned} \tag{3}$$

where $c \in \mathbb{R}$. In what follows, we shall denote by f_h the function obtained from (3) by setting $c = 0$, that is, we write

$$f_h(x) := e^{x^2/2} \int_{-\infty}^{x} \{h(t) - \mathbb{E}[h(N)]\} e^{-t^2/2} \mathrm{d}t, \quad x \in \mathbb{R}, \tag{4}$$

in such a way that f_h is the only solution to the Stein's equation (2) verifying the asymptotic relation $\lim_{x \to \pm\infty} e^{-x^2/2} f_h(x) = 0$. One should note that, in general, the function f_h defined in (4) might be only almost everywhere differentiable: from now on, we stipulate that the symbol f_h' indicates the version of the derivative of f_h given by

$$f_h'(x) = xf_h(x) + h(x) - \mathbb{E}[h(N)], \quad x \in \mathbb{R}. \tag{5}$$

Stein's equations provide the perfect tool for bridging the gap between the differential characterization of the Gaussian distribution given in Lemma 1, and the notion of distance introduced in Definition 1. Consider indeed a generic random variable F, as well as $N \sim \mathcal{N}(0,1)$. Select a function $h: \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}\,|h(N)| < \infty$ and $\mathbb{E}\,|h(F)| < \infty$, and let the function f_h be defined as in (4). Taking expectations (with respect to the law of F) on both sides of (2) yields

$$\mathbb{E}\,[h(F)] - \mathbb{E}\,[h(N)] = \mathbb{E}\left[f_h'(F) - Ff_h(F)\right].$$

In particular, if $\mathscr{H}$ is a separating class of functions such that $\mathbb{E}\,|h(N)| < \infty$ and $\mathbb{E}\,|h(F)| < \infty$ for every $h \in \mathscr{H}$, one infers that

$$d_{\mathscr{H}}(F,N) = \sup_{h\in\mathscr{H}} \left|\mathbb{E}\left[f_h'(F) - Ff_h(F)\right]\right|. \tag{6}$$

Note that the right-hand side of the previous identity does not involve the target random variable N: indeed, the Gaussian distribution plays a role in such an expression only *via* the characterizing differential operator $f(x) \mapsto f'(x) - xf(x)$ (see Lemma 1). The key point is now that if one chooses a separating class $\mathscr{H}$ whose components verify a uniform bound of some sort (as it happens for the sets $\mathscr{H}_{\mathrm{TV}}$, $\mathscr{H}_K$ and $\mathscr{H}_W$ introduced in Definition 2), then the elements of the class $\{f_h : h \in \mathscr{H}\}$ will also satisfy some uniform estimates, that one can put into use for assessing the right-hand side of (6). We will see that, in general, the mapping f_h associated with a given test function h possesses some additional degree of smoothness that makes the supremum in (6) quite amenable to analysis. Depending on the class $\mathscr{H}$ specifying the distance $d_{\mathscr{H}}$, the properties of the functions f_h significantly change, and the bounds that can be derived from (6) differ accordingly.

In the next three sections, we will discuss in some detail the (one-dimensional) bounds that one can deduce when working with $\mathscr{H} = \mathscr{H}_{\mathrm{TV}}$, $\mathscr{H}_K$ and $\mathscr{H}_W$. Note that the bounds in the total variation distance are difficult to exploit in a Poisson context (mainly because one is naturally led to deal with discrete random variables, whose total variation distance from the normal distribution is equal by definition to the maximal value of 1); however, we decided to present them for the sake of completeness.

2.2.1 Stein's Bounds for the One-Dimensional Total Variation Distance

The following statement provides some classical bounds on the total variation distance. A proof can be found, e.g., in [21, Proof of Theorem 3.3.1].

Theorem 1 *Let $N \sim \mathscr{N}(0,1)$ and let $h\colon \mathbb{R} \to [0,1]$ be a Borel function. Then, the solution f_h to the Stein's equation (2) is such that*

$$\|f_h\|_\infty \le \sqrt{\frac{\pi}{2}} \quad \textit{and} \quad \|f_h'\|_\infty \le 2. \tag{7}$$

In particular, for any real-valued random variable $F \in L^1(\Omega)$ one has the following bound

$$d_{\mathrm{TV}}(F,N) \le \sup_{f\in\mathscr{F}_{\mathrm{TV}}} \left|\mathbb{E}\left[f'(F)\right] - \mathbb{E}\left[Ff(F)\right]\right|, \tag{8}$$

where

$$\mathscr{F}_{\mathrm{TV}} := \left\{f\colon \|f\|_\infty \le \sqrt{\frac{\pi}{2}},\ \|f'\|_\infty \le 2\right\}. \tag{9}$$

Equations (8) and (9) must be formally interpreted in the following sense: (a) the class $\mathscr{F}_{\mathrm{TV}}$ is composed of all absolutely continuous functions f that are bounded by $\sqrt{\pi/2}$ and such that there exists a version of f' that is bounded by 2, and (b) the supremum on the right-hand side of (8) stands for the quantity $\sup |\mathbb{E}[u(F)] - \mathbb{E}[Ff(F)]|$, where the supremum is taken over all pairs (f, u) such that $f \in \mathscr{F}_{\mathrm{TV}}$ and u is a version of f' bounded by 2.

2.2.2 Stein's Bounds for the One-Dimensional Kolmogorov Distance

For every $z \in \mathbb{R}$, we let f_z denote the function f_h, as defined in (4), solving the Stein's equation (2) associated with the indicator function $h = \mathbb{1}_{(-\infty,z]}$. In this case, the integral in (4) can be explicitly computed, yielding that, for every real x,

$$f_z(x) = \begin{cases} \sqrt{2\pi}e^{x^2/2}\Phi(x)[1-\Phi(z)] & \text{if } x \le z \\ \sqrt{2\pi}e^{x^2/2}\Phi(z)[1-\Phi(x)] & \text{if } x \ge z, \end{cases} \tag{10}$$

where $\Phi(a) := \mathbb{P}(N \le a)$. Note that, according to (10), the function f_z is everywhere differentiable, except for the point $x = z$. According to our convention (5), we shall therefore write f_z' to indicate the version of the derivative of f_z satisfying (2) for every real x, that is: $f_z'(x) = xf_z(x) + \mathbb{1}_{(-\infty,z]}(x) - \Phi(x)$, $x \in \mathbb{R}$.

Theorem 2 *Let $z \in \mathbb{R}$ and $N \sim \mathscr{N}(0,1)$. The function f_z is such that*

$$\|f_z\|_\infty \le \frac{\sqrt{2\pi}}{4}, \quad \|f_z'\|_\infty \le 1. \tag{11}$$

Moreover, for all $u, v, w \in \mathbb{R}$,

$$|(w+u)f_z(w+u) - (w+v)f_z(w+v)| \le \left(|w| + \frac{\sqrt{2\pi}}{4}\right)(|u|+|v|) \tag{12}$$

and the following local estimate holds for every $x, h \in \mathbb{R}$:

$$\begin{aligned} &\left|f_z(x+h) - f_z(x) - hf_z'(x)\right| \\ &\le \frac{h^2}{2}\left(|x| + \frac{\sqrt{2\pi}}{4}\right) + h\left(\mathbb{1}_{[x,x+h)}(z) - \mathbb{1}_{[x+h,x)}(z)\right) \\ &= \frac{h^2}{2}\left(|x| + \frac{\sqrt{2\pi}}{4}\right) + |h|\left(\mathbb{1}_{[x,x+h)}(z) + \mathbb{1}_{[x+h,x)}(z)\right). \end{aligned} \tag{13}$$

In particular, for any integrable random variable F,

$$d_K(F,N) \le \sup_{f \in \mathscr{F}_K} \left|\mathbb{E}\left[f'(F)\right] - \mathbb{E}\left[Ff(F)\right]\right|, \tag{14}$$

where

$$\mathscr{F}_K := \left\{ f \colon \|f\|_\infty \le \frac{\sqrt{2\pi}}{4},\ \|f'\|_\infty \le 1 \right\}. \tag{15}$$

A proof of the estimates (11) and (12) can be found in [9, Proof of Lemma 2.3] (see also [21, Exercise 3.4.4]). The local bound (13) (for which a complete proof is provided below) can be found in [12, Proof of Theorem 3.1] (where it is used in a different form) and [17, Proposition 3.1]. One should also notice that (13) refines previous findings from [31]. Equations (14) and (15) must be interpreted as follows: (a) the class $\mathscr{F}_K$ is composed of all absolutely continuous functions f that are bounded by $\sqrt{2\pi}/4$, that are differentiable everywhere except for at most a finite number of points, and such that there exists a version of f' that is bounded by 1, and (b) the supremum on the right-hand side of (14) stands for the quantity $\sup |\mathbb{E}[u(F)] - \mathbb{E}[Ff(F)]|$, where the supremum is taken over all pairs (f, u) such that $f \in \mathscr{F}_K$ and u is a version of f' bounded by 2.

Proof (Theorem 2, Estimate (13)*)* Fix $z \in \mathbb{R}$. Observe that for every $x, h \in \mathbb{R}$, we can write

$$f_z(x+h) - f_z(x) - hf_z'(x) = \int_0^h \left(f'(x+t) - f'(x) \right) \mathrm{d}t.$$

As f_z is a solution to the Stein's equation (2), it satisfies, for all $y \in \mathbb{R}$,

$$f'(y) = yf(y) + \mathbb{1}_{\{y \le z\}} - \Phi(z),$$

which yields, for all $x, h \in \mathbb{R}$,

$$\begin{aligned} &f_z(x+h) - f_z(x) - hf_z'(x) \\ &\quad = \int_0^h ((x+t)f(x+t) - xf(x))\,\mathrm{d}t + \int_0^h \left(\mathbb{1}_{\{x+t \le z\}} - \mathbb{1}_{\{x \le z\}} \right) \mathrm{d}t := I_1 + I_2 \end{aligned}$$

and hence, by the triangle inequality,

$$\left| f_z(x+h) - f_z(x) - hf_z'(x) \right| \le |I_1| + |I_2|. \tag{16}$$

Using (12), we have

$$|I_1| \le \int_0^h \left(|x| + \frac{\sqrt{2\pi}}{4} \right) |t| \mathrm{d}t = \frac{h^2}{2} \left(|x| + \frac{\sqrt{2\pi}}{4} \right). \tag{17}$$

Furthermore, observe that

$$
\begin{aligned}
|I_2| &= \mathbb{1}_{\{h<0\}} \left| \int_0^h \left(\mathbb{1}_{\{x+t\le z\}} - \mathbb{1}_{\{x\le z\}} \right) \mathrm{d}t \right| + \mathbb{1}_{\{h\ge 0\}} \left| \int_0^h \left(\mathbb{1}_{\{x+t\le z\}} - \mathbb{1}_{\{x\le z\}} \right) \mathrm{d}t \right| \\
&= \mathbb{1}_{\{h<0\}} \left| - \int_h^0 \mathbb{1}_{\{x+t\le z<x\}} \mathrm{d}t \right| + \mathbb{1}_{\{h\ge 0\}} \left| - \int_0^h \mathbb{1}_{\{x\le z<x+t\}} \mathrm{d}t \right| \\
&= \mathbb{1}_{\{h<0\}} \int_h^0 \mathbb{1}_{\{x+t\le z<x\}} \mathrm{d}t + \mathbb{1}_{\{h\ge 0\}} \int_0^h \mathbb{1}_{\{x\le z<x+t\}} \mathrm{d}t.
\end{aligned}
$$

Bounding t by h in both integrals provides the following upper bound:

$$
\begin{aligned}
|I_2| &\le \mathbb{1}_{\{h<0\}}(-h)\mathbb{1}_{[x+h,x)}(z) + \mathbb{1}_{\{h\ge 0\}} h \mathbb{1}_{[x,x+h)}(z) \\
&\le h \left(\mathbb{1}_{[x,x+h)}(z) - \mathbb{1}_{[x+h,x)}(z) \right) = |h| \left(\mathbb{1}_{[x,x+h)}(z) + \mathbb{1}_{[x+h,x)}(z) \right). \qquad (18)
\end{aligned}
$$

Using estimates (17) and (either side of the equality in) (18) in (16) concludes the proof. □

2.2.3 Stein's Bounds for the Wasserstein Distance

Normal approximations in the Wasserstein distance are dealt with using the following result:

Theorem 3 *Let $h: \mathbb{R} \to \mathbb{R}$ be such that $\|h\|_{\mathrm{Lip}} \le 1$, and let $N \sim \mathscr{N}(0,1)$. Then, the function f_h defined in (4) (solving the Stein's equation (2)) is everywhere continuously differentiable and such that $\|f_h'\|_\infty \le \sqrt{2/\pi}$. Also, the derivative f_h' is almost everywhere differentiable, and there exists a version f_h'' of the derivative of f_h' such that $\|f_h''\|_\infty \le 2$. In particular, for every square-integrable random variable F, one has the bound*

$$
d_W(F, N) \le \sup_{f \in \mathscr{F}_W} \left| \mathbb{E}\left[f'(F) \right] - \mathbb{E}\left[F f(F) \right] \right|, \qquad (19)
$$

where

$$
\mathscr{F}_W := \left\{ f: \mathbb{R} \to \mathbb{R} \in \mathcal{C}^1 : \|f'\|_\infty \le \sqrt{2/\pi},\ \|f''\|_\infty \le 2 \right\}, \qquad (20)
$$

where $\mathcal{C}^1$ indicates the collection of all continuously differentiable functions on $\mathbb{R}$.

A proof of the bound $\|f_h'\|_\infty \le \sqrt{2/\pi}$ can be found in [21, Sect. 3.5], whereas the bound on f_h'' follows, e.g., from [7, Lemma 4.3]. The definition (20) formally indicates that the class $\mathscr{F}_W$ is composed of all $f \in \mathcal{C}^1$ that are bounded by $\sqrt{2/\pi}$, and whose derivative is a Lipschitz function with Lipschitz constant ≤ 2. Note that the supremum on the right-hand side of (19) is unambiguously defined, since the derivative f' exists everywhere for every $f \in \mathscr{F}_W$.

2.3 Multidimensional Stein's Bounds for Normal Approximations

This subsection provides some extensions of the results of Sect. 2.2, allowing one to deal with the normal approximation of d-dimensional random vectors, for $d \geq 2$. As a general rule, one has that the multidimensional Stein's method requires test functions that are smoother than those one can consider in the one-dimensional case. This is due to the fact that the differential operators appearing in the multidimensional Stein's method are second order operators (see [21, Chap. 4] for a full discussion of this point). As a consequence, we will only be able to derive bounds in the Wasserstein distance (that are presented for the sake of completeness, but cannot be directly applied in a Poisson context) and in the distance d_2. The distance d_3 will appear in Sect. 6.2, in connection with interpolation techniques.

Some further (standard) notation is needed. Fix an integer $d \geq 2$, and write $\mathcal{M}_d(\mathbb{R})$ to indicate the collection of all real $d \times d$ matrices. The Hilbert–Schmidt inner product and the Hilbert–Schmidt norm on $\mathcal{M}_d(\mathbb{R})$, denoted respectively by $\langle \cdot, \cdot \rangle_{H.S}$ and $\| \cdot \|_{H.S}$, are defined as follows: for every pair of matrices A and B, $\langle A, B \rangle_{H.S} = \mathrm{Tr}\left(AB^T\right)$ and $\|A\|_{H.S} = \sqrt{\langle A, A \rangle_{H.S}}$, with $\mathrm{Tr}(\cdot)$ the usual trace operator and $\cdot^T$ the usual transposition operator.

The next statement is the exact multidimensional counterpart of Lemma 1. Given $m \in \mathbb{R}^d$ and a $d \times d$ covariance matrix Σ, we shall denote $\mathscr{N}(m, \Sigma)$ the d-dimensional Gaussian distribution with mean m and covariance Σ.

Lemma 2 (Multidimensional Stein's Lemma) *Let $\Sigma = \{\Sigma(i,j) \colon i,j = 1, \ldots, d\}$ be a nonnegative definite $d \times d$ symmetric matrix. Let $N = (N_1, \ldots, N_d)$ be a random vector with values in $\mathbb{R}^d$. Then, $N \sim \mathscr{N}(0, \Sigma)$ if and only if*

$$\mathbb{E}\left[\langle N, \nabla f(N) \rangle_{\mathbb{R}^d}\right] = \mathbb{E}\left[\langle \Sigma, \mathrm{Hess} f(N) \rangle_{H.S}\right],$$

for every $\mathcal{C}^2$ function $f \colon \mathbb{R}^d \to \mathbb{R}$ having bounded first and second derivatives. Here, Hess$f$ *denotes the Hessian matrix of f.*

There are several ways of proving this result: one of the most instructive can be found in [21, Proof of Lemma 4.1.3], as it is based on the same interpolation technique we shall explore in Sect. 6.2. As in the previous subsection, the next step is to define (and solve) an appropriate Stein's equation linking the multidimensional Gaussian characterization stemming from Stein's Lemma and an appropriate notion of distance.

Definition 4 (Multidimensional Stein's Equations) Let $N = (N_1, \ldots, N_d)$ be a centered Gaussian random vector with positive definite covariance matrix Σ, and let $h: \mathbb{R}^d \to \mathbb{R}$ be such that $\mathbb{E}\,|h(N)| < \infty$. The Stein's equation associated with h and N is the partial differential equation

$$\langle \Sigma, \operatorname{Hess} f(x) \rangle_{H.S} - \langle x, \nabla f(x) \rangle_{\mathbb{R}^d} = h(x) - \mathbb{E}\,[h(N)]\,. \tag{21}$$

A solution to the equation (21) is a $\mathcal{C}^2$ function f_h verifying (21) for every $x \in \mathbb{R}$.

It is not difficult to check that, whenever h is Lipschitz, a solution to (21) is given by

$$f_h(x) = \int_0^1 \frac{1}{2t} \mathbb{E}[h(N) - h(\sqrt{t}x + \sqrt{1-t}N]\,\mathrm{d}t, \quad x \in \mathbb{R}^d, \tag{22}$$

see, e.g., [21, Proposition 4.3.2].

2.3.1 Stein's Bounds for the Wasserstein Distance

The following statement allows to deal with normal approximations in the Wasserstein distance. It represents a quantitative version of Lemma 2.

Theorem 4 *Fix $d \geq 2$. Let $h: \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function with constant $K > 0$. Then, the solution f_h to the Stein's equation (21), as defined in (22), is of class $\mathscr{C}^2$ and such that*

$$\sup_{x \in \mathbb{R}^d} \| \operatorname{Hess} f_h(x) \|_{H.S} \leq K \|\Sigma^{-1}\|_{\mathrm{op}} \|\Sigma\|_{\mathrm{op}}^{1/2}. \tag{23}$$

In particular, for N a centered d-dimensional Gaussian vector with covariance matrix Σ, where Σ is a positive definite matrix, and for any square-integrable $\mathbb{R}^d$-valued random vector F,

$$d_W(F, N) \leq \sup_{f \in \mathscr{F}_W^d(\Sigma)} \left| \mathbb{E}\,[\langle \Sigma, \operatorname{Hess} f(x) \rangle_{H.S}] - \mathbb{E}\,[\langle x, \nabla f(x) \rangle_{\mathbb{R}^d}] \right|, \tag{24}$$

where

$$\mathscr{F}_W^d(\Sigma) := \left\{ f: \mathbb{R}^d \to \mathbb{R} \in \mathcal{C}^2 : \sup_{x \in \mathbb{R}^d} \| \operatorname{Hess} f(x) \|_{H.S} \leq \|\Sigma^{-1}\|_{\mathrm{op}} \|\Sigma\|_{\mathrm{op}}^{1/2}. \right\}.$$

Details on how to prove (23) can be found in [21, Sect. 4.3]. The bound on the Wasserstein distance is immediately obtained by taking expectations on both sides of (21) (with respect to the law of F) in the case where h is a 1-Lipschitz function and $f = f_h$, and then by applying the definition of d_W.

2.3.2 Stein's Bounds for the d_2 Distance

Bounds analogous to those in the previous subsection can be deduced in the case of the d_2 distance. A proof can be found in [26, Proof of Lemma 2.17].

Theorem 5 *Fix $d \geq 2$. Let $h \in \mathscr{H}_2$ (see Definition 2–2). Then, the solution f_h to the Stein's equation (21), as defined in (22), is of class $\mathscr{C}^2$ and such that*

$$\sup_{x \in \mathbb{R}^d} \| \operatorname{Hess} f_h(x) \|_{H.S} \leq \| \Sigma^{-1} \|_{\text{op}} \| \Sigma \|_{\text{op}}^{1/2}, \tag{25}$$

and

$$M_3(f_h) := \sup_{x, \in \mathbb{R}^d, x \neq y} \frac{\| \operatorname{Hess} f_h(x) - \operatorname{Hess} f_h(x) \|_{\text{op}}}{\| x - y \|_{\mathbb{R}^d}} \leq \frac{\sqrt{2\pi}}{4} \| \Sigma^{-1} \|_{\text{op}}^{3/2} \| \Sigma \|_{\text{op}}. \tag{26}$$

As a consequence, for $N \sim \mathscr{N}(0, \Sigma)$, where Σ is a positive definite matrix, and for any square-integrable $\mathbb{R}^d$-valued random vector F,

$$d_2(F, N) \leq \sup_{f \in \mathscr{F}_2^d(\Sigma)} \left| \mathbb{E} \left[\langle \Sigma, \operatorname{Hess} f(x) \rangle_{H.S} \right] - \mathbb{E} \left[\langle x, \nabla f(x) \rangle_{\mathbb{R}^d} \right] \right|, \tag{27}$$

where

$$\mathscr{F}_2^d(\Sigma) := \left\{ f \in \mathcal{C}^2 \colon \sup_{x \in \mathbb{R}^d} \| \operatorname{Hess} f(x) \|_{H.S} \leq \| \Sigma^{-1} \|_{\text{op}} \| \Sigma \|_{\text{op}}^{1/2}, \right.$$
$$\left. M_3(f) \leq \frac{\sqrt{2\pi}}{4} \| \Sigma^{-1} \|_{\text{op}}^{3/2} \| \Sigma \|_{\text{op}}. \right\}.$$

As anticipated, the next section deals with some "discrete" variant of Stein's method.

2.4 *The One-Dimensional Chen–Stein Method for Poisson Approximations*

The Chen–Stein method is an analogue of Stein's method in the framework of Poisson approximations. Similar to Stein's original method for normal approximations (see Sect. 2.2), the goal of the Chen–Stein method is to provide *quantitative* *bounds* on the distance (in a certain strong sense) between the law of a random variable with values in $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ and the Poisson distribution with a given parameter $\lambda > 0$, denoted by $\mathrm{Po}(\lambda)$. As usual, we shall say that a random variable X has the Poisson distribution with parameter $\lambda > 0$ (in symbols, $X \sim \mathrm{Po}(\lambda)$), if

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k \in \mathbb{Z}_+.$$

Classic references for Poisson approximations are [4, 13]; see also [8, 28].

Our first elementary remark is that if F and X are two random variables with values in $\mathbb{Z}_+$, then the total variation distance between the laws of F and X (as introduced in Definition 2) can be rewritten as

$$\begin{aligned} d_{\mathrm{TV}}(F, X) &= \sup_{B \in \mathscr{B}(\mathbb{R})} |\mathbb{P}(F \in B) - \mathbb{P}(X \in B)| \\ &= \sup_{B \in \mathscr{B}(\mathbb{R})} |\mathbb{P}(F \in B \cap \mathbb{Z}_+) - \mathbb{P}(X \in B \cap \mathbb{Z}_+)| \\ &= \sup_{A \subset \mathbb{Z}_+} |\mathbb{P}(F \in A) - \mathbb{P}(X \in A)| \,. \end{aligned}$$

The following statement is the Poisson equivalent of Stein's Lemma for the Gaussian distribution.

Lemma 3 (Chen–Stein Lemma) *A random variable W with values in $\mathbb{Z}_+$ has the* $\mathrm{Po}(\lambda)$ *distribution if and only if, for every bounded $f \colon \mathbb{Z}_+ \to \mathbb{R}$,*

$$\mathbb{E}\left[Wf(W) - \lambda f(W + 1)\right] = 0.$$

A proof of this lemma can be found e.g. in [13, Proof of Theorem 2.2] or [24, Proof of Lemma 3.3.3]. As in the case of normal approximations, the Chen–Stein Lemma suggests the following question: assume that W is a random variable with values in $\mathbb{Z}_+$ and such that the quantity

$$\mathbb{E}\left[Wf(W) - \lambda f(W + 1)\right]$$

is close to zero for a large family of functions f; can we conclude that the distribution of W is close to $\mathrm{Po}(\lambda)$? In order to give a rigorous answer to this question, one has to introduce the concept of a *Chen–Stein (difference) equation*.

Definition 5 (Chen–Stein's Equations) Let $Z \sim \mathrm{Po}(\lambda)$, and let $h \colon \mathbb{Z}_+ \to \mathbb{R}$ be such that $\mathbb{E}\left[h(Z)\right] < \infty$. The *Chen–Stein's equation* associated with h is given by

$$\lambda f(k + 1) - kf(k) = h(k) - \mathbb{E}\left[h(Z)\right], \quad k \in \mathbb{Z}_+. \tag{28}$$

Any solution f to (28) necessarily verifies

$$f(k) = \frac{(k-1)!}{\lambda^k} \sum_{i=0}^{k-1} (h(i) - \mathbb{E}[h(Z)]) \frac{\lambda^i}{i!}, \quad k = 1, 2, \ldots, \tag{29}$$

while the value of $f(0)$ can be chosen arbitrarily. In what follows, we shall denote by f_h the unique solution to (29) verifying $f_h(0) = 0$, where the symbol Δ indicates the forward difference operator $\Delta f(k) := f(k+1) - f(k)$, $k = 0, 1, 2, \ldots$, and, for $j \geq 2$, $\Delta^j f := \Delta(\Delta^{j-1} f)$.

The explicit representation (29) is derived, e.g., in [13, Theorem 2.1]. The next statement provides useful bounds on f_h, Δf_h and $\Delta^2 f_h$. Given a function $g\colon \mathbb{Z}_+ \to \mathbb{R}$, we write $\|g\|_\infty = \sup_{i \geq 0} |g(i)|$.

Proposition 1 *Let the above notation prevail, and consider a bounded function $h\colon \mathbb{Z}_+ \to \mathbb{R}$. Then, the function f_h verifies the following estimates:*

$$\|f_h\|_\infty \leq \left(1 \wedge \sqrt{\frac{2}{e\lambda}}\right) \left[\sup_{i \in \mathbb{Z}_+} h(i) - \inf_{i \in \mathbb{Z}_+} h(i)\right]; \tag{30}$$

$$\|\Delta f_h\|_\infty \leq \left(\frac{1 - e^{-\lambda}}{\lambda}\right) \left[\sup_{i \in \mathbb{Z}_+} h(i) - \inf_{i \in \mathbb{Z}_+} h(i)\right], \tag{31}$$

$$\|\Delta^2 f_h\|_\infty \leq \left(\frac{2 - 2e^{-\lambda}}{\lambda}\right) \left[\sup_{i \in \mathbb{Z}_+} h(i) - \inf_{i \in \mathbb{Z}_+} h(i)\right]. \tag{32}$$

The estimates (30) and (31) are standard, see, e.g., [13, Theorem 2.3], whereas (32) is a consequence of the triangle inequality. We also mention a remarkable estimate by Daly [10, Theorem 1.3], according to which any function f verifying (29) also satisfies the relation

$$\sup_{k \geq 1} |\Delta^j f(k)| \leq \frac{2}{\lambda} \sup_{k \geq 0} |\Delta^{j-1} h(k)|,$$

holding for any integer $j \geq 2$. Such higher order estimates are not needed in our analysis.

2.4.1 Chen–Stein Bounds for the Total Variation Distance

Given $A \subset \mathbb{Z}_+$, we denote by f_A the function f_h (as defined in (29), and satisfying the boundary condition $f_h(0) = 0$) associated with the test function $h(k) = \mathbb{1}_A(k)$, $k \in \mathbb{Z}_+$. As explained above, one has that f_A solves the Chen–Stein's equation

$$\lambda f(k+1) - k f(k) = \mathbb{1}_A(k) - \mathbb{P}(Z \in A), \quad k \in \mathbb{Z}_+, \tag{33}$$

where $Z \sim \mathrm{Po}(\lambda)$. Now let W be any random variable with values in $\mathbb{Z}_+$. Taking expectations with respect to the law of W on both sides of (33) yields therefore that

$$d_{\mathrm{TV}}(W, Z) = \sup_{A \subset \mathbb{Z}_+} \left| \mathbb{E}\left[W f_A(W) - \lambda f_A(W+1) \right] \right| . \tag{34}$$

Since, by virtue of (30)–(32),

$$\|f_A\|_\infty \le 1 \wedge \sqrt{\frac{2}{e\lambda}}, \quad \|\Delta f_A\|_\infty \le \frac{1 - e^{-\lambda}}{\lambda}, \quad \|\Delta^2 f_A\|_\infty \le \frac{2 - 2e^{-\lambda}}{\lambda},$$

we immediately deduce the following statement, allowing one to deal with Poisson approximations in the total variation distance:

Theorem 6 *Let* $Z \sim \mathrm{Po}(\lambda)$, $\lambda > 0$, *and let* W *be a random variable with values in* $\mathbb{Z}_+$. *Then,*

$$d_{\mathrm{TV}}(W, Z) \le \sup_{f \in \Psi_{\mathrm{TV}}} \left| \mathbb{E}\left[W f(W) - \lambda f(W+1) \right] \right| , \tag{35}$$

where

$$\Psi_{\mathrm{TV}} := \left\{ f \colon \mathbb{Z}_+ \to \mathbb{R} \colon \|f\|_\infty \le 1 \wedge \sqrt{\frac{2}{e\lambda}},\ \|\Delta f\|_\infty \le \frac{1 - e^{-\lambda}}{\lambda},\ \|\Delta^2 f\|_\infty \le \frac{2 - 2e^{-\lambda}}{\lambda} \right\} .$$

The power of the bound (35) will be demonstrated in Sect. 7 below and, in much more detail, in the survey [32].

3 Relevant Elements of Malliavin Calculus on the Poisson Space

For the rest of the chapter, we shall demonstrate how the previous bounds based on the Stein's and Chen–Stein methods can be combined with the Malliavin operators discussed in [19]. For the convenience of the reader, we shall briefly recall the relevant definitions and results.

We work within the general framework outlined in [19], namely: $(\mathbb{X}, \mathscr{X}, \mu)$ is a σ-finite measure space, and η is a Poisson random measure on $(\mathbb{X}, \mathscr{X})$ with intensity measure μ. To simplify the discussion, for the rest of this survey we assume that the space $(\mathbb{X}, \mathscr{X})$ is such that η is *proper*, in the sense formally explained in [19, Sect. 1.1] (see in particular formula (1.6) therein). For $p \ge 1$, we denote by L^p_η the class of those random variables F such that $\mathbb{E}|F|^p < \infty$ and $F = f(\eta)$, $\mathbb{P}$-a.s., where f is a representative of F. Recall that f is a measurable function on $\mathbf{N}_\sigma$ (the class of all σ-finite measures on $(\mathbb{X}, \mathscr{X})$ taking values in $\mathbb{Z}_+ \cup \{+\infty\}$)—see [19, Sect. 1.2] for more details.

The following objects will appear in the subsequent analysis:

– For every $n \geq 1$ and every function $g \in L_s^2(\mu^n)$, the symbol $I_n(g)$ denotes the multiple Wiener–Itô integral of order n, of g with respect to $\hat{\eta} = \eta - \mu$. We also adopt the usual notational convention: $L_s^2(\mu^0) = \mathbb{R}$, and $I_0(c) = c$, for every $c \in \mathbb{R}$. See [19, Sect. 1.3]. Recall that every $F \in L_\eta^2$ admits a unique chaotic expansion of the type

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} I_n(g_n), \tag{36}$$

where $g_n \in L_s^2(\mu^n)$. See [19, Sect. 1.4].

– The Malliavin derivative operator, denoted by D, transforms random variables into random functions. Formally, the domain of D, written $\operatorname{dom} D$, is the set of those random variables $F \in L_\eta^2$ admitting a chaotic decomposition (36) such that

$$\sum_{n=1}^{\infty} n n! \|g_n\|_{L^2(\mu^n)}^2 < \infty. \tag{37}$$

If $F \in \operatorname{dom} D$, then the random function $z \mapsto D_z F$ is defined as

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}\left(g_n(z, \cdot)\right), \quad z \in \mathbb{X}. \tag{38}$$

By exploiting the isometric properties of multiple integrals, and thanks to (37), one sees that $DF \in L^2(\mathbb{P} \otimes \mu)$, for every $F \in \operatorname{dom} D$. Fix $z \in X$. Given a random variable $G \in L_\eta^2$ with representative $\mathfrak{v}$, we define $G_z = \mathfrak{v}(\eta + \delta_z)$ to be the random variable obtained by adding the Dirac mass δ_z to the argument of $\mathfrak{v}$. Since the representative $\mathfrak{v}$ is $\mathbb{P}$-a.s. uniquely defined, the definition of G_z is $\mathbb{P} \otimes \mu$-a.e. independent of the choice of $\mathfrak{v}$. The following result is proved in [19, Theorem 3].

Lemma 4 *For every $F \in L_\eta^2$, one has that $F \in \operatorname{dom} D$ if and only if the mapping $(\omega, z) \mapsto (F_z - F)(\omega)$ is an element of $L^2(\mathbb{P} \otimes \mu)$. In this case, one has also that, $\mathbb{P}$-a.s,*

$$D_z F = F_z - F, \quad \text{a.e.} - \mu(dz).$$

A consequence of this representation of D is that, if $F, G \in \operatorname{dom} D$ are such that $FG \in \operatorname{dom} D$, then $D(FG) = FDG + GDF + DGDF$.

– Due to the chaotic representation property (36), every random function $u \in L^2(\mathbb{P} \otimes \mu)$ admits a (unique up to negligible sets) representation of the type

$$u_z = \sum_{n=0}^{\infty} I_n\left(g_n(z, \cdot)\right), \quad z \in \mathbb{X}, \tag{39}$$

where, for every z, the kernel $g_n(z, \cdot)$ is an element of $L_s^2(\mu^n)$. The domain of the divergence operator, denoted by $\operatorname{dom}\delta$, is defined as the collection of those $u \in L^2(\mathbb{P} \otimes \mu)$ such that the chaotic expansion (39) verifies the condition

$$\sum_{n=0}^{\infty}(n+1)!\|g_n\|^2_{L^2(\mu^{n+1})} < \infty.$$

If $u \in \operatorname{dom}\delta$, then the random variable $\delta(u)$ is defined as

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}\left(\tilde{g}_n\right),$$

where $\tilde{g}_n$ stands for the canonical symmetrization of g_n (as a function in $n+1$ variables). The following classic result, proved in [19, Theorem 4], provides a characterization of δ as the adjoint of the derivative operator D.

Lemma 5 (Integration by Parts Formula) *For every $G \in \operatorname{dom} D$ and every $u \in \operatorname{dom}\delta$, one has that*

$$\mathbb{E}\left[G\delta(u)\right] = \mathbb{E}[\langle DG, u\rangle_{L^2(\mu)}], \tag{40}$$

where

$$\langle DG, u\rangle_{L^2(\mu)} = \int_{\mathbb{X}} D_z G \times u(z)\, \mu(\mathrm{d}z).$$

– The domain of the Ornstein–Uhlenbeck generator (see [22, Chap. 1]), written $\operatorname{dom} L$, is given by those $F \in L^2_\eta$ such that their chaotic expansion (36) verifies

$$\sum_{n=1}^{\infty} n^2 n!\|g_n\|^2_{L^2(\mu^n)} < \infty.$$

If $F \in \operatorname{dom} L$, then the random variable LF is given by

$$LF = -\sum_{n=1}^{\infty} n I_n\left(g_n\right).$$

Note that $\mathbb{E}[LF] = 0$, by definition. The following result is a direct consequence of the definitions of D, δ, and L—see also [19, Proposition 3]

Lemma 6 *Consider $F \in L^2_\eta$. Then, $F \in \operatorname{dom} L$ if and only if $F \in \operatorname{dom} D$ and $DF \in \operatorname{dom} \delta$. In this case,*

$$\delta DF = -LF. \tag{41}$$

– The domain of the pseudo-inverse L^{-1} of L is the whole space L^2_η. If $F \in L^2_\eta$ and $F = \sum_{n=0}^{\infty} I_n(g_n)$, then

$$L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(g_n). \tag{42}$$

Observe that $LF = L(F - \mathbb{E}[F])$ and $L^{-1}F = L^{-1}(F - \mathbb{E}[F])$; also, $F - \mathbb{E}[F] = LL^{-1}F$. The following elementary result is one of the staples of the analysis to follow:

Lemma 7 *Let $G \in \operatorname{dom} D$ and $F \in L^2_\eta$. Then,*

$$\mathbb{E}[FG] = \mathbb{E}[F]\mathbb{E}[G] + \mathbb{E}[\langle DG, -DL^{-1}F\rangle_{L^2(\mu)}]. \tag{43}$$

Proof Using (41), one deduces that $F = \mathbb{E}[F] + LL^{-1}F = \mathbb{E}[F] + \delta(-L^{-1}F)$. It follows that

$$\mathbb{E}[FG] = \mathbb{E}[F]\mathbb{E}[G] + \mathbb{E}[G\delta(-DL^{-1}F)].$$

The conclusion is achieved by applying (40) in the case $u = -DL^{-1}F$. □

We also recall the following representation: for $F \in L^2_\eta$ as in (36),

$$L^{-1}F = -\int_0^1 s^{-1} P_s F \, \mathrm{d}s, \tag{44}$$

where

$$P_s F = \mathbb{E}[F] + \sum_{n=1}^{\infty} s^n I_n(g_n), \quad s \in [0, 1].$$

See [19, Theorem 7].

4 One-Dimensional Malliavin–Stein Bounds in the Normal Approximation on the Poisson Space

As in Sect. 3, we work within the general framework of a Poisson measure η defined on the measurable space $(\mathbb{X}, \mathscr{X})$. We denote by μ the σ-finite intensity measure of η. We shall now show how one can combine the one-dimensional Stein's method with the operators of Malliavin calculus presented above, in order to study the normal approximation of random variables of the type $F \in \operatorname{dom} D$. As already recalled, the seed of the ideas developed below originated in the paper [27], which was the first one to combine Stein's method and Malliavin operators in the framework of point measures.

4.1 Bounds on the Wasserstein Distance

The following result provides a useful bound on one-dimensional normal approximations in the Wasserstein distance. It corresponds to the main finding of Peccati et al. [27], with the difference that here we work without any topological assumptions on the measure space $(\mathbb{X}, \mathscr{X})$ and only assume that μ is σ-finite (whereas the results of [27] are stated for μ σ-finite and non-atomic).

Theorem 7 *Let $F \in \operatorname{dom} D$ be such that $\mathbb{E}[F] = 0$, and let $N \sim \mathscr{N}(0,1)$. Then,*

$$
\begin{aligned}
&d_W(F,N) \qquad (45)\\
&\le \sqrt{\frac{2}{\pi}}\,\mathbb{E}\left[\left|1 - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)}\right|\right] + \int_{\mathbb{X}} \mathbb{E}\left[|D_zF|^2\left|D_zL^{-1}F\right|\right]\mu(\mathrm{d}z)\\
&\le \sqrt{\frac{2}{\pi}\,\mathbb{E}\left[\left(1 - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)}\right)^2\right]} + \int_{\mathbb{X}} \mathbb{E}\left[|D_zF|^2\left|D_zL^{-1}F\right|\right]\mu(\mathrm{d}z).
\end{aligned}
$$

Proof Let f be an element of the class of functions $\mathscr{F}_W$, as defined in (20). Observe that, for every $a, b \in \mathbb{R}$, one has that

$$|f(a) - f(b)| \le \sqrt{\frac{2}{\pi}}|a-b|, \qquad (46)$$

and also, since there exists a version of f'' that is bounded by 2,

$$|f(a) - f(b) - f'(a)(b-a)| \le (b-a)^2, \qquad (47)$$

which is a consequence of the elementary relation

$$f(b) = f(a) + f'(a)(b-a) + \int_a^b (f'(y) - f'(a))\mathrm{d}y.$$

Relation (46) implies that

$$\int_{\mathbb{X}} \mathbb{E}[(f(F)_z - f(F))^2]\mu(\mathrm{d}z) = \int_{\mathbb{X}} \mathbb{E}[(f(F_z) - f(F))^2]\mu(\mathrm{d}z) \leq \frac{2}{\pi} \int_{\mathbb{X}} \mathbb{E}[(D_z F)^2]\mu(\mathrm{d}z) < \infty,$$

and therefore, according to Lemma 4, $f(F) \in \operatorname{dom} D$. Since F is centered, one can now apply (43) in the case $G = f(F)$ to deduce that

$$\mathbb{E}\,[Ff(F)] = \mathbb{E}\left[\left\langle Df(F), -DL^{-1}F\right\rangle_{L^2(\mu)}\right] = \mathbb{E} \int_{\mathbb{X}} (f(F_z) - f(F))(-D_z L^{-1}F)\mu(\mathrm{d}z).$$

Applying (47) in the case $a = F$ and $b = F_z$ yields

$$\left|\mathbb{E}\,[Ff(F)] - \mathbb{E}\left[f'(F)\left\langle DF, -DL^{-1}F\right\rangle_{L^2(\mu)}\right]\right| \leq \mathbb{E}\left[\left\langle |DF|^2, |DL^{-1}F|\right\rangle\right],$$

where $\left\langle |DF|^2, |DL^{-1}F|\right\rangle$ is shorthand for

$$\int_{\mathbb{X}} |D_z F|^2 |D_z L^{-1} F|\, \mu(\mathrm{d}z).$$

Since $f \in \mathscr{F}_W$, and consequently f' is bounded by $\sqrt{2/\pi}$, we deduce that

$$\left|\mathbb{E}\left[Ff(F) - f'(F)\right]\right| \leq \sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|\left\langle DF, -DL^{-1}F\right\rangle_{L^2(\mu)} - 1\right|\right] + \int_{\mathbb{X}} \mathbb{E}\left[|D_z F|^2 \left|D_z L^{-1} F\right|\right] \mu(\mathrm{d}z),$$

and the first inequality in the statement is immediately deduced from the bound (19). The second inequality follows from a standard application of the Cauchy–Schwarz inequality. □

Remark 1 Note that, in general, the bounds in Theorem 7 can be infinite. We now exhibit a first (elementary) example of a random variable in $\operatorname{dom} D$ such that these bounds are finite.

Example 1 Consider a centered Poisson measure $\hat{\eta}$ on $X = \mathbb{R}_+$, with intensity measure equal to the Lebesgue measure ℓ. Then, for every integer $k \geq 1$, the random variable $F_k = k^{-1/2}\,\hat{\eta}([0,k]) \in \operatorname{dom} D$ and $DF_k = -DL^{-1}F_k = k^{-1/2}\mathbb{1}_{[0,k]}$. One immediately deduces from Theorem 7 that

$$d_W(F,N) \leq \frac{1}{k^{1/2}},$$

which proves the asymptotic normality of the random variables F_k as $k \to \infty$ and is consistent with the usual Berry–Esseen estimates in the Central Limit Theorem.

4.2 Bounds on the Kolmogorov Distance

When dealing with the Kolmogorov distance in the framework of Poisson functionals, it is usually very difficult to apply the uniform bound (14), since such a bound does not exploit the fine second order behavior of the Stein solution f_z. Refining an idea first developed by Schulte in [31], Eichelsbacher and Thäle [12] have obtained the following powerful estimate, whose proof uses the collection of local inequalities (13). As in the proof of Theorem 7, given two (possibly random) functions g, h on $\mathbb{X}$, we write $\langle h, g\rangle$ to indicate the integral $\int_{\mathbb{X}} h(z)g(z)\mu(\mathrm{d}z)$, whenever it is well defined. As before, whenever both h and g are in $L^2(\mu)$, we shall adopt the more precise notation $\langle h, g\rangle_{L^2(\mu)}$.

Theorem 8 *Let F be a centered element of* $\operatorname{dom} D$*, and let $N \sim \mathcal{N}(0,1)$. Then,*

$$\begin{aligned} d_K(F,N) \leq \mathbb{E}\left[\left|1 - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)}\right|\right] + \frac{\sqrt{2\pi}}{8}\mathbb{E}\left[\langle |DF|^2, |DL^{-1}F|\rangle\right] \\ + \frac{1}{2}\mathbb{E}\left[\langle |DF|^2, |F \times DL^{-1}F|\rangle\right] \\ + \sup_{z\in\mathbb{R}} \mathbb{E}\left[\langle (DF)D\mathbb{1}\{F > z\}, |DL^{-1}F|\rangle_{L^2(\mu)}\right], \end{aligned} \tag{48}$$

where $D_a\mathbb{1}\{F > z\} = \mathbb{1}\{F_a > z\} - \mathbb{1}\{F > z\}$, $a \in \mathbb{X}$.

Remark 2 In view of Lemma 4, $\mathbb{P}$-a.s. one has that

$$(D_aF)D_a\mathbb{1}\{F > z\} = (F_a - F)(\mathbb{1}\{F_a > z\} - \mathbb{1}\{F > z\}) \geq 0$$

for μ-almost every a in $\mathbb{X}$.

Remark 3 Let $F_k = k^{-1/2}\,\hat{\eta}\,([0,k])$, $k \geq 1$, be the collection of random variables studied in Example 1. Then, it is an easy exercise to show that relation (48) yields that, for some finite constant $C > 0$,

$$d_{\mathrm{Kol}}(F_k, N) \leq \frac{C}{k^{1/2}},$$

which is once again consistent with the usual Berry–Esseen estimates.

Proof (Theorem 8) Fix $z \in \mathbb{R}$. By using the explicit form of f_z (see (10)) together with Lemma 4, one proves immediately that $f_z(F) \in \mathrm{dom}\, D$, and therefore, by integrating by parts,

$$\mathbb{E}\,[Ff_z(F)] = \mathbb{E}\int_{\mathbb{X}} (f_z(F_a) - f_z(F))(-D_aL^{-1}F)\mu(\mathrm{d}a). \tag{49}$$

Using (13) in the case $x = F$ and $h = D_aF$, one sees that, for every $a \in \mathbb{X}$,

$$\begin{aligned}&\left|f_z(F_a) - f_z(F) - D_aFf_z'(F)\right|\\ &\quad\leq \frac{(D_aF)^2}{2}\left(|F| + \frac{\sqrt{2\pi}}{4}\right) + D_aFD_a\mathbb{1}\{F > z\}.\end{aligned}$$

Plugging this estimate into (49) and taking the supremum over all $z \in \mathbb{R}$ yields the desired conclusion. □

If $DF|DL^{-1}F| \in \mathrm{dom}\,\delta$ and $\mathbb{1}\{F > z\} \in \mathrm{dom}\, D$, for every $z \in \mathbb{R}$, then integration by parts yields

$$\begin{aligned}0 &\leq \mathbb{E}\left[\left\langle (DF)D\mathbb{1}\{F > z\}, |DL^{-1}F|\right\rangle_{L^2(\mu)}\right]\\ &= \mathbb{E}\left[\left\langle D\mathbb{1}\{F > z\}, DF|DL^{-1}F|\right\rangle_{L^2(\mu)}\right]\\ &= \mathbb{E}[\mathbb{1}\{F > z\}\delta(DF|DL^{-1}F|)] \leq \mathbb{E}[\delta(DF|DL^{-1}F|)^2]^{1/2}.\end{aligned}$$

As observed in [12, 31] the latter expectation can be controlled by applying standard moment estimates for Skorohod integrals, as stated in the forthcoming Proposition 2. See [20, Proposition 2.3] for a proof. See [25] for similar computations in the context of Gamma approximations.

Proposition 2 *Let* $u \in \mathrm{dom}\,\delta$. *Then,*

$$\mathbb{E}[\delta(u)^2] \leq \mathbb{E}\int_{\mathbb{X}} u(x)^2\mu(\mathrm{d}x) + \mathbb{E}\int_{\mathbb{X}}\int_{\mathbb{X}} (D_yu(x))^2\mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

The bounds appearing in this section involve the inverse L^{-1} of the Ornstein–Uhlenbeck operator, which is for the moment a quite abstract object one may find difficult to deal with. In the next section, we will discuss some results allowing to explicitly assess expectations involving such a mapping.

5 How Can One Deal with L^{-1}?

The aim of this section is to briefly describe and illustrate some strategies that one can implement, in order to explicitly compute the bounds appearing in the previous section. As anticipated, one of the main technical difficulties in order to deal with such bounds as (45) or (48) is the presence of the operator L^{-1}, which is consequently the main focus of the forthcoming discussion.

5.1 Using Chaotic Expansions

Our first elementary remark is that, in view of formulae (38) and (42), for every $F \in L^2_\eta$ having a chaotic decomposition of the type (36), the random variable $L^{-1}F$ is an element of $\operatorname{dom} D$, and one has that

$$D_z L^{-1} F = \sum_{n=1}^{\infty} I_n(g_n(z,\cdot)), \quad z \in \mathbb{X}. \tag{50}$$

It follows that expectations involving L^{-1} can be explicitly studied, whenever one has access to some detailed information about the kernels g_n appearing in the chaotic decomposition of F: typically, such a study starts with an application of the fundamental formula $g_n = (n!)^{-1} T_n f$, where f is a representative of F and the operators T_n are defined in [19, formula (1.16)] (see [19, Theorem 2]). This strategy is illustrated in [18, 32] (and the references discussed therein) in the important case where F is a so-called *U-statistic* based on η, so that, in particular, F lives in a finite sum of Wiener chaoses. For an application of this strategy in the case of random variables having an infinite chaotic expansion, see, e.g., [14].

As an important illustration, in this section we discuss how the bound on the Wassertein distance (45) can be used in the special case where $F = I_q(g)$, where $q \geq 1$ and $g \in L^2_s(\mu^q)$ (bounds in the Kolmogorov distance can be dealt with in a similar way). Our starting point is the following statement, whose proof follows immediately from (45) and (50).

Lemma 8 *For $q \geq 1$, let F be an element of the qth Wiener chaos of η, that is: $F = I_q(g)$, for some $g \in L^2_s(\mu^q)$. Then, if $N \sim \mathcal{N}(0,1)$,*

$$d_W(F,N) \leq \sqrt{\frac{2}{\pi}}\mathbb{E}\left|1 - \frac{1}{q}\|DF\|^2_{L^2(\mu)}\right| + \frac{1}{q}\int_{\mathbb{X}} \mathbb{E}\,|D_z F|^3\,\mu(\mathrm{d}z). \tag{51}$$

Combining (51) with the multiplication formula proved in [19, formula (1.67)], in [27] the following upper bound was obtained. The proof is standard but very long and quite technical, and consequently falls outside the scope of the present survey. As shown in [27], the assumption on the support and boundedness of g can be considerably weakened, at the cost of some additional technical assumption. Malliavin-Stein bounds!one-dimensional normal approximation!multiple integrals

Theorem 9 (See Theorem 4.2 in [27]) *Let F verify the assumptions of Lemma 8 for some $q \geq 2$, and assume in addition that the kernel g is bounded and has support contained in a set of the form $A \times \cdots \times A$, where A is a measurable set such that $\mu(A) < \infty$. Then, $g \in L^4_s(\mu^q)$ and, for every pair of integers (r, l) such that $1 \leq r \leq q$ and $1 \leq l \leq r \wedge (q-1)$, the kernel $g \star_r^l g$ (as defined in [19, Sect. 1.6]) are well defined and square-integrable. Moreover, there exists a universal constant C, uniquely depending on q, such that*

$$d_W(F, N) \leq \sqrt{1 - q!\|g\|^2_{L^2(\mu^q)}} + C \max\left\{\|g \star_r^l g\|_{L^2(\mu^{2q-r-l})}, \|g\|^2_{L^4(\mu^q)}\right\}, \quad (52)$$

where the maximum runs over all pairs (r, l) such that $1 \leq r \leq q$ and $1 \leq l \leq r \wedge (q-1)$.

The estimate (52) should be compared with analogous bounds for multiple integrals with respect to a Gaussian measure, as discussed, e.g., in [21, Sect. 5.2.2] and the references therein. An explicit application of (52) is developed in the subsequent section.

5.2 CLTs for Multiple Integrals: Necessary and Sufficient Conditions

As a nontrivial application of (52), we now establish two *fourth moment theorems* for Poisson multiple integrals of order 2 and 3. The case of order 2 has been dealt with in [23], where the authors assume different hypotheses than those appearing in the upcoming statement. The case of order 3 is new. Notice that the question of whether similar results can be proved for sequences of multiple Poisson integrals of order ≥ 4 stays open. Other fourth moment theorems for sums of multiple integrals whose kernels have a constant sign can be found in [15, 16]. As already observed for Theorem 9, the findings of the present section should be compared with analogous results on a Gaussian space—where fourth moment theorems hold for multiple integrals of arbitrary orders and are at the core of countless applications—see, e.g., [21, Theorem 5.2.7] and the discussion therein.

Theorem 10 (Fourth Moment Theorem for Double Integrals) *Let $\{f_n\}_{n\geq 1}$ be a sequence in $L^2_s(\mu^2)$ such that, for each n, the kernel f_n verifies the same assumptions as the kernel g in Theorem 9 (for $q = 2$ and for some measurable set A_n such that*

$\mu(A_n) < \infty$), and also

$$\lim_{n\to\infty} 2\|f_n\|^2_{L^2(\mu^2)} = 1.$$

Furthermore, assume that $\lim_{n\to\infty} \|f_n\|^4_{L^4(\mu^2)} = 0$. Then, it holds that, for $N \sim \mathcal{N}(0,1)$,

$$\mathbb{E}\left[I_2(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3 = \mathbb{E}\left[N^4\right] \Longrightarrow d_W(I_2(f_n), N) \underset{n\to\infty}{\longrightarrow} 0.$$

Moreover, if the sequence $\left\{I_2(f_n)^4\right\}_{n\geq 1}$ is uniformly integrable, then

$$\mathbb{E}\left[I_2(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3 = \mathbb{E}\left[N^4\right] \Longleftrightarrow d_W(I_2(f_n), N) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof First notice that if $I_2(f_n) \underset{n\to\infty}{\longrightarrow}$ converges in distribution to N and the sequence $\left\{I_2(f_n)^4\right\}_{n\geq 1}$ is uniformly integrable, then necessarily $\mathbb{E}\left[I_2(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left[N^4\right] = 3$. Conversely, using the product formula for Poisson multiple integrals [19, formula (1.67)], we have

$$\begin{aligned} I_2(f_n)^2 &= 2\|f_n\|^2_{L^2(\mu^2)} + I_4\left(\widetilde{f_n \otimes f_n}\right) + 4I_3\left(\widetilde{f_n \star_1^0 f_n}\right) \\ &\quad + 2I_2\left(2\widetilde{f_n \star_1^1 f_n} + \widetilde{f_n \star_2^0 f_n}\right) + 4I_1\left(\widetilde{f_n \star_2^1 f_n}\right), \end{aligned}$$

which yields, using the orthogonality of multiple integrals of different orders,

$$\begin{aligned} \mathbb{E}\left[I_2(f_n)^4\right] &= 4\|f_n\|^4_{L^2(\mu^2)} + 24\|\widetilde{f_n \otimes f_n}\|^2_{L^2(\mu^2)} + 96\|\widetilde{f_n \star_1^0 f_n}\|^2_{L^2(\mu^3)} \\ &\quad + 16\|\widetilde{f_n \star_2^1 f_n}\|^2_{L^2(\mu)} + 32\|\widetilde{f_n \otimes_1 f_n} + \frac{1}{2}\widetilde{f_n \star_2^0 f_n}\|^2_{L^2(\mu^2)}. \end{aligned}$$

As $24\|\widetilde{f_n \otimes f_n}\|^2_{L^2(\mu^2)} = 8\|f_n\|^4_{L^2(\mu^2)} + 16\|f_n \star_1^1 f_n\|^2_{L^2(\mu^2)}$ by Formula 11.6.30 in [24], we finally get

$$\begin{aligned} \mathbb{E}\left[I_2(f_n)^4\right] &= 3\left(4\|f_n\|^4_{L^2(\mu^2)}\right) + 16\|f_n \star_1^1 f_n\|^2_{L^2(\mu^2)} + 96\|\widetilde{f_n \star_1^0 f_n}\|^2_{L^2(\mu^3)} \\ &\quad + 16\|\widetilde{f_n \star_2^1 f_n}\|^2_{L^2(\mu)} + 32\|\widetilde{f_n \star_1^1 f_n} + \frac{1}{2}\widetilde{f_n \star_2^0 f_n}\|^2_{L^2(\mu^2)}. \end{aligned}$$

The fact that $\mathbb{E}\left[I_2(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3$ along with the condition $2\|f_n\|^2_{L^2(\mu^2)} \underset{n\to\infty}{\longrightarrow} 1$ implies that all the contractions appearing in the above expression go to zero as n goes to infinity. Furthermore, as $\|\widetilde{f_n \star_1^1 f_n}\|^2_{L^2(\mu^2)} \leq \|f_n \star_1^1 f_n\|^2_{L^2(\mu^2)}$, we get that $\|\widetilde{f_n \star_1^1 f_n}\|^2_{L^2(\mu^2)} \underset{n\to\infty}{\longrightarrow} 0$, which in turn implies that $\|\widetilde{f_n \star_2^0 f_n}\|^2_{L^2(\mu^2)} \underset{n\to\infty}{\longrightarrow} 0$. Using the bound provided in (52) concludes the proof. □

The next statement contains the announced result for triple integrals.

Theorem 11 (Fourth Moment Theorem for Triple Integrals) *Let $\{f_n\}_{n\geq 1}$ be a sequence in $L^2_s(\mu^3) \cap L^4_s(\mu^3)$ such that, for each n, the kernel f_n verifies the same assumptions as the kernel g in Theorem 9 (for $q = 3$ and for some measurable set A_n such that $\mu(A_n) < \infty$), and also $\lim_{n\to\infty} 6\|f_n\|^2_{L^2(\mu^3)} = 1$. Furthermore, assume that $\lim_{n\to\infty} \|f\|^4_{L^4(\mu^3)} = 0$. Then, it holds that, for $N \sim \mathcal{N}(0,1)$,*

$$\mathbb{E}\left[I_3(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3 = \mathbb{E}\left[N^4\right] \Longrightarrow d_W(I_3(f_n), N) \underset{n\to\infty}{\longrightarrow} 0.$$

Moreover, if the sequence $\left\{I_3(f_n)^4\right\}_{n\geq 1}$ is uniformly integrable, then

$$\mathbb{E}\left[I_3(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3 = \mathbb{E}\left[N^4\right] \Longleftrightarrow d_W(I_3(f_n), N) \underset{n\to\infty}{\longrightarrow} 0.$$

Proof As before, if $I_3(f_n)$ converges in distribution to N and the sequence $\left\{I_3(f_n)^4\right\}_{n\geq 1}$ is uniformly integrable, then necessarily $\mathbb{E}\left[I_3(f_n)^4\right] \underset{n\to\infty}{\longrightarrow} 3$. Conversely, using the multiplication formula [19, formula (1.67)] and the combinatorial relation [24, formula (11.6.30)] as in the proof of Theorem 10, we get

$$\begin{aligned}\mathbb{E}\left[I_3(f_n)^4\right] &= 3\left(36\|f\|^4_{L^2(\mu^3)}\right) + 1296\|f_n \star_1^1 f_n\|^2_{L^2(\mu^4)} + 1296\|f_n \star_2^2 f_n\|^2_{L^2(\mu^2)} \\ &+9720\|\widetilde{f_n \star_1^0 f_n}\|^2_{L^2(\mu^5)} + 324\|\widetilde{f_n \star_3^2 f_n}\|^2_{L^2(\mu)} + 1944\|\widetilde{2f_n \star_2^1 f_n} + \widetilde{f_n \star_3^0 f_n}\|^2_{L^2(\mu^3)} \\ &+1944\|\widetilde{f_n \star_1^1 f_n} + 2\widetilde{f_n \star_2^0 f_n}\|^2_{L^2(\mu^4)} + 648\|\widetilde{f_n \star_2^2 f_n} + \widetilde{f_n \star_3^1 f_n}\|^2_{L^2(\mu^2)} \end{aligned} \tag{53}$$

(note that the above relation can be deduced—as an interesting exercise!—by a careful use of the multiplication formula). The condition $6\|f_n\|^2_{L^2(\mu^2)} \underset{n\to\infty}{\longrightarrow} 1$ ensures that $3\left(36\|f\|^4_{L^2(\mu^3)}\right) \underset{n\to\infty}{\longrightarrow} 3$, hence implying that all the norms appearing in the expression of the fourth moment converge to zero as n goes to infinity. Furthermore, it can be seen that

$$\|f_n \otimes_1 f_n\|^2_{L^2(\mu^4)} + \|\widetilde{f_n \star_1^1 f_n + 2 f_n \star_2^0 f_n}\|^2_{L^2(\mu^4)} \underset{n\to\infty}{\longrightarrow} 0 \Longrightarrow \|\widetilde{f_n \star_2^0 f_n}\|^2_{L^2(\mu^4)} \underset{n\to\infty}{\longrightarrow} 0,$$

as $\|\widetilde{f_n \star_1^1 f_n}\|^2_{L^2(\mu^4)} \leq \|f_n \star_1^1 f_n\|^2_{L^2(\mu^4)}$.

The same argument yields $\|\widetilde{f_n \star_3^1 f_n}\|^2_{L^2(\mu^2)} \underset{n\to\infty}{\longrightarrow} 0$. Observe that $f_n \star_3^0 f_n = f_n^2$ and hence

$$\|\widetilde{f_n \star_3^0 f_n}\|^2_{L^2(\mu^3)} = \|f\|^4_{L^4(\mu^3)} \underset{n\to\infty}{\longrightarrow} 0.$$

Combining this with $\|2\widetilde{f_n \star_2^1 f_n} + \widetilde{f_n \star_3^0 f_n}\|^2_{L^2(\mu^3)} \underset{n\to\infty}{\longrightarrow} 0$ yields

$$\|\widetilde{f_n \star_2^1 f_n}\|^2_{L^2(\mu^3)} \underset{n\to\infty}{\longrightarrow} 0.$$

Using the bound (52) along with the conclusion that all the norms of the contractions norms in the fourth moment expression converge to zero as n goes to infinity concludes the proof. □

5.3 Mehler's Formula and Second Order Inequalities

For many random variables $F \in L^2_\eta$ having a chaotic decomposition as in (36) (especially in the context of stochastic geometry), the task of explicitly computing the kernels $g_n = T_n f$ turns out to be technically very challenging. As a consequence, it might be preferable to work directly with the integral representation (44), combined with the Mehler's formula [19, formula (1.71)], providing an explicit probabilistic representation of the operators $\{P_t\}$. This line of research has been successfully pursued in [20], where Berry–Esseen bounds (in the Kolmogorov distance) not displaying L^{-1} have been deduced from (48) *via* Mehler's formula. These estimates are expressed in terms of the following quantities γ_i, $i = 1, \ldots, 6$:

$$\gamma_1 := 2\left[\int \left[\mathbb{E}(D_{x_1}F)^2(D_{x_2}F)^2\right]^{1/2}\left[\mathbb{E}(D^2_{x_1,x_3}F)^2(D^2_{x_2,x_3}F)^2\right]^{1/2} \lambda^3(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},$$

$$\gamma_2 := \left[\int \mathbb{E}(D^2_{x_1,x_3}F)^2(D^2_{x_2,x_3}F)^2 \,\lambda^3(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},$$

$$\gamma_3 := \int \mathbb{E}|D_xF|^3 \,\lambda(\mathrm{d}x)$$

$$\gamma_4 := \frac{1}{2}\big[\mathbb{E}F^4\big]^{1/4} \int \big[\mathbb{E}(D_xF)^4\big]^{3/4}\,\lambda(\mathrm{d}x),$$

$$\gamma_5 := \left[\int \mathbb{E}(D_xF)^4\,\lambda(\mathrm{d}x)\right]^{1/2},$$

$$\gamma_6 := \left[\int 6\big[\mathbb{E}(D_{x_1}F)^4\big]^{1/2}\big[\mathbb{E}(D^2_{x_1,x_2}F)^4\big]^{1/2} + 3\mathbb{E}(D^2_{x_1,x_2}F)^4\,\lambda^2(\mathrm{d}(x_1,x_2))\right]^{1/2}.$$

The next bound can be seen as an extension, to the Poisson setting, of the second order Poincaré inequalities on the Gaussian space—see, e.g., [21, Theorem 5.3.3].

Theorem 12 (See [20]) *Let $F \in \operatorname{dom} D$ be such that $\mathbb{E}[F] = 0$ and $\mathbf{Var}[F] = 1$, and let N be a standard Gaussian random variable. Then,*

$$d_K(F,N) \leq \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5 + \gamma_6.$$

As explained in [20], the content of Theorem 12 has striking connections with the theory of *stabilization*, as initiated in the seminal paper [29]. See again [20] for several applications to nearest neighbor graph statistics, as well as to intrinsic volumes of k-faces arising in Voronoi tessellations and to nonlinear functionals of shot-noise processes.

5.4 A Connection with Logarithmic Sobolev Inequalities

We conclude this section by showing how Mehler's formula [19, formula (1.71)] can be used to provide a direct, intrinsic proof, of an important *modified logarithmic Sobolev inequality* proved by Wu in [33]. Recall that the *entropy* of a given random variable F such that $F > 0$, a.s.-$\mathbb{P}$, and $\mathbb{E}F < \infty$ is defined as

$$\operatorname{Ent}(F) := \mathbb{E}(F\log F) - \mathbb{E}(F)\log\mathbb{E}(F).$$

Theorem 13 (Modified Logarithmic Sobolev Inequality—See [33]) *Let $F \in \operatorname{dom} D$ be such that $F > 0$ with probability one. Then, writing $\Phi(x) := x\log x$, $x > 0$,*

$$\operatorname{Ent}(F) = \mathbb{E}[\Phi(F)] - \Phi(\mathbb{E}(F)) \leq \mathbb{E}\int_{\mathbb{X}} \big(D_z\Phi(F) - \Phi'(F)D_zF\big)\,\mu(dz). \tag{54}$$

Proof By a standard approximation argument we can assume that there exist finite constants ϵ, η such that $0 < \epsilon < F < \eta$ with probability one. In this way, all computations appearing below—involving in particular exchanging derivations and expectations—are formally justified by classical measure-theoretical results. We shall use the relation $d/dtP_tF = -t^{-1}LP_tF$, as well as the fact that the mapping

$$(x, y) \mapsto y(\Phi'(x + y) - \Phi'(x))$$

is convex (see also [6] for a broader analysis of this property). We have

$$\begin{aligned}
\mathbb{E}[\Phi(F)] - \Phi(\mathbb{E}(F)) &= \mathbb{E}(\Phi(P_1F) - \Phi(P_0F)) \\
&= \int_0^1 \mathbb{E}\left(\frac{d}{dt}\Phi(P_tF)\right)\mathrm{d}t = \int_0^1 \mathbb{E}\left(\Phi'(P_tF)\frac{d}{dt}P_tF\right)\mathrm{d}t \\
&= -\int_0^1 \mathbb{E}\left(\Phi'(P_tF)LP_tF\right)t^{-1}\mathrm{d}t = \int_0^1 \mathbb{E}\left(\Phi'(P_tF)\delta DP_tF\right)t^{-1}\mathrm{d}t \\
&= \int_{\mathbb{X}} \mathbb{E}\left[\int_0^1 (D_z\Phi'(P_tF) \times D_zP_tF)\, t^{-1}\mathrm{d}t\right]\mu(\mathrm{d}z).
\end{aligned}$$

Using convexity and Mehler's formula, together with the relation $DP_tF = tP_tDF$, we deduce that, for all z,

$$\begin{aligned}
\mathbb{E}\left[\int_0^1 (D_z\Phi'(P_tF) \times D_zP_tF)\, t^{-1}\mathrm{d}t\right] &\leq \mathbb{E}\left[\int_0^1 (\Phi'(F + tD_zF) - \Phi'(F))\,\mathrm{d}t \times D_zF\right] \\
&= \mathbb{E}\left[\left(\int_0^1 \Phi'(F + tD_zF)\mathrm{d}t - \Phi'(F)\right) \times D_zF\right].
\end{aligned}$$

Since

$$\int_0^1 \Phi'(F + tD_zF)\mathrm{d}t = \frac{1}{D_zF}(\Phi(F + D_zF) - \Phi(F)),$$

we deduce the desired conclusion. □

Several applications of (54) in the context of concentration estimates for Poisson functionals can be found in [2, 3].

6 Multidimensional Malliavin–Stein Bounds in the Normal Approximation on the Poisson Space

As in the previous sections, we continue working within a general Poisson measure η, on the measurable space $(\mathbb{X}, \mathscr{X})$, with σ-finite intensity measure given by μ. We will now discuss some multidimensional extension of the bounds presented in Sect. 3.

6.1 Bounds Using Stein's Method

The following bound, that is based on the estimates of Theorem 5, allows one to deal with normal approximations, in the sense of the distance d_2, where the target Gaussian distribution has a non-singular covariance matrix.

Theorem 14 *Fix $d \geq 2$ and let $\Sigma = \{\sigma(i,j) : i,j = 1, \dots, d\}$ be a $d \times d$ positive definite matrix. Suppose that $N \sim \mathscr{N}(0, \Sigma)$ and that $F = (F_1, \dots, F_d)$ is a $\mathbb{R}^d$-valued random vector such that $\mathbb{E}[F_i] = 0$ and $F_i \in \operatorname{dom} D$, $i = 1, \dots, d$. Then,*

$$d_2(F,N) \leq \|\Sigma^{-1}\|_{\text{op}} \|\Sigma\|_{\text{op}}^{1/2} \sqrt{\sum_{i,j=1}^{d} \mathbb{E}[(\Sigma(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{L^2(\mu)})^2]} \quad (55)$$

$$+ \frac{\sqrt{2\pi}}{8} \|\Sigma^{-1}\|_{\text{op}}^{3/2} \|\Sigma\|_{\text{op}} \int_{\mathbb{X}} \mathbb{E}\left[\left(\sum_{i=1}^{d} |D_z F_i|\right)^2 \left(\sum_{i=1}^{d} |D_z L^{-1} F_i|\right)\right] \mu(\mathrm{d}z). \quad (56)$$

Proof If either one of the expectations in (55) and (56) is infinite, there is nothing to prove. We shall therefore work under the assumption that both expressions (55)–(56) are finite. By the definition of the distance d_2, and by using a standard approximation argument, one sees that it is sufficient to show the following estimate:

$$\begin{aligned} &|\mathbb{E}[h(X)] - \mathbb{E}[h(F)]| \\ &\leq A \|\Sigma^{-1}\|_{\text{op}} \|\Sigma\|_{\text{op}}^{1/2} \sqrt{\sum_{i,j=1}^{d} \mathbb{E}[(\Sigma(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{L^2(\mu)})^2]} \qquad (57) \\ &\quad + \frac{\sqrt{2\pi}}{8} B \|\Sigma^{-1}\|_{\text{op}}^{3/2} \|\Sigma\|_{\text{op}} \int_{\mathbb{X}} \mathbb{E}\left[\left(\sum_{i=1}^{d} |D_z F_i|\right)^2 \left(\sum_{i=1}^{d} |D_z L^{-1} F_i|\right)\right] \mu(\mathrm{d}z), \end{aligned}$$

for any $h \in \mathscr{C}^\infty$ with first and second bounded derivatives, such that $\|h\|_{\text{Lip}} \le A$ and $M_2(h) \le B$. To prove (57), we use Theorem 5 to infer that

$$
\begin{aligned}
&|\mathbb{E}[h(X)] - \mathbb{E}[h(F)]| \\
&= |\mathbb{E}[\langle \Sigma, \text{Hess} f_h(F)\rangle_{H.S.} - \langle F, \nabla f_h(F)\rangle_{\mathbb{R}^d}]| \\
&= \left| \mathbb{E}\left[\sum_{i,j=1}^d \Sigma(i,j) \frac{\partial^2}{\partial x_i \partial x_j} f_h(F) - \sum_{k=1}^d F_k \frac{\partial}{\partial x_k} f_h(F) \right] \right| \\
&= \left| \sum_{i,j=1}^d \mathbb{E}\left[\Sigma(i,j) \frac{\partial^2}{\partial x_i \partial x_j} f_h(F) \right] + \sum_{k=1}^d \mathbb{E}\left[\delta(DL^{-1}F_k) \frac{\partial}{\partial x_k} f_h(F) \right] \right| \\
&= \left| \sum_{i,j=1}^d \mathbb{E}\left[\Sigma(i,j) \frac{\partial^2}{\partial x_i \partial x_j} f_h(F) \right] - \sum_{k=1}^d \mathbb{E}\left[\left\langle D\left(\frac{\partial}{\partial x_k} f_h(F) \right), -DL^{-1}F_k \right\rangle_{L^2(\mu)} \right] \right|.
\end{aligned}
$$

We write $\frac{\partial}{\partial x_k} f_h(F) := \varphi_k(F_1, \ldots, F_d) = \varphi_k(F)$. By using Lemma 4, we deduce that, $\mathbb{P}$-a.s. and for every $z \in \mathbb{X}$ (except at most for a set of μ measure 0),

$$
D_z \varphi_k(F_1, \ldots, F_d) = \sum_{i=1}^d \frac{\partial}{\partial x_i} \varphi_k(F)(D_z F_i) + R_k,
$$

with $R_k = \sum_{i,j=1}^d R_{i,j,k}(D_z F_i, D_z F_j)$, and

$$
|R_{i,j,k}(y_1, y_2)| \le \frac{1}{2} \sup_{x \in \mathbb{R}^d} \left| \frac{\partial^2}{\partial x_i \partial x_j} \varphi_k(x) \right| \times |y_1 y_2|.
$$

It follows that

$$
\begin{aligned}
&|\mathbb{E}[h(X)] - \mathbb{E}[h(F)]| \\
&= \left| \sum_{i,j=1}^d \mathbb{E}\left[\Sigma(i,j) \frac{\partial^2}{\partial x_i \partial x_j} f_h(F) \right] - \sum_{i,k=1}^d \mathbb{E}\left[\frac{\partial^2}{\partial x_i \partial x_k} (f_h(F)) \langle DF_i, -DL^{-1}F_k \rangle_{L^2(\mu)} \right] \right. \\
&\quad \left. + \sum_{i,j,k=1}^d \mathbb{E}\left[\langle R_{i,j,k}(DF_i, DF_j), -DL^{-1}F_k \rangle_{L^2(\mu)} \right] \right| \\
&\le \sqrt{\mathbb{E}[\|\text{Hess} f_h(F)\|^2_{H.S.}]} \times \sqrt{\sum_{i,j=1}^d \mathbb{E}\left[\left(\Sigma(i,j) - \langle DF_i, -DL^{-1}F_j \rangle_{L^2(\mu)} \right)^2 \right]} + |R_2|,
\end{aligned}
$$

where

$$R_2 = \sum_{i,j,k=1}^{d} \mathbb{E}[\langle R_{i,j,k}(DF_i, DF_j), -DL^{-1}F_k\rangle_{L^2(\mu)}].$$

Theorem 5 yields that $\|\mathrm{Hess} f_h(F)\|_{H.S.} \leq \|\Sigma^{-1}\|_{\mathrm{op}} \|\Sigma\|_{\mathrm{op}}^{1/2} \|h\|_{\mathrm{Lip}}$. Using the elementary fact that all partial derivatives of order 3 of h are bounded by $M_3(h)$, we have

$$\begin{aligned} |R_{i,j,k}(y_1, y_2)| &\leq \frac{1}{2} \sup_{x\in\mathbb{R}^d} \left| \frac{\partial^3}{\partial x_i \partial x_j \partial x_k} f_h(y) \right| \times |y_1 y_2| \\ &\leq \frac{\sqrt{2\pi}}{8} M_2(h) \|\Sigma^{-1}\|_{\mathrm{op}}^{3/2} \|\Sigma\|_{\mathrm{op}} \times |y_1 y_2| \\ &\leq \frac{\sqrt{2\pi}}{8} B \|\Sigma^{-1}\|_{\mathrm{op}}^{3/2} \|\Sigma\|_{\mathrm{op}} \times |y_1 y_2|, \end{aligned}$$

from which we deduce the desired conclusion. □

In the next subsection, we shall show how one can deal with singular covariance matrices.

6.2 Bounds Obtained by Interpolation

The proof of the next result uses an interpolation technique, sometimes called "smart path method" that represents a valid alternative to Stein's method, and allows in particular to deal with covariance matrices that are degenerate. The price to pay is the fact that one has to deal with smoother test functions.

Theorem 15 *Fix $d \geq 1$ and let $\Sigma = \{\Sigma(i,j) : i,j = 1, \dots, d\}$ be a $d\times d$ covariance matrix (not necessarily positive definite). Suppose that $N = (N_1, \dots, N_d) \sim \mathscr{N}(0, \Sigma)$ and that $F = (F_1, \dots, F_d)$ is a $\mathbb{R}^d$-valued random vector such that $\mathbb{E}[F_i] = 0$ and $F_i \in \mathrm{dom}\, D$, $i = 1, \dots, d$. Then,*

$$d_3(F, X) \leq \frac{d}{2} \sqrt{\sum_{i,j=1}^{d} \mathbb{E}[(\Sigma(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{L^2(\mu)})^2]} \tag{58}$$

$$+\frac{1}{4} \int_{\mathbb{X}} \mathbb{E}\left[\left(\sum_{i=1}^{d} |D_z F_i| \right)^2 \left(\sum_{i=1}^{d} |D_z L^{-1} F_i| \right) \right] \mu(\mathrm{d}z). \tag{59}$$

Proof We will work under the assumption that both expectations in (58) and (59) are finite. By the definition of the distance d_3, we need only to show the following estimate:

$$|\mathbb{E}[\varphi(X)] - \mathbb{E}[\varphi(F)]| \leq \frac{1}{2}\sum_{i,j=1}^{d} \mathbb{E}[|\Sigma(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{L^2(\mu)}|] + \frac{1}{4}\int_{\mathbb{X}} \mathbb{E}\left[\left(\sum_{i=1}^{d} |D_z F_i|\right)^2 \left(\sum_{i=1}^{d} |D_z L^{-1} F_i|\right)\right] \mu(\mathrm{d}z)$$

for any $\varphi \in \mathscr{C}^3$ with second and third derivatives bounded by 1. Without loss of generality, we may assume that F and N are independent. For $t \in [0, 1]$, we set

$$\Psi(t) = \mathbb{E}[\varphi(\sqrt{1-t}(F_1, \ldots, F_d) + \sqrt{t}N)].$$

We have immediately

$$|\Psi(1) - \Psi(0)| \leq \sup_{t\in(0,1)} |\Psi'(t)|.$$

Indeed, due to the assumptions on φ, the function $t \mapsto \Psi(t)$ is differentiable on $(0, 1)$, and one has also

$$\begin{aligned}\Psi'(t) &= \sum_{i=1}^{d} \mathbb{E}\left[\frac{\partial}{\partial x_i}\varphi\left(\sqrt{1-t}(F_1, \ldots, F_d) + \sqrt{t}N\right)\left(\frac{1}{2\sqrt{t}}N_i - \frac{1}{2\sqrt{1-t}}F_i\right)\right] \\ &:= \frac{1}{2\sqrt{t}}\,\mathbb{A} - \frac{1}{2\sqrt{1-t}}\,\mathbb{B}.\end{aligned}$$

On the one hand, we have (by integration by parts)

$$\begin{aligned}\mathbb{A} &= \sum_{i=1}^{d} \mathbb{E}\left[\frac{\partial}{\partial x_i}\varphi(\sqrt{1-t}(F_1, \ldots, F_d) + \sqrt{t}N)N_i\right] \\ &= \sum_{i=1}^{d} \mathbb{E}\left[\mathbb{E}\left[\frac{\partial}{\partial x_i}\varphi(\sqrt{1-t}a + \sqrt{t}N)N_i\right]_{|a=(F_1,\ldots,F_d)}\right] \\ &= \sqrt{t}\sum_{i,j=1}^{d} \Sigma(i,j)\mathbb{E}\left[\mathbb{E}\left[\frac{\partial^2}{\partial x_i \partial x_j}\varphi(\sqrt{1-t}a + \sqrt{t}N)\right]_{|a=(F_1,\ldots,F_d)}\right] \\ &= \sqrt{t}\sum_{i,j=1}^{d} \Sigma(i,j)\mathbb{E}\left[\frac{\partial^2}{\partial x_i \partial x_j}\varphi(\sqrt{1-t}(F_1, \ldots, F_d) + \sqrt{t}N)\right].\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{B} &= \sum_{i=1}^{d} \mathbb{E}\left[\frac{\partial}{\partial x_i}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}N)F_i\right]\\ &= \sum_{i=1}^{d} \mathbb{E}\left[\mathbb{E}\left[\frac{\partial}{\partial x_i}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}b)F_i\right]_{|b=N}\right].\end{aligned}$$

We now write $\varphi_i^{t,b}(\cdot)$ to indicate the function on $\mathbb{R}^d$ defined by

$$\varphi_i^{t,b}(F_1,\ldots,F_d) = \frac{\partial}{\partial x_i}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}b).$$

Integrating by parts

$$\begin{aligned}&\mathbb{E}[\varphi_i^{t,b}(F_1,\ldots,F_d)F_i]\\ &= \mathbb{E}\left[\sum_{j=1}^{d}\frac{\partial}{\partial x_j}\varphi_i^{t,b}(F_1,\ldots,F_d)\langle DF_j,-DL^{-1}F_i\rangle_{L^2(\mu)}\right] + \mathbb{E}\left[\langle R_b^i,-DL^{-1}F_i\rangle\right],\end{aligned}$$

where R_b^i is a residue verifying

$$\begin{aligned}&|\mathbb{E}[\langle R_b^i,-DL^{-1}F_i\rangle]| \qquad (60)\\ &\leq \frac{1}{2}\left(\max_{k,l}\sup_{x\in\mathbb{R}^d}\left|\frac{\partial}{\partial x_k\partial x_l}\varphi_i^{t,b}(x)\right|\right)\int_{\mathbb{X}}\mathbb{E}\left[\left(\sum_{j=1}^{d}|D_zF_j|\right)^2|D_zL^{-1}F_i|\right]\mu(\mathrm{d}z).\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{B} &= \sqrt{1-t}\sum_{i,j=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\frac{\partial^2}{\partial x_i\partial x_j}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}b)\langle DF_i,-DL^{-1}F_j\rangle_{L^2(\mu)}\right]_{|b=X}\right]\\ &\quad+\sum_{i=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\langle R_b^i,-DL^{-1}F_i\rangle_{L^2(\mu)}\right]_{|b=N}\right]\\ &= \sqrt{1-t}\sum_{i,j=1}^{d}\mathbb{E}\left[\frac{\partial^2}{\partial x_i\partial x_j}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}N)\langle DF_i,-DL^{-1}F_j\rangle_{L^2(\mu)}\right]\\ &\quad+\sum_{i=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\langle R_b^i,-DL^{-1}F_i\rangle\right]_{|b=N}\right].\end{aligned}$$

Putting the estimates on $\mathbb{A}$ and $\mathbb{B}$ together, we infer

$$\begin{aligned}
&\Psi'(t) \\
&= \frac{1}{2}\sum_{i,j=1}^{d} \mathbb{E}\left[\frac{\partial^2}{\partial x_i \partial x_j}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}X)(\Sigma(i,j)-\langle DF_i,-DL^{-1}F_j\rangle_{L^2(\mu)})\right] \\
&\quad -\frac{1}{2\sqrt{1-t}}\sum_{i=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\langle R_b^i,-DL^{-1}F_i\rangle_{L^2(\mu)}\right]_{|b=X}\right].
\end{aligned}$$

We notice that

$$\left|\frac{\partial^2}{\partial x_i \partial x_j}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}b)\right| \le 1,$$

and also

$$\begin{aligned}
\left|\frac{\partial^2}{\partial x_k \partial x_l}\varphi_i^{t,b}(F_1,\ldots,F_d)\right| &= (1-t)\times\left|\frac{\partial^3}{\partial x_i \partial x_k \partial x_l}\varphi(\sqrt{1-t}(F_1,\ldots,F_d)+\sqrt{t}b)\right| \\
&\le (1-t).
\end{aligned}$$

To conclude, we can apply inequality (60) and deduce the estimates

$$\begin{aligned}
&|\mathbb{E}[\varphi(N)]-\mathbb{E}[\varphi(F)]| \\
&\quad \le \sup_{t\in(0,1)}|\Psi'(t)| \\
&\quad \le \frac{1}{2}\sum_{i,j=1}^{d}\mathbb{E}[|\Sigma(i,j)-\langle DF_i,-DL^{-1}F_j\rangle_{L^2(\mu)}|] \\
&\qquad +\frac{1-t}{4\sqrt{1-t}}\|\varphi'''\|_\infty\int_{\mathbb{X}}\mathbb{E}\left[\left(\sum_{i=1}^{d}|D_zF_i|\right)^2\left(\sum_{i=1}^{d}|D_zL^{-1}F_i|\right)\right]\mu(\mathrm{d}z) \\
&\quad \le \frac{d}{2}\sqrt{\sum_{i,j=1}^{d}\mathbb{E}[(\Sigma(i,j)-\langle DF_i,-DL^{-1}F_j\rangle_{L^2(\mu)})^2]} \\
&\qquad +\frac{1}{4}\int_{\mathbb{X}}\mathbb{E}\left[\left(\sum_{i=1}^{d}|D_zF_i|\right)^2\left(\sum_{i=1}^{d}|D_zL^{-1}F_i|\right)\right]\mu(\mathrm{d}z),
\end{aligned}$$

thus concluding the proof. □

7 Poisson Approximation on the Poisson Space

As before, the framework of this section is the one of a general Poisson measure η on the measurable space $(\mathbb{X}, \mathscr{X})$, with σ-finite intensity measure given by μ. The aim of this section is to discuss the combination of the Chen–Stein method (see Sect. 2.4) and the Malliavin calculus of variations (see Sect. 3 and [19]) in order to study Poisson approximations for functionals of η, both in the one-dimensional case (Sect. 7.1) and the multidimensional case (Sect. 7.2).
The main achievement in the one-dimensional case is a general inequality on the Poisson space (see Theorem 16) assessing the distance in total variation between the law of a Poisson random variable and the law of a (sufficiently regular) integer-valued Poisson functional. The multidimensional case is treated as part of a wider result allowing one to assess the distance (in a certain sense) between the law of vector of functionals of η and the law of a vector composed by Poisson and Gaussian elements. A strong motivation for this set of results comes from applications in stochastic geometry (as illustrated by Lachièze-Rey and Reitzner [18] and Schulte and Thäle [32]).

7.1 *One-Dimensional Chen–Stein–Malliavin Method*

Based on the Chen–Stein bounds discussed in Sect. 2.4.1, the following result provides a general inequality (in terms of Malliavin operators) assessing the distance (in total variation) between the law of a Poisson random variable and the law of an integer-valued functional of η.

Theorem 16 *Let $Z \sim \mathrm{Po}(\lambda)$, $\lambda > 0$ and assume that $F \in L^2(P)$ is an element of* $\mathrm{dom}\, D$ *such that $\mathbb{E}(F) = \lambda$ and F takes values in $\mathbb{Z}_+$. Then,*

$$d_{\mathrm{TV}}(F, Z) \leq \frac{1-e^{-\lambda}}{\lambda}\mathbb{E}\left|\lambda - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)}\right| + \frac{1-e^{-\lambda}}{\lambda}\mathbb{E}\left[\int_{\mathbb{X}} \left|D_zF\,(D_zF-1)\,D_zL^{-1}F\right|\mu(\mathrm{d}z)\right] \tag{61}$$

$$\leq \frac{1-e^{-\lambda}}{\lambda}\sqrt{\mathbb{E}\left[\left(\lambda - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)}\right)^2\right]} + \frac{1-e^{-\lambda}}{\lambda}\mathbb{E}\left[\int_{\mathbb{X}} \left|D_zF\,(D_zF-1)\,D_zL^{-1}F\right|\mu(\mathrm{d}z)\right]. \tag{62}$$

If $F = \eta(A)$, *where* $\mu(A) = \lambda$, *then one has that* $D_zF = -D_zL^{-1}F = \mathbf{1}_A(z)$, *and*

$$\lambda - \langle DF, -DL^{-1}F\rangle_{L^2(\mu)} = \int_{\mathbb{X}} \left|D_zF\,(D_zF - 1)\,D_zL^{-1}F\right| \mu(\mathrm{d}z) = 0.$$

Proof Let the notation of Sect. 2.4.1 prevail. Using the Chen–Stein bound (35) of Theorem 6 and recalling the relation $\delta D = -L$, one infers that for all $f \in \Psi_{\mathrm{TV}}$, it holds that

$$\begin{aligned}\mathbb{E}[Ff(F) - \lambda f(F+1)] &= \mathbb{E}[(F-\lambda)f(F) - \lambda\Delta f(F)]\\ &= \mathbb{E}[\delta(-DL^{-1}F)f(F) - \lambda\Delta f(F)].\end{aligned}$$

Integrating by parts yields

$$E[\delta(-DL^{-1}F)f(F)] = \mathbb{E}[\langle Df(F), -DL^{-1}F\rangle_{L^2(\mu)}],$$

where, by virtue of Lemma 4, $D_zf(F) = f(F + D_zF) - f(F)$. Observe that Lemma 4 implies that, since F takes values in $\mathbb{Z}_+$, then one can always choose a version of D_zF with values in $\mathbb{Z}$, in such a way that $F + D_zF = F_z$ takes values in $\mathbb{Z}_+$. Furthermore, observe that for every $f : \mathbb{Z}_+ \to \mathbb{R}$ and every $k, a \in \mathbb{Z}_+$ such that $k > a$, one has that

$$f(k) = f(a) + \Delta f(a)(k-a) + \sum_{j=a}^{k-1} \Delta^2 f(j)(k-1-j);$$

on the other hand, when $k, a \in \mathbb{Z}_+$ are such that $k < a$,

$$f(k) = f(a) + \Delta f(a)(k-a) + \sum_{j=k}^{a-1} \Delta^2 f(j)(j+1-k).$$

These two relations yield that, for every $k, a \in \mathbb{Z}_+$,

$$|f(k) - f(a) - \Delta f(a)(k-a)| \le \frac{\|\Delta^2 f\|_\infty}{2}\,|(k-a)(k-a-1)|\,.$$

Taking $a = F$ and $k = F_z$, one therefore deduces that

$$D_zf(F) = \Delta f(F)D_zF + R_z,$$

where R_z is a residual random function verifying

$$|R_z| \le \frac{\|\Delta^2 f\|_\infty}{2}\,|D_zF\,(D_zF - 1)|\,, \quad z \in Z.$$

As a consequence,

$$\mathbb{E}[\langle Df(F), -DL^{-1}F\rangle_{L^2(\mu)} - \lambda \Delta f(F)]$$
$$= \mathbb{E}\big[\Delta f(F)(\langle DF, -DL^{-1}F\rangle_{L^2(\mu)} - \lambda)\big] - \mathbb{E}\int_{\mathbb{X}} (R_z \times D_z L^{-1}F)\, \mu(\mathrm{d}z),$$

and the desired conclusion follows by taking absolute values on both sides, as well as by applying the estimates available on $\|\Delta f\|_\infty$ and $\|\Delta^2 f\|_\infty$. Inequality (62) follows from the Cauchy–Schwarz inequality. □

The following statement is a consequence of Theorem 16.

Proposition 3 *Let* $Z \sim \mathrm{Po}(\lambda)$*,* $\lambda > 0$ *and let* $\{F_n : n \geq 1\} \subset \mathrm{dom}\, D$ *be a sequence of random variables with values in* $\mathbb{Z}_+$ *such that* $\mathbb{E}[F_n] \to \lambda$*, as* $n \to \infty$*. Assume that, as* $n \to \infty$*,*

1. $\mathbb{E}\left|\lambda - \langle DF_n, -DL^{-1}F_n\rangle_{L^2(\mu)}\right| \to 0$*, and*
2. $\mathbb{E}\left[\int_{\mathbb{X}} \left|D_z F_n\, (D_z F_n - 1)\, D_z L^{-1} F_n\right| \mu(\mathrm{d}z)\right] \to 0.$

Then, $d_{\mathrm{TV}}(F_n, Z) \underset{n\to\infty}{\to} 0$ *and* F_n *converges in distribution to* Z*.*

Note that in Proposition 3 we do not assume that $\mathbb{E}[F_n] = \lambda$, for every n. In order to apply Theorem 16, one has therefore to use the triangle inequality to write

$$d_{\mathrm{TV}}(F_n, Z) \leq d_{\mathrm{TV}}(F_n, Z_n) + d_{\mathrm{TV}}(Z_n, Z)$$

where Z_n has a Poisson distribution with mean $\lambda_n := \mathbb{E}[F_n]$, and then use the classical fact that $d_{\mathrm{TV}}(Z_n, Z) \leq |\lambda - \lambda_n|$. The effectiveness of Proposition 3 compared to more classical existing methods to prove convergence in law towards a Poisson distribution lies in the fact that it only involves two sequences of mathematical expectations. The so-called method of moments (as the Poisson distribution is determined by its moments, it follows that, in order to prove that a given sequence $\{F_n\}$ converges in distribution to $\mathrm{Po}(\lambda)$, it is sufficient to prove that $\mathbb{E}[F_n^k]$ converges to $\mathbb{E}[\mathrm{Po}(\lambda)^k]$, for every integer $k \geq 1$) is extremely demanding (and very little used) in the framework of Poisson measures. This is mainly due to the fact that the combinatorial structures involved in the so-called diagram formulae (that are mnemonic devices used to compute moments by means of combinatorial enumerations—see [24, Chap. 4]) become quickly too complex to be effectively put into use. When compared to the method of moments, the simplicity of Proposition 3 provides a powerful alternative to computing the moments of all orders of the sequence under consideration.

7.2 Portmanteau Inequalities on the Poisson Space

This section is aimed at presenting recent *portmanteau inequalities* on the Poisson space proved in [5, Theorem 2.1], involving vectors of random variables that are functionals of η (the term *Portmanteau* is used to indicate the encompassing of several results of different nature into one statement). This estimate—which is formally stated in formula (67) below—is expressed in terms of Malliavin operators, and basically allows one to measure the distance between the laws of a general random element and of a random vector whose components are in part Gaussian and in part Poisson random variables. The inequality (67) is a genuine "portmanteau statement"—in the sense that it can be used to directly deduce a number of disparate results about the convergence of random variables defined on a Poisson space (such as a multi-dimensional version of Theorem 16 but also assess convergence to Poisson–Gaussian limits), as well as to recover known ones (such as Theorems 7, 14–16). These results span a wide spectrum of asymptotic behaviors that are dealt with in a completely unified way in these Portmanteau inequalities. Apart from Malliavin calculus, one of the main technical tools in the proof of these inequalities is an interpolation technique used in [1] for proving multidimensional Poisson results.

Before giving a statement of the Portmanteau inequalities, it is necessary to introduce some additional notation. Fix two integers d, m. Observe that, in the discussion to follow, one can take either d or m to be zero, and in this case every expression involving such an index is set equal to zero by convention. The main objects appearing in the upcoming statement of the Portmanteau inequalities are:

– A metric d_1^0 between the laws of two $\mathbb{Z}_+^d \times \mathbb{R}$-valued random vectors X and Y such that $\mathbb{E}\|X\|_{\mathbb{Z}_+^d \times \mathbb{R}}, \mathbb{E}\|Y\|_{\mathbb{Z}_+^d \times \mathbb{R}} < \infty$, is given by

$$d_1^0(X, Y) = \sup_{h \in \mathcal{H}_1^0} |\mathbb{E}(h(X)) - \mathbb{E}(h(Y))|,$$

where $\mathcal{H}_1^0$ indicates the collection of all functions

$$\psi : \mathbb{Z}_+^d \times \mathbb{R} \mapsto \mathbb{R} : (j_1, \ldots, j_d; x) \mapsto \psi(j_1, \ldots, j_d; x)$$

such that ψ is bounded by 1 and, for all $j_1, \ldots, j_d$, the mapping $x \mapsto \psi(j_1, \ldots, j_d; x)$ is in Lip(1).

– A metric d_2^0 between the laws of two $\mathbb{Z}_+^d \times \mathbb{R}^m$—valued random vectors X and Y such that $\mathbb{E}\|X\|_{\mathbb{Z}_+^d \times \mathbb{R}^m}, \mathbb{E}\|Y\|_{\mathbb{Z}_+^d \times \mathbb{R}^m} < \infty$, is given by

$$d_2^0(X, Y) = \sup_{h \in \mathcal{H}_2^0} |\mathbb{E}(h(X)) - \mathbb{E}(h(Y))|,$$

where $\mathcal{H}_2^0$ indicates the collection of all functions

$$\psi : \mathbb{Z}_+^d \times \mathbb{R}^m \mapsto \mathbb{R} : (j_1, \ldots, j_d; x_1, \ldots, x_m) \mapsto \psi(j_1, \ldots, j_d; x_1, \ldots, x_m)$$

such that $|\psi|$ is bounded by 1 and for all $j_1, \ldots, j_d$, the mapping $(x_1, \ldots, x_m) \mapsto \psi(j_1, \ldots, j_d; x_1, \ldots, x_m)$ is bounded and admits continuous bounded partial derivatives up to the order three with $\|g\|_{\text{Lip}} \leq 1$, $\|g''\|_\infty \leq 1$ and $\|g'''\|_\infty \leq 1$.

- A vector $\lambda = (\lambda_1, \ldots, \lambda_d)$ of strictly positive real numbers, as well as a random vector $Z = (Z_1, \ldots, Z_d) \sim \text{Po}_d(\lambda_1, \ldots, \lambda_d)$, that is, the elements of Z are independent and such that Z_i has a Poisson distribution with parameter λ_i, for every $i = 1, \ldots, d$.
- A $m \times m$ covariance matrix $\Sigma = \{\Sigma(i,j) \colon i,j = 1, \ldots, m\}$, and a vector $N = (N_1, \ldots, N_m) \sim \mathscr{N}(0, \Sigma)$, that is, N is a m-dimensional centered Gaussian vector with covariance matrix Σ.
- A vector H defined as the $(d + m)$-dimensional random element $H = (Z, N)$. It is assumed that Z is independent of N and that H is independent of η.
- A vector $F = (F_1, \ldots, F_d)$ of random variables with values in $\mathbb{Z}_+$ such that, for every $i = 1, \ldots, d$, $F_i \in \operatorname{dom} D$ and $\mathbb{E}[F_i] = \lambda_i$.
- A vector $G = (G_1, \ldots, G_m)$ of centered elements of $\operatorname{dom} D$.
- A vector V defined as the $(d + m)$-dimensional random element $V = (F, G)$. Note that, by definition, V is $\sigma(\eta)$-measurable.

Note that the two metrics d_1^0 and d_2^0 induce a topology, on the class of probability distributions, respectively, on $\mathbb{Z}_+^d \times \mathbb{R}$ and $\mathbb{Z}_+^d \times \mathbb{R}^m$, that is stronger than the topology of weak convergence. The following quantities used in the statement of the Portmanteau inequalities are defined in terms of Malliavin operators and play a specific role in the quantification of the distance separating the laws of the vector V and H. After the definition of each quantity of interest, an interpretation of what they measure is provided

$$\alpha_1(\lambda, F) := \sum_{i=1}^d \mathbb{E}\left|\lambda_i - \langle DF_i, -DL^{-1}F_i\rangle_{L^2(\mu)}\right| + \sum_{i=1}^d \mathbb{E}\int_{\mathbb{X}} \left|D_zF_i\,(D_zF_i - 1)\,D_zL^{-1}F_i\right| \mu(\mathrm{d}z). \tag{63}$$

The quantity $\alpha_1(\lambda_d, F)$ has the form $\sum_{i=1}^d a_i$ where each a_i measures the distance between the laws of F_i and Z_i. Note that each a_i is the term appearing in the general bound of Theorem 16, applied to F_i.

$$\alpha_2(F) := \sum_{1\leq i\neq j\leq d} \mathbb{E}\left|\langle DF_i, -DL^{-1}F_j\rangle_{L^2(\mu)}\right| \tag{64}$$

$$+\sum_{1\le i\ne j\le d}\mathbb{E}\int_{\mathbb{X}}\left|D_zF_j\left(D_zF_j-1\right)D_zL^{-1}F_i\right|\mu(\mathrm{d}z)$$

$$+\sum_{1\le j\ne k\le d}\sum_{i=1}^{d}\mathbb{E}\int_{\mathbb{X}}\left|D_zF_jD_zF_kD_zL^{-1}F_i\right|\mu(\mathrm{d}z).$$

The quantity $\alpha_2(F)$ measures the independence between the elements of F. This quantity is clearly specific to the multidimensional case and does not have any one-dimensional equivalent. Note that it vanishes if one takes d to be one.

$$\gamma(\Sigma,G):=\sum_{k,j=1}^{m}\mathbb{E}\left|\Sigma(j,k)-\langle DG_j,-DL^{-1}G_k\rangle_{L^2(\mu)}\right|$$
$$+\mathbb{E}\int_{\mathbb{X}}\left(\sum_{j=1}^{m}\left|D_zG_j\right|\right)^2\left(\sum_{j=1}^{m}\left|D_zL^{-1}G_j\right|\right)\mu(\mathrm{d}z). \qquad (65)$$

The quantity $\gamma(\Sigma,G)$ measures the distance between the laws of G and N, and plays the same role as the bounds appearing in Theorems 14 and 15.

$$\beta(F,G):=\sum_{i=1}^{d}\sum_{j=1}^{m}\mathbb{E}\left\langle|DL^{-1}G_j|,|DF_i|\right\rangle_{L^2(\mu)}. \qquad (66)$$

The quantity $\beta(F,G)$ provides an estimate of how independent F and G are by quantifying the degree of independence between their elements. This quantity is clearly specific to the multidimensional case and does not have any one-dimensional equivalent. Note that it vanishes if one takes $d=1, m=0$ or $d=0, m=1$. A further connection between the quantity $\beta(F,G)$ and the "degree of independence" of F and G (the same can be said about $\alpha_2(F)$ and the degree of dependence of the elements of F) can be obtained by combining the integration by parts formula of Lemma 5 with the standard relation $L=-\delta D$, yielding that, for every $j=1,\dots,m$ and $i=1,\dots,d$,

$$\mathbb{E}\left[\langle DG_j,-DL^{-1}F_i\rangle_{L^2(\mu)}\right]=\mathbb{E}\left[\langle -DL^{-1}G_j,DF_i\rangle_{L^2(\mu)}\right]=\mathrm{Cov}(G_j,F_i).$$

The fact that the dependence structure of the elements of the vector V can be assessed by means of a small number of parameters is a remarkable consequence of the use of the Stein and Chen–Stein methods, as well as of the integration by parts formulae of Malliavin calculus. In general, characterizing independence on the Poisson space is a very delicate (and mostly open) issue.
The following theorem states the Portmanteau inequalities.

Theorem 17 *Let the above assumptions and notation prevail. Then, for every d, m there exists an adequate distance $d_\star(\cdot,\cdot)$, as well as a universal constant K (solely depending on λ and Σ), such that*

$$d_\star(H,V) \leq K\{\alpha_1(\lambda,F) + \alpha_2(F) + \gamma(\Sigma,G) + \beta(F,G)\}, \tag{67}$$

where $d_\star = d_1^0$ if $d \geq 0, m = 1$, $d_\star = d_2^0$ if $d \geq 0, m \geq 2$ and $d_\star = d_{\mathrm{TV}}$ if $d \geq 1, m = 0$. Furthermore, the distances d_1^0 and d_2^0 provide a stronger topology than the one of convergence in distribution on $\mathbb{R}^{d+m}$.

The remarkable fact pointed out in the statement of this theorem is that the above introduced coefficients can be linearly combined in order to measure the overall proximity of the laws of H and V.

Proof Due to its high technicality and length, we only give the key ideas of the proof of Theorem 17. For the complete details of the proof, see [5, Proof of Theorem 3.1]. Even though the statement is multidimensional, an interpolation technique is used in order to deal with each component of the vector V one-dimensionally. Let ψ be a function belonging to the class of functions associated with the three distances appearing in the statement. The main goal is to bound a quantity of the type $|\mathbb{E}(\psi(F,G)) - \mathbb{E}(\psi(Z,N))|$ with the bound (67). We can assess such a quantity in the following way:

$$\begin{aligned}|\mathbb{E}(\psi(F,G)) - \mathbb{E}(\psi(Z,N))| \leq{}& |\mathbb{E}(\psi(F,G)) - \mathbb{E}(\psi(F,N))| \\ &+|\mathbb{E}(\psi(F,N)) - \mathbb{E}(\psi(Z,N))|.\end{aligned}$$

The first step will be to deal with $\mathbb{E}(\psi(F,N)) - \mathbb{E}(\psi(Z,N))$ and the second step will be to assess $\mathbb{E}(\psi(F,G)) - \mathbb{E}(\psi(F,N))$.
We give the general method used to deal with those two steps. For Step 1, the term $\mathbb{E}(\psi(F,N)) - \mathbb{E}(\psi(Z,N))$ can be decomposed further in the following way:

$$\mathbb{E}[\psi(F,N) - \psi(Z,N)] = \sum_{k=1}^{d} \mathbb{E}\left[\psi\left(Z_{(1,k-1)}, F_{(k,d)}, N\right) - \psi\left(Z_{(1,k)}, F_{(k+1,d)}, N\right)\right].$$

Each term appearing in the sum can be assessed independently by using the one-dimensional Chen–Stein method, providing a recursive way of using a one-dimensional method to prove a multidimensional result.
Step 2 is slightly more delicate, as one has to take into account the dependence between F and G. The case $m = 1$ is quite direct using Taylor expansions and the case $m = 2$ relies on an interpolation technique inspired by Arratia et al. [1]. □

References

1. Arratia, R., Goldstein, L., Gordon, L.: Two moments suffice for Poisson approximations: the Chen–Stein method. Ann. Probab. **17**, 9–25 (1989)
2. Bachmann, S., Peccati, G.: Concentration bounds for geometric Poisson functionals: logarithmic Sobolev inequalities revisited. Electron. J. Probab. **21**(6), 1–44 (2016)
3. Bachmann, S., Reitzner, M.: Concentration for Poisson U-statistics: subgraph counts in random geometric graphs. Manuscript (2015)
4. Barbour, A.D., Holst, L., Janson, S.: Poisson Approximation. Oxford University Press, Oxford (1992)
5. Bourguin, S., Peccati, G.: Portmanteau inequalities on the Poisson space: mixed limits and multidimensional clustering. Electron. J. Probab. **19**(66), 1–42 (2014)
6. Chafaï, D.: Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities. J. Math. Kyoto Univ. **44**(2), 325–363 (2004)
7. Chatterjee, S.: A new method of normal approximation. Ann. Probab. **36**, 1584–1610 (2008)
8. Chatterjee, S., Diaconis, P., Meckes, E.: Exchangeable pairs and Poisson approximation. Probab. Surv. **2**, 64–106 (2005)
9. Chen, L.H.Y., Goldstein, L., Shao, Q.M.: Normal Approximation by Stein's Method. Springer, Berlin (2011)
10. Daly, F.: Upper bounds for Stein-type operators. Electron. J. Probab. **13**(20), 566–587 (2008)
11. Dudley, R.M.: Real Analysis and Probability, 2nd edn. Cambridge University Press, Cambridge (2003)
12. Eichelsbacher, P., Thäle, C.: New Berry-Esseen bounds for non-linear functionals of Poisson random measures. Electron. J. Probab. **19**(102), 1–25 (2013)
13. Erhardsson, T.: Stein's method for Poisson and compound Poisson approximation. In: Barbour, A.D., Chen, L.H.Y. (eds.) An Introduction to Stein's Method. Lecture Notes Series. Institute for Mathematical Sciences, National University of Singapore, vol. 4, pp. 61–113. Singapore University Press/World Scientific, Singapore (2005)
14. Hug, D., Last, G., Schulte, M.: Second order properties and central limit theorems for geometric functionals of Boolean models. Ann. Appl. Probab. **26**(1), 73–135 (2016)
15. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and random geometric graphs. Electron. J. Probab. **18**(32), 1–32 (2013)
16. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, II: rescaled kernels, marked processes and geometric U-statistics. Stoch. Process. Appl. **123**(12), 4186–4218 (2013)
17. Lachièze-Rey, R., Peccati, G.: New Kolmogorov bounds for functionals of binomial point processes. Manuscript (2015)
18. Lachièze-Rey, R., Reitzner, M.: U-statistics in stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 229–253. Springer, Cham (2016)
19. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
20. Last, G., Peccati, G., Schulte, M.: Normal approximations on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization. Probab. Theory Relat. Fields (2015, to appear)
21. Nourdin, I., Peccati, G.: Normal Approximations with Malliavin Calculus. From Stein's Method to Universality. Cambridge University Press, Cambridge (2012)
22. Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Springer, Berlin (2006)
23. Peccati, G., Taqqu, M.S.: Central limit theorems for double Poisson integrals. Bernoulli **14**(3), 791–821 (2008)

24. Peccati, G., Taqqu, M.S.: Wiener Chaos: Moments, Cumulants and Diagrams. Bocconi & Springer Series, vol. 1, Springer, Milan (2011)
25. Peccati, G., Thäle, C.: Gamma limits and U-statistics on the Poisson space. ALEA Lat. Am. J. Probab. Math. Stat. **10**(1), 525–560 (2013)
26. Peccati, G., Zheng, C.: Multi-dimensional Gaussian fluctuations on the Poisson space. Electron. J. Probab. **15**(48), 1487–1527 (2010)
27. Peccati, G., Solé, J.L., Taqqu, M.S., Utzet, F.: Stein's method and normal approximation of Poisson functionals. Ann. Probab. **38**(2), 443–478 (2010)
28. Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
29. Penrose, M.D., Yukich, J.E.: Central limit theorems for some graphs in computational geometry. Ann. Appl. Probab. **11**, 1005–1041 (2001)
30. Privault, N.: Combinatorics of Poisson stochastic integrals with random integrals. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 37–80. Springer, Cham (2016)
31. Schulte, M.: Normal approximation of Poisson functionals in Kolmogorov distance. J. Theor. Probab. **29**(1), 96–117 (2016)
32. Schulte, M., Thäle, C.: Poisson point process convergence and extreme values in stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 255–294. Springer, Cham (2016)
33. Wu, L.: A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. Probab. Theory Relat. Fields **118**(3), 427–438 (2000)

U-Statistics in Stochastic Geometry

Raphaël Lachièze-Rey and Matthias Reitzner

Abstract A U-statistic of order k with kernel $f : \mathbb{X}^k \to \mathbb{R}^d$ over a Poisson process η is defined as

$$\sum_{(x_1,\ldots,x_k)} f(x_1, \ldots, x_k),$$

where the summation is over k-tuples of distinct points of η, under appropriate integrability assumptions on f. U-statistics play an important role in stochastic geometry since many interesting functionals can be written as U-statistics, like intrinsic volumes of intersection processes, characteristics of random geometric graphs, volumes of random simplices, and many others. It turns out that the Wiener–Ito chaos expansion of a U-statistic is finite and thus Malliavin calculus is a particularly suitable method. Variance estimates, approximation of the covariance structure, and limit theorems which have been out of reach for many years can be derived. In this chapter we state the fundamental properties of U-statistics and investigate moment formulae. The main object of the chapter is to introduce the available limit theorems.

1 U-Statistics and Decompositions

1.1 Definition

Let $(\mathbb{X}, \mathcal{X})$ be a Polish space, $k \geq 1$, and $f : \mathbb{X}^k \to \mathbb{R}$ be a measurable function. The U-statistic of order k with kernel f over a point set $\xi \in \mathbf{N}(\mathbb{X})$ is 0 if ξ has strictly

R. Lachièze-Rey
Université René Descartes – Paris 5, Laboratoire MAP 5, 45 rue des Saints Pères, 75270 Paris Cedex 06, France
e-mail: raphael.lachieze-rey@parisdescartes.fr

M. Reitzner (✉)
Institut für Mathematik, Universität Osnabrück, Albrechtstraße 28a, 49086 Osnabrück, Germany
e-mail: matthias.reitzner@uni-osnabrueck.de

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_7

less than k points and the formal sum

$$U(f,\xi) = \sum_{\mathbf{x}_k \in \xi^k_{\neq}} f(\mathbf{x}_k)$$

otherwise, where $\xi^k_{\neq}$ is the class of k-tuples $\mathbf{x}_k = (x_1, \ldots, x_k)$ of distinct points from ξ; we recall that $\mathbf{N}(\mathbb{X})$ stands for the class of all locally finite counting measures on $(\mathbb{X}, \mathcal{X})$. Remark that, since the sum is over all such k-tuples, f can be assumed to be symmetric without loss of generality.

An abundant literature deals with the asymptotic study of the random variable $U(f, \tilde{\eta}_p)$ as $p \to \infty$ when $\tilde{\eta}_p$ is a binomial process of intensity p, i.e., a set of p iid variables over $\mathbb{X}$. Here we are concerned with a Poisson process η over $\mathbb{X}$ which intensity measure is a non-atomic locally finite measure μ on $\mathbb{X}$. If $f \in L^1_s(\mu^k) = L^1_s(\mathbb{X}^k; \mu^k)$, the expectation $\mathbb{E}U(|f|, \eta)$ is finite a.s, whence the definition of $U(f, \eta)$ makes sense. We want to point out that f depends on η only via the k-tuples $\eta^k_{\neq}$ but may depend on parameters like the dimension of the space, the intensity measure μ of η, etc.

The Poisson point process can equivalently be introduced as the random measure $\eta = \sum_{x \in \eta} \delta_x$, since μ is assumed to have no atoms. Below we rather adopt the vision of point processes as random point sets, as it eases certain formulations and highlights the geometric point of view in the applications.

1.2 Chaotic Decomposition and Multiple Wiener–Itô Integrals

Recall that $I_n(\cdot)$ is the n-th order multiple Wiener–Itô integral over η defined in [20, Sect. 1.3], and that, by virtue of [20, Theorem 2], every $L^2(\mathbb{P}_\eta)$ functional of η admits a Wiener–Itô decomposition, i.e., a representation as an infinite series of orthogonal multiple Wiener–Itô integrals. The decomposition of a Poisson U-statistic is finite and has been computed in [27, Lemma 3.3].

Theorem 1 *Let $f \in L^1_s(\mu^k)$ such that $U(f, \eta) \in L^2(\mathbb{P}_\eta)$. We have the $L^2(\mathbb{P}_\eta)$-decomposition*

$$U(f, \eta) = \sum_{n=0}^{k} I_n(h_n), \tag{1}$$

where

$$h_n(\mathbf{x}_n) = \binom{n}{k} \int_{\mathbb{X}^n} f(\mathbf{x}_n, \mathbf{x}_{k-n}) \mu^{k-n}(\mathrm{d}\mathbf{x}_{k-n}); \mathbf{x}_n \in \mathbb{X}^n. \tag{2}$$

Furthermore, h_n *is a function of* $L_s^1(\mu^n) \cap L_s^2(\mu^n)$.

Remark 1 Somewhat counterintuitively, $f \in L_s^1(\mu^k) \cap L_s^2(\mu^k)$ does not imply that $\mathbb{E}U(f,\eta)^2 < \infty$ (see [27]), but in most examples f is bounded and has a bounded support, which makes the latter condition automatically satisfied.

As is apparent from Theorem 1, each U-statistic of order k is a finite sum of multiple Wiener–Itô integrals of order $n \leq k$, and it is not difficult to prove that conversely any multiple Wiener–Itô integral of order $n \geq 1$ can be written as a finite sum of U-statistics whose orders are smaller or equal to n. From a formal point of view, it is therefore equivalent to study the asymptotics of finite sums of U-statistics or of finite sums of multiple Wiener–Itô integrals. U-statistics are more likely to appear in applications, but the homogeneity of multiple Wiener–Itô integrals makes them easier to deal with, and some of the Malliavin operators take a particularly intuitive form. Consider for instance the case where $F = I_k(f)$ is a multiple Wiener–Itô integral of order $k \geq 1$, and $f \in L^2(\mu^k)$. The Malliavin derivative in $x \in \mathbb{X}$, Ornstein–Uhlenbeck operator, and inverse Ornstein–Uhlenbeck operator, defined in [20, Sect. 1.5], are well defined for μ-a.e. x and take the following elementary form

$$D_xF = kI_{k-1}(f(x,\cdot)), x \in \mathbb{X}, \quad LF = -kI_k(f), \quad L^{-1}F = -k^{-1}I_k(f). \tag{3}$$

For a U-statistic F, one can still derive $D_xF, LF, L^{-1}F$ using the linearity of these operators and the decomposition (1).

The object of this section is the study of sums of multiple Wiener–Itô integrals whose orders are bounded by some $k \geq 1$. The chaotic decomposition also yields that any $L^2(\mathbb{P}_\eta)$ variable can be approximated by such a sum, allowing us in some cases to pass on limit theorems stated here to infinite sums. The following result gives the first two moments of U-statistics.

Proposition 1 *Let* $f \in L_s^1(\mu^k)$. *Then* $\mathbb{E}|U(f,\eta)| < \infty$ *and*

$$\mathbb{E}U(f,\eta) = \int_{\mathbb{X}^k} f(\mathbf{x}_k)\, \mu^k(\mathrm{d}\mathbf{x}_k). \tag{4}$$

If furthermore $U(f,\eta) \in L^2(\mathbb{P}_\eta)$,

$$\mathrm{Var}(U(f,\eta)) = \sum_{n=1}^{k} n!\|h_n\|^2 \tag{5}$$

where h_n *is given in Theorem 1 and* $\|\cdot\|$ *denotes the usual* $L^2(\mu^n)$*-norm.*

Proof The first statement is a direct consequence of the multivariate Mecke equation, while the second stems from the orthogonality between multiple Wiener–Itô integrals $I_n(h_n), 0 \leq n \leq k$, see [20, Lemma 4]. □

1.3 Hoeffding Decomposition

Assume that μ is a probability distribution. Let $p \geq 1, \tilde{\eta}_p = \{X_1, \dots, X_p\}$ be a family of i.i.d. variables with common distribution μ on $\mathbb{X}$. Given a measurable kernel h over $\mathbb{X}^k, k \geq 1$, the traditional Hoeffding decomposition (see, e.g., Vitale [33]) is written

$$U(h, \tilde{\eta}_p) = k! \binom{p}{k} \sigma_k^p(h) = k! \binom{p}{k} \sum_{m=0}^{k} \binom{k}{m} \sigma_m^p(H_m),$$

where

$$\sigma_m^p(H_m) = \frac{1}{\binom{p}{m}} \sum_{1 \leq i_1 < i_2 < \cdots < i_m \leq p} H_m(X_{i_1}, \dots, X_{i_m}), \quad 0 \leq m \leq k,$$

and each kernel H_m is symmetric and *completely degenerate*, i.e.,

$$\mathbb{E} H_m(x_1, \dots, x_{m-1}, X_m) = \int_{\mathbb{X}} H_m(x_1, \dots, x_{m-1}, y) \mu(\mathrm{d}y) = 0$$

for $\mu^{(m-1)}$-a.e. $x_1, \dots, x_{m-1}$. This property implies in particular the orthogonality of the $\sigma_m^p(H_m)$, $1 \leq n \leq k$. If μ is a probability measure, the H_m are uniquely determined and can be expressed explicitly via an inclusion–exclusion formula

$$H_m(x_1, \dots, x_m) = \sum_{n=0}^{m} (-1)^{m-n} \sum_{1 \leq i_1 < \cdots < i_n \leq m} \binom{k}{n}^{-1} h_n(x_{i_1}, \dots, x_{i_n}), \tag{6}$$

where h_n is defined in (2). As is clear in this last formula, this decomposition is different from (1) because in the latter, the integration is performed with respect to the compensated measure $\eta - \mu$, while in $\sigma_m^p(H_m)$ the compensation occurs in the kernel H_m.

The Hoeffding rank m_1 is defined as the smallest index m such that $\|H_m\| \neq 0$, and we can see through (6) that it is equal to the smallest index n such that $\|h_n\| \neq 0$. We furthermore have $H_{m_1} = \binom{k}{m_1}^{-1} h_{m_1}$. As proved in [8] for binomial processes or [19] for Poisson processes, the stochastic integral of order m_1 dominates the sum in certain asymptotic regimes, and limit theorems for geometric U-statistics can then be derived by studying this term, see Sect. 2.1.2.

1.4 Contraction Operators

Let $f \in L_s^2(\mu^q), g \in L_s^2(\mu^k)$. Let $0 \leq l \leq r \leq \min(q,k)$. The contraction kernel of f and g of index (r,l), denoted $f \star_r^l g$ is a function of $k+q-r-l$ arguments, decomposed in $(\mathbf{x}_{r-l}, \mathbf{y}_{q-r}, \mathbf{z}_{k-r})$, where $\mathbf{x}_{r-l} \in \mathbb{X}^{r-l}, \mathbf{y}_{q-r} \in \mathbb{X}^{q-r}$, and $\mathbf{z}_{k-r} \in \mathbb{X}^{k-r}$. It is properly defined $\mu^{k+q-r-l}$-a.e. by

$$f \star_r^l g(\mathbf{x}_{r-l}, \mathbf{y}_{q-r}, \mathbf{z}_{k-r}) := \int f(\mathbf{x}_l, \mathbf{x}_{r-l}, \mathbf{y}_{q-r}) g(\mathbf{x}_l, \mathbf{x}_{r-l}, \mathbf{z}_{k-r})\, \mu^l(\mathrm{d}\mathbf{x}_l).$$

Contraction operators are used below to assess the distance between a multiple Wiener–Itô integral and the normal law. For conditions ensuring that the contraction functions are well defined everywhere and twice integrable, see for instance [22].

2 Rates of Convergence

Let $F \in L^2(\mathbb{P}_\eta)$ be a random variable of the form

$$F = \sum_{n=0}^{k} I_n(h_n) \tag{7}$$

with kernels $h_n \in L_s^2(\mu^n), n \geq 1$. This model encompasses U-statistics, as outlined by Theorem 1, as well as finite sums of U-statistics and multiple Wiener–Itô integrals.

In applied situations, the set-up consists of a fixed integer $k \geq 1$, a family of measures $\mu_t, t > 0$ on $\mathbb{X}$, and a family of kernels $h_{n,t} \in L_s^2(\mu_t^n)$, $1 \leq n \leq k$, $t > 0$. We study the random variables

$$F_t := \sum_{n=0}^{k} I_n(h_{n,t}), t > 0, \tag{8}$$

where the stochastic integration is performed with respect to μ_t^n. Our limit theorems are about the existence of families $(a_t)_{t>0}, (b_t)_{t>0}$, and of a random variable V such that

$$\tilde{F}_t := \frac{F_t - a_t}{\sqrt{b_t}} \to V$$

as $t \to \infty$ in the weak topology. If not stated otherwise, we consider $a_t = \mathbb{E}F_t, b_t = \mathrm{Var}(F_t)$. In the applications we consider in this chapter, μ_t is either of the form

- $\mu_t = t\mu$ for some reference measure μ on the space $\mathbb{X}$, or
- $\mu_t = 1_{\mathbb{X}_t}\mu$ where $\mathbb{X}_t \subset \mathbb{X}$ depends on t.

The following two settings occur in the most important applications.

If $\eta = \eta_t$ is a Poisson point process on $\mathbb{X} = \mathbb{R}^d$, the measure μ will often be the Lebesgue measure ℓ_d, or, for $\mathbb{X} = \mathbb{R}^d \times M$, a product measure $\mu = \ell_d \otimes \nu$ with a probability measure ν on a topological *marks space* $(M, \mathscr{M})$. See Sect. 2.1.1 for more on marked U-statistics.

If $\eta = \eta_t$ is a Poisson 'flat' process on the Grassmannian $\mathbb{X} = A(d, i)$ of affine i-dimensional subspaces (flats) of $\mathbb{R}^d$, the intensity measure $\mu(\cdot)$ will be a translation invariant measure on $A(d, i)$. The Poisson flat process is only observed in a compact convex window $W \subset \mathbb{R}^d$ with interior points. Thus, we can view η_t as a Poisson process on the set $[W]$ defined by

$$[W] = \{h \in A(d, i) : \; h \cap W \neq \emptyset\} .$$

2.1 Central Limit Theorems

Let F be of the form (7) satisfying $\mathbb{E}F = h_0 = 0$. Let $N \sim \mathcal{N}(0, 1)$, $\sigma^2 = \mathrm{Var}(F)$. The next result gives bounds on the distance between F and N in terms of the contractions between the kernels of F. The result in the Wasserstein distance has been established in [18], and the one in Kolmogorov distance in [9].

Theorem 2 *Put*

$$B(F) = \max\left(\max_1 \|h_n \star_r^l h_m\|, \max_{n=1,\dots,k} \|h_n\|^2_{L^4(\mu^n)} \right)$$

$$B'(F) = \max(|1 - \sigma^2|, B(F), B(F)^{3/2})$$

where $\max_1$ *is over* $1 \le l \le r \le n \le m \le k$ *with* $l \neq m$. *There exists a constant* $C_k > 0$ *not depending on the kernels of* F *such that*

$$d_W(F, N) \le \sigma^{-1} C_k B(F) \tag{9}$$

$$d_K(F, N) \le C_k B'(F). \tag{10}$$

We reproduce here the important steps of the proof for the Wasserstein bound. The main result, due to Peccati et al. [22], is a general inequality on the Wasserstein distance between a Poisson functional F with expectation $\mathbb{E}F = 0$ and variance $\sigma^2 > 0$ and the normal law. The fact that F has a finite chaos expansion and that its kernels are twice integrable yields that the operators $D_x F, LF, L^{-1}F$ are well defined for μ-a.e. x. We have

$$d_W(F,N) \le \frac{1}{\sigma}\sqrt{\mathbb{E}[(\sigma^2 - \langle D_x F, -D_x L^{-1}F\rangle_{L^2(\mu)})^2]} + \frac{1}{\sigma^2}\int_{\mathbb{X}} \mathbb{E}[(D_x F)^2 |D_x L^{-1}F|]\mu(\mathrm{d}x). \tag{11}$$

See the surveys [20] and [4] (in this volume) for a proof and more insights on this result. To translate these inequalities into bounds on the contraction norms, we use the multiplication formula from [22], see also [20], which yields that the multiplication of multiple Wiener–Itô integrals is a linear combination again of multiple Wiener–Itô integrals. For $k, q \ge 1, f \in L_s^2(\mu^q), g \in L_s^2(\mu^k)$,

$$I_q(f)I_k(g) = \sum_{r=0}^{q\vee k} r! \binom{q}{r}\binom{k}{r}\sum_{l=0}^{r}\binom{r}{l} I_{q+p-r-l}(f \tilde{\star}_r^l g), \tag{12}$$

where the symmetrized contraction kernels $f \tilde{\star}_r^l g$ are the average of kernels $f \star_r^l g$ over all possible permutations of the arguments.

If for instance $F = I_k(f)$ is a multiple Wiener–Itô integral of order $k \ge 1$, (3) gives the value of the Malliavin operators, and a computation then yields the bound (9) with $f_k = f; f_i = 0$ for $i \ne k$, see [24, Proposition 5.5]. If F is a general functional with a finite decomposition, such as a U-statistic (see (1)), Malliavin operators are computed using linearity and yield the bound (9), see the proof of Theorem 3.5 in [18].

Concerning Kolmogorov distance, Schulte [31] has derived a Stein's bound similar to (11), but with more terms on the right-hand side (Theorem 1.1), reflecting the effect that test functions are indicator functions, more irregular than the Lipschitz functions involved in Wasserstein distance. This bound was later improved by Eichelsbacher and Thäle [9, Theorem 3.1], reducing the number of additional terms. With similar computations as in the Wasserstein case, one can then prove [9, Theorem 4.1] that these additional terms only add contraction norms $\|f_i \star_l^r f_j\|^{3/2}$ up to a constant, yielding the bound $B'(F)$.

Remark 2 The terms in $B'(F)$ bounding the Kolmogorov distance are smaller than the original terms present in $B(F)$ if the bound goes to 0, and don't change the bound magnitude or its eventual convergence to 0.

Remark 3 The constant C_k explodes as $k \to \infty$. In other papers [21, 27] similar bounds are derived in more specific cases, with a different method. The constants are more tractable and allow for instance to approximate accurately the distance from a Gaussian to an infinite series of multiple Wiener–Itô integrals by that of its truncation at some order (see for instance [30]).

Theorem 3 (Fourth Moment Theorem) *Assume that F is of the form (7) and that the kernels h_n are nonnegative. Then for some $C'_k > 0$*

$$B(F) \leq C'_k \sqrt{\mathbb{E}F^4 - 3\sigma^4}.$$

- In view of (9), the convergence of the fourth moment to that of a Gaussian therefore implies central limit, with a bound for Wasserstein distance. In this case, as noted in [9], using (10) yields a similar bound for Kolmogorov distance. The positiveness of the kernels is adapted to U-statistics with a nonnegative kernel.
- It is remarkable that, in case of an infinite collection $(F_t)_{t>0}$, the convergence of the fourth moment to that of the Gaussian variable as $t \to \infty$ is sufficient for such variables to converge to the normal law. The only technical requirement is that the variables $F_t^4, t > 0$, are uniformly integrable.

Example 1 (De Jong's Theorem) Assume that μ is a probability distribution. Let f_2 be a nonzero degenerate symmetric kernel of $L_s^1(\mu^2)$, i.e., such that

$$\int_{\mathbb{X}} f_2(x,y)\mu(\mathrm{d}x) = 0 \text{ for } \mu\text{-a.e. } y \in \mathbb{X}.$$

This degeneracy property implies that $U(f_2, \eta) = I(f_2)$, we also assume that $f_2 \in L_s^4(\mu^2)$. De Jong [6] derived a fourth moment central limit theorem for binomial U-statistics of the form $U(f_2, \tilde{\eta}_p)$, where $p \in \mathbb{N}$ goes to infinity and $\tilde{\eta}_p$ is a sequence of p iid variables with law μ. In the Poisson framework, (10) yields Berry–Esseen bounds between $F = U(f_2, \eta) = I_2(f_2)$ and N:

$$d_W(\tilde{F}, N) \leq C_2 \frac{1}{\|f_2\|^2} b(f_2)$$
$$d_K(\tilde{F}, N) \leq C_2 \frac{1}{\|f_2\|^2} \max(b(f_2), b(f_2)^{3/2})$$

where

$$b(f_2) = \max\left(\|f_2 \star_2^0 f_2\|, \|f_2 \star_1^1 f_2\|, \|f_2 \star_2^1 f_2\|\right).$$

See Eichelsbacher and Thäle [9, Theorem 4.5] for details. In [23], Peccati and Thäle derive bounds on the Wasserstein distance between such a U-statistic and a target Gamma variable, also in terms of contraction operators.

2.1.1 Local U-Statistics with Marks

For many applications, it is useful to assume that the state space is of the form $S \times M$ where S is a subset of $\mathbb{R}^d$ containing the points t_i of η, and $(M, \mathscr{M})$ is a *marks space*, i.e. a locally compact space endowed with some probability measure ν. The space M contains marks m_i that will be randomly assigned to each point of the process. Define $\mathbb{X} = \mathbb{R}^d \times M, \mu = \ell \otimes \nu$.

For $t > 0$, introduce $\mathbb{X}_t = [-t^{1/d}, t^{1/d}]^d \times M$, $\mu_t = \mathbb{1}_{\mathbb{X}_t}\mu$, and let η_t be a Poisson measure with intensity μ_t. Let f be a real function on $\mathbb{X}$ locally integrable, i.e., such that for every compact $S \subset \mathbb{R}^d$, $\int_{S\times M} |f| \mathrm{d}\mu^k < \infty$. Assume also that f is a spatially stationary function of $\mathbb{X}$, i.e., such that for μ^k-almost all $(\mathbf{t}_k, \mathbf{m}_k) \in \mathbb{X}^k, z \in \mathbb{R}^d$,

$$f(\mathbf{t}_k + z, \mathbf{m}_k) = f(\mathbf{t}_k, \mathbf{m}_k), \tag{13}$$

where $\mathbf{t}_k + z$ is the result of the addition of z to each member of $\mathbf{t}_k$. We consider the U-statistic $F_t = U(f, \eta_t)$, well defined for each t. The tail behavior of the function f is fundamental regarding the asymptotic behavior of F_t as $t \to \infty$.

Definition 1 Let $f : (\mathbb{R}^d \times M)^k \to \mathbb{R}$ locally integrable. Then f is *rapidly decreasing* if it is stationary and satisfies the following integrability condition: There exists a non-vanishing probability density κ on $(\mathbb{R}^d)^{k-1}$ such that for $p = 2, 4$,

$$A_p(f) = \int\limits_{(\mathbb{R}^d)^{k-1}\times M^k} f(0, \mathbf{t}_{k-1}, \mathbf{m}_k)^p \kappa(\mathbf{t}_{k-1})^{1-p} \ell_d^{k-1}(\mathrm{d}\mathbf{t}_{k-1}) \nu^k(\mathrm{d}\mathbf{m}_k) < \infty.$$

The slight abuse of notation $f(0, \mathbf{t}_{k-1}, \mathbf{m}_{k-1})$ means that $\mathbf{t}_k = (0, \mathbf{t}_{k-1}) = (0, t_2, \ldots, t_{k-1})$, and $\mathbf{m}_k = (m_1, \ldots, m_k)$.

We have in this case the following result, which is a consequence of Theorem 6.2 and Example 2.12-(ii) in [19]:

Theorem 4 *Let $F_t = U(f, \eta_t)$ where f is a rapidly decreasing locally integrable function. Then we have for some $C_1, C_2, C_3 > 0$ not depending on t,*

$$C_1 t \le \mathrm{Var}(F_t) \le C_2 t$$
$$d_W(\tilde{F}_t, N_1) \le C_3 t^{-1/2}.$$

Remark 4 Reitzner and Schulte [27] first established this result in the case where f is the indicator function of a ball of $\mathbb{R}^d$ (any non-vanishing continuous density κ can be chosen in this case because $f(0, \cdot)$ has a compact support).

Remark 5 A similar result holds if F is simply assumed to be a finite sum of multiple Wiener–Itô integrals whose kernels are rapidly decreasing functions, the U-statistics being a particular case.

2.1.2 Geometric U-Statistics

Coming back to the general framework, let $\mu_t = t\mu$, and η_t a Poisson process with intensity measure μ_t on $\mathbb{X}$. Let $F_t = U(f, \eta_t)$ where $f \in L_s^1(\mu^k)$ is fixed and is such that F has a finite variance. Then $F_t - \mathbb{E}F_t$ admits the decomposition (8) with the kernels $h_n = h_{n,t}, n = 1, \ldots, k$, defined by

$$h_{n,t}(\mathbf{x}_n) = t^{k-n}\binom{k}{n}\int_{\mathbb{X}^{k-n}} f(\mathbf{x}_n, \mathbf{x}_{k-n})\mu^{k-n}(\mathrm{d}\mathbf{x}_{k-n}),\ \mathbf{x}_n \in \mathbb{X}^n.$$

An important feature is the Hoeffding rank of the U-statistic

$$n_1 := \inf\{n : \|h_{n,t}\| \neq 0\},$$

which does not depend on t. The variance expression (4) yields that $I_{n_1}(h_{n_1,t})$ is the predominant term in (8), in the sense that $F_t - I_{n_1}(h_{n_1,t}) = o(F_t)$ for the $L^2(\mathbb{P}_\eta)$ norm as $t \to \infty$. It yields the following result (Theorem 7.3 in [19]).

Theorem 5 *Let* $\tilde{F}_t = \mathrm{Var}(F_t)^{-1/2}(F_t - \mathbb{E}F_t)$*. For some* $C_1, C_2, C_3 > 0$ *not depending on t,*

$$C_1 t^{2k-n_1} \leq \mathrm{Var}(F_t) \leq C_2 t^{2k-n_1}.$$

(i) If $n_1 = 1$*,* $U(f, \eta_t)$ *follows a central limit theorem and*

$$d_W(\tilde{F}_t, N) \leq C_3 t^{-1/2},$$
$$d_K(\tilde{F}_t, N) \leq C_3 t^{-1/2}.$$

(ii) If $n_1 > 1$*,* $U(f, \eta_t)$ *does not follow a CLT and* $\tilde{F}_t$ *converges to a Gaussian chaos of order* n_1 *(see* [19, Theorem 7.3–2]*).*

This result is also a particular case of Theorem 6.

Remark 6 Point (i) first appears in [27].

Remark 7 Point (ii) crucially uses the results of Dynkin and Mandelbaum [8].

Remark 8 The speed of convergence to the Gaussian chaos in (ii) is studied by Peccati and Thäle [23] in case the limit is a Gamma distributed random variable.

2.1.3 Regimes Classification

The crucial difference in Theorems 4 and 5 is the area of influence of a given point $x \in \eta_t$. In the case of a local U-statistic, for any $t > 0$, a typical point $x \in \eta_t$ interacts with other points of η_t via the contributions $f(x, x_1, \ldots, x_{k-1})$ for points x_i of η_t. Given that f satisfies (13), $f(x, x_1, \ldots, x_{k-1})$ is likely to be small if one of the x_i is far from x, independently of t. Therefore, the points of η which interaction with x via the kernel f gives a significant contribution for t large will be near x. The situation is different for a geometric U-statistic, where a point of η_t interacts with any other point via f, regardless of their distance. Both these regimes can be seen as two particular cases of a continuum.

Let $f : \mathbb{X}^k \to \mathbb{R}$ be a rapidly decreasing function. Let $\alpha_t > 0$, defined to be the *scaling factor*, $\mathbb{X}_t = [-t^{1/d}, t^{1/d}]^d \times M$, $\mu_t = \mathbb{1}_{\mathbb{X}_t}\ell_d \otimes \nu$. Let f_t be the kernel obtained by rescaling with the stationary function f,

$$f_t(\mathbf{x}_k) = f(\alpha_t \mathbf{x}_k), \mathbf{x}_k \in \mathbb{X}_t^k, \tag{14}$$

and $F_t = U(f_t, \eta_t)$, where η_t is the Poisson measure with intensity μ_t.

Say that f has *non-degenerate projections* if none of the functions

$$f_n(\mathbf{x}_n) = \int_{\mathbb{X}_t^{k-n}} f(\mathbf{x}_n, \mathbf{x}_{k-n}) \mathrm{d}\mu_t^{k-n}, \mathbf{x}_n \in \mathbb{X}^n,$$

well defined in virtue of (13) is μ-a.e. equal to 0. It is trivially the case if for instance $\|f\| \neq 0$ and $f \geq 0$ μ-a.e. Concerning notation, every spatial transformation of a point $x = (t, m) \in \mathbb{R}^d \times M$, such as translation, rotation, or multiplication by a scalar, is only applied to the spatial component t.

Subsequently, any spatial transformation applied to a k-tuple of points $\mathbf{x}_k = (x_1, \ldots, x_k)$ is applied to the spatial components of the x_i's. The quantity $v_t = \alpha_t^{-d}$ is relevant because it gives the magnitude of the number of points interacting with a typical point x. The case $v_t = \alpha_t = 1$ is that of local U-statistic. If $v_t = t$ is roughly the volume of $\mathbb{X}_t$, it corresponds to geometric U-statistics. In the latter case it is useless to assume that f is rapidly decreasing, as only the behavior of f over $\mathbb{X}_1$ is relevant.

Theorem 6 *Assume that f_t is of the form (14), where f is a rapidly decreasing function with nondegenerate projections. With the notation above, there are $C_1, C_2, C_3 > 0$ such that*

$$C_1 \leq \frac{\mathrm{Var}(F_t)}{t v_t^{2k-2} \max(1, v_t^{-k+1})} \leq C_2,$$

and

$$d_W(\tilde{F}_t, N) \leq C_3 t^{-1/2} \max(1, v_t^{-k+1})^{1/2}$$
$$d_K(\tilde{F}_t, N) \leq C_3 t^{-1/2} \max(1, v_t^{-k+1})^{1/2}.$$

Concerning the bound for Kolmogorov distance, it is not formally proved in the literature. It relies on the fact that in Theorem 2, $B'(F) \leq CB(F)$ for some $C > 0$ in the case where $\sigma \to 1$ and $B(F) \to 0$. Then one can simply reproduce the proof of [19], entirely based on an upper bound for $B(F)$.

Remark 9 Theorems 4 and 5-(i) can be retrieved from this theorem by setting, respectively, $v_t = 1$ or $v_t = t$.

Remark 10 If some projections do vanish, the convergence might be modified, and the limit might not even be Gaussian, as it is the case for the degenerate geometric U-statistics of Theorem 5-(ii).

Remark 11 Depending on the asymptotic behavior of v_t, we can identify four different regimes:

1. **Long interactions**: $v_t \to \infty$, CLT at speed $t^{-1/2}$, the first chaos $I_{1,t}(h_{t,1})$ dominates (geometric U-statistics).
2. **Constant size interactions**: $v_t = 1$, CLT at speed $t^{-1/2}$, all chaoses have the same order of magnitude (local U-statistics).
3. **Small interactions**: $v_t \to 0, tv_t^{-k+1} \to \infty$, CLT at speed $(tv_t^{-k+1})^{-1/2}$, higher order chaoses dominate. In the case of random graphs ($k = 2$), the corresponding bound in $(tv_t)^{-1/2}$ has been obtained in [18].
4. **Rare interactions**: $tv_t^{-k+1} \to c < \infty$, the bound does not converge to 0. In the case $k = 2$, it has been shown in [18] that there is no CLT but a Poisson limit in the case $c > 0$ (see [4, Chap. 6] in this book for more on Poisson limits).

2.2 Other Limit Laws and Multi-Dimensional Convergence

Besides the Gaussian chaoses appearing in Theorem 5-(ii), some characterizations of non-central limits have also been derived for Poisson U-statistics.

2.2.1 Multidimensional Convergence

We consider in this section the conjoint behavior of random variables $F_t = (F_{1,t}, \ldots, F_{k,t})$ where, for $t > 0, 1 \le m \le k, q_m \ge 1$, and $F_{m,t} = I_{q_m}(h_{m,t})$ for some $h_{m,t} \in L_s^2(\mu^{q_m})$.

Any candidate for the limit of F_t which is in $L^2(\mathbb{P}_\eta)$ should have covariance matrix

$$C_{m,n} = \lim_{t\to\infty} \mathbb{E} F_{m,t} F_{n,t}, \quad 1 \le m, n \le k,$$

if the limit exists. In this case there is indeed asymptotic normality if all contraction norms

$$\|h_{m,t} \star_r^l h_{n,t}\|$$

go to 0 as $t \to \infty$, for $0 \le l \le r \le q_n \le q_m$ with $l \ne q_m, r \ne 0$, see [24, Theorem 5.8]. Explicit bounds on the speed of convergence with a specific distance related to thrice differentiable functions on $(\mathbb{R}^d)^k$ are contained in [4, Chap. 6] in this book, and the convergence is stable, in the sense of Bourguin and Peccati [3].

If now $F_t = (F_{1,t}, \dots, F_{k,t})$ where each $F_{m,t}$ is a U-statistic, one can consider the random vector G_t constituted by all multiple Wiener–Itô integrals with respect to kernels from the decompositions of the $F_{m,t}$, as defined in (2). One can then infer conditions for asymptotic normality of F_t by applying the previous considerations to G_t.

As noted in Remark 11, some U-statistics behave asymptotically like Poisson variables. Asymptotic joint laws of U-statistics can also converge to random vectors with marginal Poisson laws, and it can also happen that they converge to an hybrid random vector which has both Gaussian and Poisson marginals, here again the reader is referred to the survey [4], in this volume.

2.2.2 Gamma

Similar results to those of Sect. 2.1 with Gamma limits have been derived by Peccati and Thäle [23] for Poisson chaoses of even order. The distance used there is

$$d_3(U,V) = \sup_{h \in \mathcal{H}^3} |\mathbb{E}[h(U) - h(V)]|$$

for two random variables U, V, where $\mathcal{H}^3$ is the class of functions of class $\mathcal{C}^3$ with the first three derivatives uniformly bounded by 1. We again denote by $f \tilde{\star}_r^l g$ the symmetrized contraction kernels .

For $\nu > 0$, let $F(\nu/2)$ be a Gamma distribution with mean and variance both equal to $\nu/2$. We introduce the centered unit variance variable $G(\nu) := 2F(\nu/2) - \nu$.

Theorem 7 *Let $F = I_k(h_k)$ for some even integer $k \geq 2$. We have*

$$d_3(I_k(h_k), G(\nu)) \leq D_k \max\{k!\|h_k\| - 2\nu; \|h_k \star_p^p h_k\|; \|h_k \star_r^l h_k\|^{1/2}; \|h_k \tilde{\star}_{q/2}^{q/2} h_k - c_k h_k\|\}$$

where the maximum is taken over all $p = 1, \dots, k-1$ such that $p \neq k/2$ and all (r,l) such that $r \neq l$ and $l = 0$, or $r \in \{1, \dots, k\}$ and $l \in \{1, \dots, \min(r, k-1)\}$. Also

$$c_k = \frac{4}{(q/2)!\binom{q}{q/2}^2}.$$

Remark 12 In the case of double Wiener–Itô integrals ($k = 2$), the authors of [23] provide a fourth moment theorem, in the sense that under some technical conditions, a sequence of double Wiener–Itô integrals converges to a Gamma variable if their first moments converge to that of a Gamma variable.

Remark 13 This result enables to give an upper bound on the speed of convergence to the second Gaussian chaos in Theorem 5 in the case $n_1 = 2$, if this limit is indeed a Gamma variable.

Remark 14 The case $q = 4$ has also been settled in a recent paper by Fissler and Thäle [10].

3 Large Deviations

There are only few investigations concerning concentration inequalities for Poisson U-statistics. Most results require a nice bound on $\sup_{\eta\in\mathbf{N}(\mathbb{X}),\, z\in\mathbb{X}} D_zF(\eta) < \infty$. For U-statistics of order ≥ 2 this condition is usually not satisfied, even if the kernel f is bounded. For U-statistics of order 1, this holds if $\|f\|_\infty < \infty$. Therefore we split our investigations into a section on U-statistics of order one and another one on higher order local U-statistics. We start with a general result. Throughout this section we assume that $f \geq 0$ and $f \neq 0$.

3.1 A General LDI

In this section we sketch an approach developed in [28] leading to a general concentration inequality. Here it is necessary to view η as a random counting measure $\sum \delta_x$ (and continue writing $x \in \eta$ if x is in the support of η). For two counting measures η and ν we define the difference $\eta\backslash\nu$ by

$$\eta\backslash\nu = \sum_{x\in\mathbb{X}} (\eta(\{x\}) - \nu(\{x\}))_+ \,\delta_x\,. \tag{15}$$

For $x \in \eta$ and $f \in L^1_s(\mu^k)$, we recall that

$$U(f,\eta) = \sum_{x\in\eta} F(x,\eta) \quad \text{with} \quad F(x,\eta) = \sum_{\mathbf{x}_{k-1}\in(\eta\backslash\{\mathbf{x}\})^{k-1}_{\neq}} f(x,\mathbf{x}_{k-1}).$$

Assume that in addition to η a second point set $\zeta \in \mathbf{N}_s(\mathbb{X})$ is chosen. The non-negativity of f yields

$$\begin{aligned} U(f,\eta) &\leq U(f;\zeta) + k\sum_{x\in\eta} F(x,\eta)\mathbb{1}(x \notin \zeta) \\ &= U(f;\zeta) + k\int F(x,\eta)\,\mathrm{d}(\eta\backslash\zeta)\,. \end{aligned}$$

The convex distance of a finite point set $\eta \in \mathbf{N}_s(\mathbb{X})$ to some $A \subset \mathbf{N}_s(\mathbb{X})$ was introduced in [26], and is given by

$$d(\eta, A) = \max_u \min_{\zeta \in A} \int u \, \mathrm{d}(\eta \backslash \zeta),$$

where for given η the maximum is taken with respect to all nonnegative measurable functions $u : \mathbb{X} \to \mathbb{R}_+$ satisfying $\|u\|_{2,\eta}^2 = \int u^2 \mathrm{d}\eta \leq 1$. To link the convex distance to the U-statistic, we insert for u the normalized function $\|F(x,\eta)\|_{2,\eta}^{-1} F(x,\eta)$ and rewrite $U(f,\eta)$ in terms of the convex distance as follows:

$$\begin{aligned} d(\eta, A) &\geq \min_{\zeta \in A} \int \frac{1}{\|F(x,\eta)\|_{2,\eta}} F(x,\eta) \mathrm{d}(\eta \backslash \zeta) \\ &\geq \frac{1}{k\|F(x,\eta)\|_{2,\eta}} \min_{\zeta \in A} \Big(U(f,\eta) - U(f;\zeta) \Big). \end{aligned}$$

If we assume $F(x,\eta) \leq B$ for some $B \in \mathbb{R}$, then $\|F(x,\eta)\|_{2,\eta}^2 \leq B \sum_{x \in \eta} F(x,\eta) = BU(f,\eta)$, which implies

$$d(\eta, A) \geq \frac{1}{k\sqrt{B}} \min_{\zeta \in A} \frac{U(f,\eta) - U(f;\zeta)}{\sqrt{U(f;\eta)}} \, \mathbb{1}(\forall x \in \eta : \; F(x,\eta) \leq B) \, . \tag{16}$$

In [26], a LDI for the convex distance was proved. For η a Poisson point process on some lcscH[1] space $\mathbb{X}$ with finite intensity measure, and for $A \subset \mathbf{N}(\mathbb{X})$, we have

$$\mathbb{P}(A)\mathbb{P}\left(d(\eta, A) \geq s\right) \leq \exp\left(-\frac{s^2}{4}\right), \; s \geq 0.$$

Precisely as in [28], this concentration inequality combined with the estimate (16) yields the following theorem.

Theorem 8 *Assume that $\varepsilon(\cdot)$ and $B \in \mathbb{R}$ satisfy $\mathbb{P}(\exists x \in \eta : \; F(x,\eta) > B) \leq \varepsilon(B)$. Let m be the median of $U(f,\eta)$. Then for $u \geq 0$*

$$\mathbb{P}(|U(f,\eta) - m| \geq u) \leq 4 \exp\left(-\frac{u^2}{4k^2 B(u+m)}\right) + 3\varepsilon(B) \, . \tag{17}$$

In the next sections we apply this to U-statistics of order one and to local U-statistics. In the applications, the crucial ingredient is a good estimate for $\varepsilon(B)$.

[1]Locally compact second countable Hausdorff space.

3.2 LDI for First Order U-Statistics

There are several concentration inequalities for integrals of functions $f \in L^1(\mu)$ over Poisson point processes, i.e., U-statistics of order one,

$$U(f,\eta) = \sum_{x\in\eta} f(x) = \int f \mathrm{d}\eta, f \geq 0,$$

in which case $D_z U = f(z)$, $z \in \mathbb{X}$. Assuming that $\|f\|_\infty = B < \infty$, we have a.s.

$$\|D_z U\|_\infty = B.$$

A result by Houdre and Privault [12] shows that for any σ-compact metric space $\mathbb{X}$

$$\mathbb{P}(U - \|f\|_1 \geq u) \leq \exp\left(-\frac{\|f\|_1}{\|f\|_\infty} g\big(\frac{u}{\|f\|_1}\big)\right) \tag{18}$$

where $g(u) = (1+u)\ln(1+u) - u, \ u \geq 0$ and because $f \geq 0$ the 1-norm equals the expectation $\mathbb{E}U(f,\eta)$. A similar result is due to Ane and Ledoux [1]. Reynaud-Bouret [29] proves an estimate involving the 2-norm $\|f\|_2$ instead of the 1-norm. A slightly more general estimate is given by Breton et al. [5].

We could also make use of Theorem 8 and choose $B = \|f\|_\infty$. This yields

$$\mathbb{P}(|U(f,\eta) - m| \geq u) \leq 4\exp\left(-\frac{u^2}{4\|f\|_\infty(u+m)}\right), \tag{19}$$

which is a slightly weaker estimate than (18).

3.3 LDI for Local U-Statistics

In this paragraph we assume that $\mathbb{X}$ is equipped with a norm $\|\cdot\|$ and $B(x,r)$ denotes the ball of radius r around $x \in \mathbb{X}$. We call $U(f,\eta)$ a local U-statistic of radius r if $f(x_1,\dots,x_k) = 0$ for $\max_{i\neq j}\|x_i - x_j\| > r$. We have

$$F(x,\eta) \leq \|f\|_\infty \eta(B(x,r))^{k-1}$$

and

$$\mathbb{P}(\exists x: \ F(x,\eta) > B) \leq \mathbb{E}\sum_{x\in\eta} \mathbb{1}(F(x,\eta) > B) = \int_{\mathbb{X}} \mathbb{P}(F(x,\eta) > B)\mathrm{d}\mu.$$

And it remains to estimate

$$\mathbb{P}\left(\eta(B(x,r)) > \left(\frac{B}{\|f\|_\infty}\right)^{\frac{1}{k-1}}\right).$$

We use the Chernoff bound for the Poisson distribution, namely

$$\mathbb{P}(\eta(B^d(x,r)) > r) \leq \inf_{s\geq 0} e^{E(e^s-1)-sr}, \tag{20}$$

since $\eta(B(x,r))$ is a Poisson distributed random variable with mean

$$\mathbb{E}\eta(B^d(x,r)) = \mu(B^d(x,r)) \leq \sup_{x\in\mathbb{X}} \mu(B^d(x,r)) =: E\,. \tag{21}$$

Because $\inf_{s\geq 0} E(e^s-1) - sr = r(1-\ln(r/E)) - E$ we estimate the right-hand side of (20) by $\exp\left(-\frac{1}{2}r\right)$ for $Ee^2 \leq r$. This leads to

$$\mathbb{P}(\exists x:\ F(x,\eta) > B) \leq \mu(\mathbb{X}) \exp\left(-\frac{1}{2}\left(\frac{B}{\|f\|_\infty}\right)^{\frac{1}{k-1}}\right) := \varepsilon(B)$$

for $B \geq E^{k-1}e^{2(k-1)}\|f\|_\infty$. We set $B = \|f\|_\infty^{\frac{1}{k}}(\frac{u^2}{(u+m)})^{\frac{k-1}{k}}$ and combine this with the general Theorem 8.

Theorem 9 *Set* $E := \sup_{x\in\mathbb{X}} \mu(B^d(x,r))$. *Then for* $\frac{u^2}{(u+m)} \geq E^k e^{2k}\|f\|_\infty$,

$$\mathbb{P}(|U(f,\eta) - m| \geq u) \leq 4\mu(\mathbb{X}) \exp\left(-\frac{1}{4k^2}\|f\|_\infty^{-\frac{1}{k}}\left(\frac{u^2}{u+m}\right)^{\frac{1}{k}}\right).$$

Clearly, in particular situations more careful choices of $\varepsilon(B)$ and B lead to more precise bounds.

4 Applications

In this section we investigate some applications of the previous theorems in stochastic geometry. In all these cases $\mathbb{X}$ is either a subset of $\mathbb{R}^d$ or a subset of the affine Grassmannian $A(d,i)$, the space of all i-dimensional spaces in $\mathbb{R}^d$, endowed with the usual hit-and-miss topology and Borel σ-algebra.

We state some normal approximation and concentration results which follow from the previous theorems. In many cases multi-dimensional convergence and convergence to other limit distributions can be proved in various regimes. We restrict our presentation to certain 'simple' cases without making any attempt for completeness. Our aim is just to indicate recent trends, we refer to further results and investigations in the literature.

4.1 Intersection Process

Let η_t be a Poisson process on the space $A(d,i)$ with an intensity measure of the form $\mu(\cdot) = t\theta(\cdot)$ with $t \in \mathbb{R}^+$ and a σ-finite non-atomic measure θ. The Poisson flat process is only observed in a compact convex window $W \subset \mathbb{R}^d$ with interior points. Thus, we can also view η_t as a Poisson process on the set $\mathbb{X} = [W]$ defined by

$$[W] = \{L \in A(d,i) : L \cap W \neq \emptyset\}.$$

Given the i-flat process η_t, we investigate the $(d - k(d-i))$-flats in W which occur as the intersection of k flats of η_t. Hence we assume $k \leq d/(d-i)$. They form the intersection process $\eta_t^{(k)}$, see [13, Sect. 3.3.1]. In particular, we are interested in the sum of the j-th intrinsic volumes given by

$$\Phi_t = \Phi_t(W,i,k,j) = \frac{1}{k!} \sum_{(L_1,\ldots,L_k)\in\eta^k_{t,\neq}} V_j(L_1 \cap \ldots \cap L_k \cap W)$$

for $j = 0,\ldots,d-k(d-i)$, $i = 0,\ldots,d-1$ and $k = 1,\ldots,\lfloor d/(d-i)\rfloor$. Here one has to restrict the sum to those k-tuples of i-flats that are in general position. This is necessary in case of a discrete directional distribution, for example.

For the definition of the j-th intrinsic volume $V_j(\cdot)$ we refer to [13]. We remark that $V_0(K)$ is the Euler characteristic of the set K, and that $V_n(K)$ of an n-dimensional convex set K is the Lebesgue measure $\ell_n(K)$. Thus $\Phi_t(W,i,1,0)$ is the number of flats in W and $\Phi_t(W,i,k,d-k(d-i))$ is the $(d-k(d-i))$-volume of their intersection process.

To ensure that the expectations of these random variables are neither 0 nor infinite, we assume that $0 < \theta([W]) < \infty$, and that $2 \leq k \leq \lfloor d/(d-i)\rfloor$ independent random i-flats on $[W]$ with probability measure $\theta(\cdot)/\theta([W])$ intersect in a $(d-k(d-i))$-flat almost surely and their intersection flat hits the interior of W with positive probability. For example, these conditions are satisfied if the flat process is stationary and isotropic.

The fact that the summands in the definition of Φ_t are bounded and have a bounded support makes sure that all moment conditions are satisfied and we can apply Theorem 5. Denote by $\tilde{\Phi}_t$ the normalized version of Φ_t.

Theorem 10 *Let N be a standard Gaussian random variable. Then constants $c = c(W, i, k, j)$ exist such that*

$$d_W(\tilde{\Phi}_t, N) \leq ct^{-1/2},$$
$$d_K(\tilde{\Phi}_t, N) \leq ct^{-1/2},$$

for $t \geq 1$.

Furthermore, it can be shown [27] that the asymptotic variances satisfies $\operatorname{Var}\Phi_t = C_\Phi t^{2k-1}(1 + o(1))$ as $t \to \infty$ with a constant $C_\Phi = C_\Phi(W, i, k, j)$. The order of magnitude already follows from the first part of Theorem 5.

For more information we refer to [11, 21]. In the second paper the Wiener–Itô chaos expansion is used to derive even multivariate central limit theorems in an increasing window for much more general functionals Φ_t.

4.2 Flat Processes

For $i < \frac{d}{2}$ two i-dimensional planes in general position will not intersect. Thus the intersection process described in the previous section will be empty with probability one. A natural way to investigate the geometric situation in this setting is to ask for the distances between these i-dimensional flats, or more general for the so-called proximity functional already introduced in [13]. The central limit theorems described in the following fits precisely into the setting of this contribution, we refer to [32] for further results.

Let η_t be a Poisson process on the space $A(d, i)$ with an intensity measure of the form $\mu_t(\cdot) = t\theta(\cdot)$ with $t \in \mathbb{R}^+$ and a σ-finite non-atomic measure θ. The Poisson flat process is observed in a compact convex window $W \subset \mathbb{R}^d$. For two i-dimensional planes in general position there is a unique segment $[x_1, x_2]$ with

$$d(L_1, L_2) = \|x_2 - x_1\| = \min_{y \in L_1, z \in L_2} \|z - y\|.$$

The midpoints $m(L_1, L_2) = \frac{1}{2}(x_1 + x_2)$ form a point process of infinite intensity, hence we restrict this to the point process

$$\{m(L_1, L_2) : \ d(L_1, L_2) \leq \delta, L_1, L_2 \in \eta^2_{\neq}\}$$

and are interested in the number of midpoints in W

$$\Pi_t = \Pi_t(W, \delta) = \frac{1}{2} \sum_{L_1, L_2 \in \eta^2_{\neq}} \mathbb{1}(d(L_1, L_2) \leq \delta, m(L_1, L_2) \in W).$$

It is not difficult to show that $\mathbb{E}\Pi_t$ is of order $t^2\delta^{d-2i}$. The U-statistic Π_t is local on the space $A(d,i)$. Thus the following theorem due to Schulte and Thäle [32] is in spirit similar to Theorem 4. Denote by $\tilde{\Pi}_t$ the normalized version of Π_t.

Theorem 11 *Let N be a standard Gaussian random variable. Then constants $c(W,i)$ exist such that*

$$d_K(\tilde{\Pi}_t, N) \leq c(W,i) t^{-\frac{d-i}{2}},$$

for $t \geq 1$.

Moreover, Schulte and Thäle proved that the ordered distances form after suitable rescaling asymptotically an inhomogeneous Poisson point process on the positive real axis. There is a generalization of Theorem 11 including powers and directional constrains in a recent paper of Hug et al. [14].

We add to this a concentration inequality which follows immediately from Theorem 9. Observe that $\mu_t(\mathbb{X}) = t\theta([W])$, and denote by $B^d(h,\delta)$ a ball with center in h and radius δ.

Theorem 12 *Denote by m_t the median of Π_t. Then*

$$\mathbb{P}(|\Pi_t - m_t| \geq u) \leq 4t\theta([W]) \exp\left(-\frac{1}{16}\frac{u}{\sqrt{u+m_t}}\right)$$

for $\frac{u}{\sqrt{u+m_t}} \geq e^2 t \sup_{h\in[W]} \theta(B^d(h,\delta))$.

4.3 Gilbert Graph

Let η_t be a Poisson point process on $\mathbb{R}^d$ with an intensity-measure of the form $\mu_t(\cdot) = t\ell_d(\cdot \cap W)$, where ℓ_d is Lebesgue measure and $W \subset \mathbb{R}^d$ a compact convex set with $\ell_d(W) = 1$. Let $(r_t : t > 0)$ be a sequence of positive real numbers such that $r_t \to 0$, as $t \to \infty$. The random geometric graph is defined by taking the points of η_t as vertices and by connecting two distinct points $x, y \in \eta_t$ by an edge if and only if $\|x-y\| \leq r_t$. The resulting graph is called Gilbert graph.

There is a vast literature on the Gilbert graph and one should have a look at Penrose's seminal book [25]. More recent developments are due to Bourguin and Peccati [3], Lachièze-Rey and Peccati [18, 19], and Reitzner et al. [28].

In a first step one is interested in the number of edges

$$N_t = N_t(W, r_t) = \frac{1}{2}\sum_{(x,y)\in\eta^2_{t,\neq}} \mathbb{1}(\|x-y\| \leq r_t)$$

of this random geometric graph. It is natural to consider instead of the norm functions $\mathbb{1}(f(y-x) \le r_t)$ and instead of counting more general functions $g(y-x)$:

$$\sum_{(x,y)\in\eta^2_{t,\neq}} \mathbb{1}(f(y-x) \le r_t)\, g(y-x).$$

For simplicity we restrict our investigations in this survey to the number of edges N_t in the thermodynamic setting where tr_t^d tends to a constant as $t \to \infty$. Further results for other regimes, multivariate limit theorems, and sharper concentration inequalities can be found in Penrose's book and the papers mentioned above.

Because of the local definition of the Gilbert graph, N_t is a local U-statistic. Theorem 6 with $\nu_t = tr_t^d$ can be applied.

Theorem 13 *Let N be a standard Gaussian random variable. Then constants $c(W)$ exist such that*

$$d_W(\tilde{N}_t, N) \le c(W)t^{-1/2},$$
$$d_K(\tilde{N}_t, N) \le c(W)t^{-1/2},$$

for $t \ge 1$.

A concentration inequality follows immediately from Theorem 9. Observe that $\mu_t(\mathbb{X}) = t\ell_d(W)$.

Theorem 14 *Denote by m_t the median of N_t. Then there is a constant c_d such that*

$$\mathbb{P}(|N_t - m_t| \ge u) \le 4t\ell_d(W) \exp\left(-\frac{1}{16}\frac{u}{\sqrt{u+m_t}}\right)$$

for $\frac{u}{\sqrt{u+m_t}} \ge c_d$.

In [28] a concentration inequality for all $u \ge 0$ is given using a similar but more detailed approach.

4.4 Random Simplicial Complexes

Given the Gilbert graph of a Poisson point process η_t we construct the Vietoris–Rips complex $R(r_t)$ by calling $F = \{x_{i_1}, \dots, x_{i_{k+1}}\}$ a k-face of $R(r_t)$ if all pairs of points in F are connected by an edge in the Gilbert graph. Observe that, e.g., counting the number $N_t^{(k)}$ of k-faces is equivalent to a particular subgraph counting. By definition this is a local U-statistics given by

$$N_t^{(k)} = N_t^{(k)}(W, r_t) = \frac{1}{(k+1)!} \sum_{(x_1,\dots,x_{k+1})\in\eta_{t,\neq}^{k+1}} \mathbb{1}(\|x_i - x_j\| \le r_t,\ \forall 1 \le i,j \le k+1).$$

Central limit theorems and a concentration inequality follow immediately from the results for local U-statistics. We restrict our statements again to the thermodynamic case where tr_t^d tends to a constant as $t \to \infty$. Results for other regimes can be found, e.g., in Penrose's book. Because of the local definition of the Gilbert graph, $N_t^{(k)}$ is a local U-statistic. Theorem 6 with $\nu_t = tr_t^d$ can be applied.

Theorem 15 *Let N be a standard Gaussian random variable. Then constants $c(W)$ exist such that*

$$d_W(\tilde{N}_t^{(k)}, N) \le c(W)t^{-1/2},$$

$$d_K(\tilde{N}_t^{(k)}, N) \le c(W)t^{-1/2},$$

for $t \ge 1$.

A concentration inequality follows immediately from Theorem 9. Observe that $\mu_t(\mathbb{X}) = t\theta([W])$.

Theorem 16 *Denote by m_t the median of $N_t^{(k)}$. Then*

$$\mathbb{P}(|N_t^{(k)} - m_t| \ge u) \le 4t\ell_d(W) \exp\left(-\frac{1}{4(k+1)^2} \frac{u^{\frac{2}{k}}}{(u+m_t)^{\frac{1}{k}}}\right)$$

for $\frac{u^2}{u+m_t} \ge c_{d,k}$.

Much deeper results concerning the topology of random simplicial complexes are contained in [7, 15, 17]. We refer the interested reader to the recent survey article by Kahle [16]

4.5 Sylvester's Constant

Again we assume that the Poisson point process η has an intensity-measure of the form $\mu_t(\cdot) = t\ell_d(\cdot \cap W)$, where ℓ_d is Lebesgue measure and $W \subset \mathbb{R}^d$ a compact convex set with $\ell_d(W) = 1$.

As a last example of a U-statistic we consider the following functional related to Sylvester's problem. Originally raised with $k = 4$ in 1864, Sylvester's original problem asks for the distribution of the number of vertices of the convex hull of four random points. Put

$$N_t = N_t(W,k) = \sum_{(x_1,\dots,x_k)\in\eta^k_{\neq}} \mathbb{1}(x_1,\dots,x_k \text{ are vertices of } \mathrm{conv}(x_1,\dots,x_k)),$$

which counts the number of k-tuples of the process such that every point is a vertex of the convex hull, i.e., the number of k-tuples in convex position.

The expected value of U is then given by

$$\mathbb{E}N_t = t^k \mathbb{P}(X_1,\dots,X_k \text{ are vertices of } \mathrm{conv}(X_1,\dots,X_k)),$$

where $X_1,\dots,X_k$ are independent random points chosen according to the uniform distribution on W.

The question to determine the probability that k random points in a convex set W are in convex position has a long history, see, e.g., the more recent development by Bárány [2]. In our setting, the function $t^{-k}N_t$ is an estimator for this probability and we are interested in its distributional properties.

The asymptotic behavior of $\mathrm{Var}(N_t)$ is of order t^{2k-1}. Together with Theorem 5, we immediately get the following result showing that the estimator H is asymptotically Gaussian. Again, by $\tilde{N}_t$ we denote the normalized version of N_t.

Theorem 17 *Let N be a standard Gaussian random variable. Then there exists a constant $c(W,k)$ such that*

$$d_W\left(\tilde{N}_t, N\right) \le c(W,k)t^{-\frac{1}{2}}.$$

For much more information on random polytopes we refer the reader to [13], Sect. 3.5, in this survey.

References

1. Ané, C., Ledoux, M.: On logarithmic Sobolev inequalities for continuous time random walks on graphs. Probab. Theory Relat. Fields **116**, 573–602 (2000)
2. Bárány, I.: Sylvester's question: the probability that n points are in convex position. Ann. Probab. **27**, 2020–2034 (1999)
3. Bourguin, S., Peccati, G: Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering. Electron. J. Probab. **19**(66), 1–42 (2014)
4. Bourguin, S., Peccati, G.: The Malliavin-Stein method on the Poisson space. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 185–228. Springer, Cham (2016)
5. Breton, J., Houdré, C., Privault, N.: Dimension free and infinite variance tail estimates on Poisson space. Acta Appl. Math. **95**, 151–203 (2007)

6. De Jong, P.: A central limit theorem for generalized multilinear forms. J. Multivariate Anal. **34**, 275–289 (1990)
7. Decreusefond, L., Ferraz, E., Randriam, H., Vergne, A.: Simplicial homology of random configurations. Adv. Appl. Probab. **46**, 1–20 (2014)
8. Dynkin, E.B., Mandelbaum, A.: Symmetric statistics, Poisson point processes, and multiple Wiener integrals. Ann. Stat. **11**(3), 739–745 (1983)
9. Eichelsbacher, P., Thäle, C.: New Berry-Essen bounds for non-linear functionals of Poisson random measures. Electron. J. Probab. **19**(102), 1–25 (2014)
10. Fissler, T., Thäle, C.: A four moments theorem for gamma limits on a Poisson chaos. arXiv:1502.01568 (2015)
11. Heinrich, L., Schmidt, H., Schmidt, V.: Central limit theorems for Poisson hyperplane tessellations. Ann. Appl. Probab. **16**, 919–950 (2006)
12. Houdré, C., Privault, N.: Concentration and deviation inequalities in infinite dimensions via covariance representations. Bernoulli **8**, 697–720 (2002)
13. Hug, D., Reitzner, M.: Introduction to stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi and Springer Series, vol. **6**, pp. 145–184. Springer/Bocconi University Press, Milan (2016)
14. Hug, D., Thäle, C., Weil, W.: Intersection and proximity for processes of flats. J. Math. Anal. Appl. **426**, 1–42 (2015)
15. Kahle, M.: Random geometric complexes. Discrete Comp. Geom. **45**, 553–573 (2011)
16. Kahle, M.: Topology of random simplicial complexes: a survey. In: Tillmann, U., Galatius, S., Sinha, D. (eds.) Algebraic Topology: Applications and New Directions. Contemporary Mathematics, vol. 620, pp. 201–221. American Mathematical Society, Providence (2014)
17. Kahle, M., Meckes, E.: Limit theorems for Betti numbers of random simplicial complexes. Homol. Homotopy Appl. **15**, 343–374 (2013)
18. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs. Electron. J. Probab. **18**(32), 1–32 (2013)
19. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space, II: rescaled kernels, marked processes and geometric U-statistics. Stoch. Process. Appl. **123**, 4186–4218 (2013)
20. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Milan (2016)
21. Last, G., Penrose, M., Schulte, M., Thäle, C.: Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**(2), 348–364 (2014)
22. Peccati, G., Solé, J.L., Taqqu, M.S., Utzet, F.: Stein's method and normal approximation of Poisson functionals. Ann. Prob. **38**, 443–478 (2010)
23. Peccati, G., Thäle, C.: Gamma limits and U-statistics on the Poisson space. ALEA Lat. Am. J. Probab. Math. Stat. **10**, 525–560 (2013)
24. Peccati, G., Zheng, C.: Multi-dimensional Gaussian fluctuations on the Poisson space. Electron. J. Probab. **15**(48), 1487–1527 (2010)
25. Penrose, M.D.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
26. Reitzner, M.: Poisson point processes: Large deviation inequalities for the convex distance. Electron. Commun. Probab. **18**(96), 1–7 (2013)
27. Reitzner, M., Schulte, M.: Central limit theorems for U-statistics of Poisson point processes. Ann. Probab. **41**(6), 3879–3909 (2013)
28. Reitzner, M., Schulte, M., Thäle, C.: Limit theory for the Gilbert graph. Preprint (2014). arXiv 1312.4861. http://arxiv.org/abs/1312.4861
29. Reynaud-Bouret, P.: Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. Probab. Theory Relat. Fields **126**, 103–153 (2003)

30. Schulte, M.: A central limit theorem for the Poisson–Voronoi approximation. Adv. Appl. Math. **49**, 285–306 (2012)
31. Schulte, M.: Normal approximation of Poisson functionals in Kolmogorov distance. J. Theor. Prob. **29**(1), 96–117 (2016)
32. Schulte, M., Thäle, C.: The scaling limit of Poisson-driven order statistics with applications in geometric probability. Stoch. Process. Appl. **122**, 4096–4120 (2012)
33. Vitale, R.: Covariances of symmetric statistics. J. Multivariate Anal. **41**, 14–26 (1992)

Poisson Point Process Convergence and Extreme Values in Stochastic Geometry

Matthias Schulte and Christoph Thäle

Abstract Let η_t be a Poisson point process with intensity measure $t\mu$, $t > 0$, over a Borel space $\mathbb{X}$, where μ is a fixed measure. Another point process ξ_t on the real line is constructed by applying a symmetric function f to every k-tuple of distinct points of η_t. It is shown that ξ_t behaves after appropriate rescaling like a Poisson point process, as $t \to \infty$, under suitable conditions on η_t and f. This also implies Weibull limit theorems for related extreme values. The result is then applied to investigate problems arising in stochastic geometry, including small cells in Voronoi tessellations, random simplices generated by non-stationary hyperplane processes, triangular counts with angular constraints, and non-intersecting k-flats. Similar results are derived if the underlying Poisson point process is replaced by a binomial point process.

1 Introduction

This chapter deals with the application of the Malliavin–Chen–Stein method for Poisson approximation to problems arising in stochastic geometry. More precisely, we will develop a general framework which yields Poisson point process convergence and Weibull limit theorems for the order-statistic of a class of functionals driven by an underlying Poisson or binomial point process on an abstract state space.

To motivate our general theory, let us describe a particular situation to which our results can be applied (see Remark 4 and also Example 4 in [29] for more details). Let K be a convex body in $\mathbb{R}^d$, $d \geq 2$, (that is a compact convex set with interior points) whose volume is denoted by $\ell_d(K)$. For $t > 0$ let η_t be the restriction to K of a translation-invariant Poisson point process in $\mathbb{R}^d$ with intensity t and let $(\theta_t)_{t>0}$ be a sequence of real numbers satisfying $t^{2/d}\theta_t \to \infty$, as $t \to \infty$. Taking η_t as vertex set of a random graph, we connect two different points of η_t by an edge if and only if

M. Schulte (✉)
Karlsruhe Institute of Technology, Institute of Stochastics, 76128 Karlsruhe, Germany
e-mail: matthias.schulte@kit.edu

C. Thäle
Faculty of Mathematics, Ruhr University Bochum, 44780 Bochum, Germany
e-mail: christoph.thaele@rub.de

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*,
Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_8

their Euclidean distance does not exceed θ_t. The so-constructed random geometric graph, or Gilbert graph, is among the most prominent random graph models (see [25] for some recent developments and [22] for an exhaustive reference). We now consider the order statistic $\xi_t = \{M_t^{(m)} : m \in \mathbb{N}\}$ defined by the edge-lengths of the random geometric graph, that is, $M_t^{(1)}$ is the length of the shortest edge, $M_t^{(2)}$ is the length of the second-shortest edge etc. Now, our general theory implies that the rescaled point process $t^{2/d}\xi_t$ converges towards a Poisson point process on $\mathbb{R}_+$ with intensity measure given by $B \mapsto \beta d \int_B u^{d-1}\,\mathrm{d}u$ for Borel sets $B \subset \mathbb{R}_+$, where $\beta = \kappa_d \ell_d(K)/2$ and κ_d stands for the volume of the d-dimensional unit ball. Moreover, it implies that there is a constant $C > 0$ only depending on K such that

$$\left| \mathbb{P}\left(t^{2/d} M_t^{(m)} > y\right) - e^{-\beta y^d} \sum_{i=0}^{m-1} \frac{(\beta y^d)^i}{i!} \right| \leq C \max\{y^{d+1}, y^{2d}\}\, t^{-2/d}$$

for any $m \in \mathbb{N}$, $y \in (0, t^{2/d}\theta_t)$ and $t \geq 1$. In particular, the distribution of the rescaled length $t^{2/d}M_t^{(1)}$ of the shortest edge of the random graph converges, as $t \to \infty$, to a Weibull distribution with survival function $y \mapsto e^{-\beta y^d}$, $y \geq 0$, at rate $t^{-2/d}$.

Our purpose here is to establish a general framework that can be applied to a broad class of examples. We also allow the underlying point process to be a Poisson *or* a binomial point process. Our main result for the Poisson case refines those in [29] or [30] and improves the rate of convergence. Its proof follows the ideas of Peccati [21] and Schulte and Thäle [29], but uses the special structure of the functional under consideration as well as recent techniques from [20] around Mehler's formula on the Poisson space. This saves some technical computations related to the product formula for multiple stochastic integrals (cf. [18], in this volume, as well as [19, 32]). In case of an underlying binomial point process we use a bound for the Poisson approximation of (classical) U-statistics from [1]. As application of our main results, we present a couple of examples, which continue and complement those studied in [29, 30]. These are

1. Cells with small (nucleus-centered) inradius in a Voronoi tessellation.
2. Simplices generated by a class of rotation-invariant hyperplane processes.
3. Almost collinearities and flat triangles in a planar Poisson or binomial process.
4. Arbitrary length-power-proximity functionals of non-intersecting k-flats.

The rest of this chapter is organized as follows. Our main results and their framework are presented in Sect. 2. The application to problems arising in stochastic geometry is the content of Sect. 3. The proofs of the main results are postponed to the final Sect. 4.

2 Results

Let η_t ($t > 0$) be a Poisson point process on a measurable space $(\mathbb{X}, \mathcal{X})$ with intensity measure $\mu_t := t\mu$, where μ is a fixed σ-finite measure on $\mathbb{X}$. To avoid technical complications, we shall assume in this chapter that $(\mathbb{X}, \mathcal{X})$ is a standard Borel space. This ensures, for example, that any point process on $\mathbb{X}$ can almost surely be represented as a sum of Dirac measures. Let further $k \in \mathbb{N}$ and $f : \mathbb{X}^k \to \mathbb{R}$ be a measurable symmetric function. Our aim here is to investigate the point process ξ_t on $\mathbb{R}$ which is induced by η_t and f as follows:

$$\xi_t := \frac{1}{k!} \sum_{(x_1,\ldots,x_k)\in\eta^k_{t,\neq}} \delta_{f(x_1,\ldots,x_k)} . \tag{1}$$

Here $\eta^k_{t,\neq}$ stands for the set of all k-tuples of distinct points of η_t and δ_x is the unit Dirac measure concentrated at the point $x \in \mathbb{R}$. We shall assume that

$$\mu_t^k(f^{-1}([-s, s])) < \infty \qquad \text{for all} \qquad s > 0\,,$$

to ensure that ξ_t is a locally finite counting measure on $\mathbb{R}$.

For $m \in \mathbb{N}$ we denote by $M_t^{(m)}$ the distance from the origin to the m-th point of ξ_t on the positive half-line $\mathbb{R}_+ := (0, \infty)$, and by $M_t^{(-m)}$ the distance from the origin to the m-th point on the negative half-line $\mathbb{R}_- := (-\infty, 0]$. If ξ_t has less than m points on the positive or negative half-line, we put $M_t^{(m)} = \infty$ or $M_t^{(-m)} = \infty$, respectively.

Fix $\gamma \in \mathbb{R}$ and for $y_1, y_2 \in \mathbb{R}$ define

$$\alpha_t(y_1, y_2) := \frac{1}{k!} \int\limits_{\mathbb{X}^k} \mathbb{1}\{t^{-\gamma} y_1 < f(x_1, \ldots, x_k) \le t^{-\gamma} y_2\}\, \mu_t^k(\mathrm{d}(x_1, \ldots, x_k))\,.$$

We remark that, as a consequence of the multivariate Mecke formula for Poisson point processes (see [18, formula (1.11)]), $\alpha_t(y_1, y_2)$ can be interpreted as

$$\alpha_t(y_1, y_2) = \frac{1}{k!}\mathbb{E} \sum_{(x_1,\ldots,x_k)\in\eta^k_{t,\neq}} \mathbb{1}\{t^{-\gamma} y_1 < f(x_1, \ldots, x_k) \le t^{-\gamma} y_2\}\,,$$

which is the expected number of points of ξ_t in $(t^{-\gamma} y_1, t^{-\gamma} y_2]$ if $y_1 < y_2$ and zero if $y_1 \ge y_2$. Moreover, let, for $k \ge 2$,

$$r_t(y) := \max_{1\le\ell\le k-1} \int\limits_{\mathbb{X}^\ell} \Bigg(\int\limits_{\mathbb{X}^{k-\ell}} \mathbb{1}\{|f(x_1, \ldots, x_k)| \le t^{-\gamma} y\}\, \mu_t^{k-\ell}(\mathrm{d}(x_{\ell+1}, \ldots, x_k)) \Bigg)^2 \mu_t^\ell(\mathrm{d}(x_1, \ldots, x_\ell))$$

for $y \ge 0$ and put $r_t \equiv 0$ if $k = 1$.

Theorem 1 *Let ν be a σ-finite non-atomic Borel measure on $\mathbb{R}$. Then, there is a constant $C \geq 1$ only depending on k such that*

$$\left|\mathbb{P}(t^\gamma M_t^{(m)} > y) - e^{-\nu((0,y])} \sum_{i=0}^{m-1} \frac{\nu((0,y])^i}{i!}\right| \leq |\nu((0,y]) - \alpha_t(0,y)| + C\, r_t(y)$$

and

$$\left|\mathbb{P}(t^\gamma M_t^{(-m)} \geq y) - e^{-\nu((-y,0])} \sum_{i=0}^{m-1} \frac{\nu((-y,0])^i}{i!}\right| \leq |\nu((-y,0]) - \alpha_t(-y,0)| + C\, r_t(y)$$

for all $m \in \mathbb{N}$ and $y \geq 0$. Moreover, if

$$\lim_{t\to\infty} \alpha_t(y_1,y_2) = \nu((y_1,y_2]) \qquad \text{for all} \qquad y_1,y_2 \in \mathbb{R} \text{ with } y_1 < y_2 \tag{2}$$

and

$$\lim_{t\to\infty} r_t(y) = 0 \qquad \text{for all} \qquad y > 0\,, \tag{3}$$

the rescaled point processes $(t^\gamma \xi_t)_{t>0}$ converge in distribution to a Poisson point process on $\mathbb{R}$ with intensity measure ν.

Remark 1 Let us comment on the particular case $k = 1$. Here, the point process ξ_t is itself a Poisson point process on $\mathbb{R}$ with intensity measure derived from α_t as a consequence of the famous mapping theorem, for which we refer to Sect. 2.3 in [16]. This is confirmed by our Theorem 1.

Remark 2 Theorem 1 generalizes earlier versions in [29, 30], which have a similar structure, but where the quantity

$$\hat{r}_t(y) := \sup_{\substack{(\hat{x}_1,\dots,\hat{x}_\ell)\in\mathbb{X}^\ell \\ 1\leq\ell\leq k-1}} \mu_t^{k-\ell}\big(\{(x_1,\dots,x_{k-\ell}) \in \mathbb{X}^{k-\ell} : |f(\hat{x}_1,\dots,\hat{x}_\ell,x_1,\dots,x_{k-\ell})| \leq t^{-\gamma}y\}\big)$$

for $y \geq 0$ is considered instead of $r_t(y)$. It is easy to see that $r_t(y)$ and $\hat{r}_t(y)$ are related by

$$r_t(y) \leq \inf_{\varepsilon>0} \alpha_t(-y-\varepsilon, y)\, \hat{r}_t(y) \quad \text{for all} \quad y \geq 0\,.$$

In particular, this means that the rate of convergence of the order statistics in Theorem 1 improves that in [29, 30] by removing a superfluous square root from $\hat{r}_t(y)$. Moreover and in contrast to [29, 30], the constant C only depends on the parameter k.

In our applications presented in Sect. 3, the function f is always strictly positive so that ξ_t is concentrated on $\mathbb{R}_+$. Moreover, the measure ν will be of a special form. The following corollary deals with this situation. To state it, we use the convention that $\alpha_t(y) := \alpha_t(0, y)$ for $y \geq 0$.

Corollary 1 *Let $\beta, \tau > 0$. Then there is a constant $C > 0$ only depending on k such that*

$$\left| \mathbb{P}(t^\gamma M_t^{(m)} > y) - e^{-\beta y^\tau} \sum_{i=0}^{m-1} \frac{(\beta y^\tau)^i}{i!} \right| \leq |\beta y^\tau - \alpha_t(y)| + C\, r_t(y)$$

for all $m \in \mathbb{N}$ and $y \geq 0$. If, additionally,

$$\lim_{t\to\infty} \alpha_t(y) = \beta y^\tau \quad \text{and} \quad \lim_{t\to\infty} r_t(y) = 0 \quad \text{for all} \quad y > 0\,, \tag{4}$$

the rescaled point processes $(t^\gamma \xi_t)_{t>0}$ converge in distribution to a Poisson point process on $\mathbb{R}_+$ with the intensity measure

$$\nu(B) = \beta\tau \int_B u^{\tau-1}\, \mathrm{d}u, \quad B \subset \mathbb{R}_+ \text{ Borel}. \tag{5}$$

Remark 3 The limiting Poisson point process appearing in the context of Corollary 1 is usually called a Weibull process on $\mathbb{R}_+$, the reason for this being that the distance from the origin to the next point follows a Weibull distribution.

If μ is a finite measure, i.e., if $\mu(\mathbb{X}) < \infty$, one can replace the underlying Poisson point process η_t by a binomial point process ζ_n having a fixed number of n points which are independent and identically distributed according to the probability measure $\mu(\,\cdot\,)/\mu(\mathbb{X})$. Without loss of generality we assume that $\mu(\mathbb{X}) = 1$ in what follows. In this situation, we consider instead of ξ_t defined at (1) the derived point process $\hat{\xi}_n$ on $\mathbb{R}$ given by

$$\hat{\xi}_n := \frac{1}{k!} \sum_{(x_1,\ldots,x_k)\in\zeta^k_{n,\neq}} \delta_{f(x_1,\ldots,x_k)}\,,$$

where $\zeta^k_{n,\neq}$ stands for the collection of all k-tuples of distinct points of ζ_n. For $m \in \mathbb{N}$ let $\widehat{M}_n^{(m)}$ and $\widehat{M}_n^{(-m)}$ be defined similarly as $M_n^{(m)}$ and $M_n^{(-m)}$ above with ξ_t replaced by $\hat{\xi}_n$. For $n, k \in \mathbb{N}$ we denote by $(n)_k$ the descending factorial $n\cdot(n-1)\cdot\ldots\cdot(n-k+1)$. Using the notation

$$\alpha_n(y_1, y_2) := \frac{(n)_k}{k!} \int_{\mathbb{X}^k} \mathbb{1}\,\{n^{-\gamma} y_1 < f(x_1,\ldots,x_k) \leq n^{-\gamma} y_2\}\ \mu^k(\mathrm{d}(x_1,\ldots,x_k))\,,$$

$$r_n(y) := \max_{1\le \ell\le k-1} (n)_{2k-\ell} \int\limits_{\mathbb{X}^\ell} \Bigg(\int\limits_{\mathbb{X}^{k-\ell}} \mathbb{1}\{|f(x_1,\dots,x_k)| \le n^{-\gamma} y\}$$

$$\mu^{k-\ell}(\mathrm{d}(x_{\ell+1},\dots,x_k))\Bigg)^2 \mu^\ell(\mathrm{d}(x_1,\dots,x_\ell))$$

for $y_1, y_2, y \in \mathbb{R}$, we can now present the binomial counterpart of Theorem 1.

Theorem 2 *Let μ be a probability measure on $\mathbb{X}$ and ν be a σ-finite non-atomic Borel measure on $\mathbb{R}$. Then, there is a constant $C \ge 1$ only depending on k such that*

$$\left|\mathbb{P}(n^\gamma \widehat{M}_n^{(m)} > y) - e^{-\nu((0,y])} \sum_{i=0}^{m-1} \frac{\nu((0,y])^i}{i!}\right|$$

$$\le |\nu((0,y]) - \alpha_n(0,y)| + C\Big(r_n(y) + \frac{\alpha_n(0,y)}{n}\Big)$$

and

$$\left|\mathbb{P}(n^\gamma \widehat{M}_n^{(-m)} \ge y) - e^{-\nu((-y,0])} \sum_{i=0}^{m-1} \frac{\nu((-y,0])^i}{i!}\right|$$

$$\le |\nu((-y,0]) - \alpha_n(-y,0)| + C\Big(r_n(y) + \frac{\alpha_n(-y,0)}{n}\Big)$$

for all $m \in \mathbb{N}$ and $y \ge 0$. Moreover, if

$$\lim_{n\to\infty} \alpha_n(y_1,y_2) = \nu((y_1,y_2]) \qquad \text{for all} \qquad y_1,y_2 \in \mathbb{R} \text{ with } y_1 < y_2$$

and

$$\lim_{n\to\infty} r_n(y) = 0 \qquad \text{for all} \qquad y > 0\,,$$

the rescaled point processes $(n^\gamma \hat{\xi}_n)_{n\ge 1}$ converge in distribution to a Poisson point process on $\mathbb{R}$ with intensity measure ν.

As in the Poisson case, Theorem 2 allows a reformulation as in Corollary 1 for the special situation in which f is nonnegative and ν has a power-law density. As above, we use the convention that $\alpha_n(y) := \alpha_n(0,y)$ for $y \ge 0$.

Corollary 2 *Let $\beta, \tau > 0$. Then there is a constant $C > 0$ only depending on k such that*

$$\left|\mathbb{P}(n^\gamma \widehat{M}_n^{(m)} > y) - e^{-\beta y^\tau} \sum_{i=0}^{m-1} \frac{(\beta y^\tau)^i}{i!}\right| \le |\beta y^\tau - \alpha_n(y)| + C\Big(r_n(y) + \frac{\alpha_n(y)}{n}\Big)$$

for all $m \in \mathbb{N}$ *and* $y \geq 0$. *If, additionally,*

$$\lim_{n\to\infty} \alpha_n(y) = \beta y^\tau \quad \text{and} \quad \lim_{n\to\infty} r_n(y) = 0 \qquad \text{for all} \qquad y > 0\,,$$

the rescaled point processes $(n^\gamma \hat{\xi}_n)_{n\geq 1}$ *converge in distribution to a Poisson point process on* $\mathbb{R}_+$ *with intensity measure given by* (5).

3 Examples

In this section we apply the results presented above to problems arising in stochastic geometry, see [11]. The minimal nucleus-centered inradius of the cells of a Voronoi tessellation is considered in Sect. 3.1. This example is inspired by the work [5] and was not previously considered in [29], although it is closely related to the minimal edge length of the random geometric graph discussed in the introduction. Our next example generalizes Example 6 of [29] from the translation-invariant case to arbitrary distance parameters $r \geq 1$. In dimension two it also sheds some new light onto the area of small cells in line tessellations. Our third example is inspired by a result in [31] and deals with approximate collinearities and flat triangles induced by a planar Poisson or binomial point process. Our last example deals with non-intersecting k-flats. The result generalizes Example 1 in [29] and one of the results in [30] to arbitrary distance powers $a > 0$.

3.1 Voronoi Tessellations

For a finite set $\chi \neq \emptyset$ of points in $\mathbb{R}^d$, $d \geq 2$, the Voronoi cell $v_\chi(x)$ with nucleus $x \in \chi$ is the (possibly unbounded) set

$$v_\chi(x) = \big\{z \in \mathbb{R}^d : \|x - z\| \leq \|x' - z\| \text{ for all } x' \in \chi \setminus \{x\}\big\}$$

of all points in $\mathbb{R}^d$ having x as their nearest neighbor in χ. The family

$$\mathcal{V}_\chi = \{v_\chi(x) : x \in \chi\}$$

subdivides $\mathbb{R}^d$ into a finite number of random polyhedra, which form the so-called Voronoi tessellation associated with χ, see [27, Chap. 10.2]. For $\chi = \emptyset$ we put $\mathcal{V}_\emptyset = \{\mathbb{R}^d\}$. One characteristic measuring the size of a Voronoi cell $v_\chi(x)$ is its nucleus-centered inradius $R(x, \chi)$. It is defined as the radius of the largest ball included in $v_\chi(x)$ and having x as its midpoint. Note that $R(x, \chi)$ takes the value ∞ if $\chi = \{x\}$.

Define

$$R(\mathcal{V}_\chi) := \min\{R(x,\chi) : x \in \chi\}$$

for nonempty χ and $R(\mathcal{V}_\emptyset) := \infty$.

In [5] the asymptotic behavior of $R(\mathcal{V}_\chi)$ has been investigated in the case that χ is a Poisson point process in a convex body K of intensity $t > 0$, as $t \to \infty$. Using Corollary 1 we can get back one of the main results of [5] and add a rate of convergence to the limit theorem (compare with [5, Eq. (2b)] in particular). Moreover, we provide a similar result for an underlying binomial point process.

Corollary 3 *Let η_t be a Poisson point process with intensity measure $t\ell_d|_K$, where $\ell_d|_K$ stands for the restriction of the Lebesgue measure to a convex body K and $t > 0$. Then, there exists a constant $C > 0$ depending on K such that*

$$\left|\mathbb{P}\big(t^{2/d}R(\mathcal{V}_{\eta_t}) > y\big) - e^{-2^{d-1}\kappa_d\ell_d(K)y^d}\right| \le C\,t^{-2/d}\max\{y^{d+1}, y^{2d}\}$$

for all $y \ge 0$ and $t \ge 1$. In addition, if ζ_n is a binomial point process with $n \ge 2$ independent points distributed according to $\ell_d(K)^{-1}\,\ell_d|_K$, then

$$\left|\mathbb{P}\big(n^{2/d}R(\mathcal{V}_{\zeta_n}) > y\big) - e^{-2^{d-1}\kappa_d\ell_d(K)y^d}\right| \le C\,n^{-2/d}\max\{y^{d}, y^{2d}\}$$

for $y \ge 0$ and with a constant $C > 0$ depending on K.

Proof To apply Corollary 1 we first have to investigate $\alpha_t(y)$ for fixed $y > 0$. For this we abbreviate $\mathcal{V}_{\eta_t}$ by $\mathcal{V}_t$ and observe that—by definition of a Voronoi cell—$R(\mathcal{V}_t)$ is half of the minimal interpoint distance of points from η_t, i.e.

$$R(\mathcal{V}_t) = \frac{1}{2}\min\left\{\|x_1 - x_2\| : (x_1, x_2) \in \eta^2_{t,\neq}\right\}.$$

Consequently, we have

$$\begin{aligned}
\alpha_t(y) &= \frac{t^2}{2}\int\limits_K\int\limits_K \mathbb{1}\{\|x_1 - x_2\| \le 2yt^{-\gamma}\}\,\mathrm{d}x_2\,\mathrm{d}x_1\\
&= \frac{t^2}{2}\int\limits_{\mathbb{R}^d}\ell_d(K\cap B^d_{2yt^{-\gamma}}(x_1))\,\mathrm{d}x_1 - \frac{t^2}{2}\int\limits_{\mathbb{R}^d\setminus K}\ell_d(K\cap B^d_{2yt^{-\gamma}}(x_1))\,\mathrm{d}x_1\,,
\end{aligned}$$

where $B^d_r(x)$ is the d-dimensional ball of radius $r > 0$ around $x \in \mathbb{R}^d$. From Theorem 5.2.1 in [27] (see Eq. (14) in particular) it follows that

$$\frac{t^2}{2}\int\limits_{\mathbb{R}^d} V_d(K\cap B^d_{2yt^{-\gamma}}(x_1))\,\mathrm{d}x_1 = \frac{t^2}{2}\,\ell_d(K)\,\kappa_d(2yt^{-\gamma})^d = 2^{d-1}\ell_d(K)\kappa_d y^d t^{2-\gamma d}\,.$$

Moreover, Steiner's formula [27, Eq. (14.5)] yields

$$\begin{aligned}
&\frac{t^2}{2} \int_{\mathbb{R}^d \setminus K} \ell_d(K \cap B^d_{2yt^{-\gamma}}(x_1))\, \mathrm{d}x_1 \\
&\quad \le \frac{\kappa_d}{2}\, t^2 (2yt^{-\gamma})^d\, \ell_d\left(\left\{z \in \mathbb{R}^d \setminus K : \inf_{z' \in K} \|z - z'\| \le 2yt^{-\gamma}\right\}\right) \\
&\quad = \frac{\kappa_d}{2} t^2 (2yt^{-\gamma})^d \sum_{j=0}^{d-1} \kappa_{d-j} V_j(K) (2yt^{-\gamma})^{d-j}\,,
\end{aligned}$$

where $V_0(K), \dots, V_{d-1}(K)$ are the so-called intrinsic volumes of K, see [11] or [27]. Choosing $\gamma = 2/d$, this implies that $\alpha_t(y)$ is dominated by its first integral term and that

$$\left|\alpha_t(y) - 2^{d-1}\kappa_d \ell_d(K) y^d\right| \le c_1\, t^{-2/d} \max\{y^{d+1}, y^{2d}\}$$

for $t \ge 1$ with a constant c_1 only depending on K.

Finally, we have to deal with $r_t(y)$. Here, we have

$$\begin{aligned}
r_t(y) &= t^3 \int_K \left(\int_K \mathbb{1}\{\|x - y\| \le 2yt^{-\gamma}\}\, \mathrm{d}y \right)^2 \mathrm{d}x \\
&\le t^3 \ell_d(K)\, (t^{-2}\kappa_d 2^d y^d)^2 = \ell_d(K)\, 4^d \kappa_d^2 y^{2d} t^{-1}\,.
\end{aligned}$$

In the binomial case, one can derive analogous bounds for $\alpha_n(y)$ and $r_n(y)$, $y > 0$. Since $\min(2/d, 1) = 2/d$ for all $d \ge 2$, application of Corollaries 1 and 2 completes the proof. □

Remark 4 We have used in the proof that $R(\mathcal{V}_{\eta_t})$ is half of the minimal inter-point distance between points of η_t in K. Thus, Corollary 3 also makes a statement about this minimal inter-point distance. Consequently, $2R(\mathcal{V}_{\eta_t})$ is also the same as the shortest edge length of a random geometric graph based on η_t as discussed in the introduction (cf. [25] and [22] for an exhaustive reference on random geometric graphs) or as the shortest edge length of a Delaunay graph (see [11] or [6, 27] for background material on Delaunay graphs or tessellations). A similar comment applies if η_t is replaced by a binomial point process ζ_n.

3.2 Hyperplane Tessellations

Let $\mathcal{H}$ be the space of hyperplanes in $\mathbb{R}^d$, fix a distance parameter $r \ge 1$ and a convex body $K \subset \mathbb{R}^d$, and define as in [12, Sect. 3.4.5] a (finite) measure μ on $\mathcal{H}$

by the relation

$$\int_{\mathcal{H}} g(H)\,\mu(\mathrm{d}H) = \int_{\mathbb{S}^{d-1}} \int_0^\infty g(u^\perp + pu)\,\mathbb{1}\{(u^\perp + pu)\cap K \neq \emptyset\}\,p^{r-1}\,\mathrm{d}p\,\mathrm{d}u\,,$$

where $g \geq 0$ is a measurable function on $\mathcal{H}$, $u^\perp$ is the linear subspace of all vectors that are orthogonal to u, and du stands for the infinitesimal element of the normalized Lebesgue measure on the $(d-1)$-dimensional unit sphere $\mathbb{S}^{d-1}$. By η_t we mean in this section a Poisson point process on $\mathcal{H}$ with intensity measure $\mu_t := t\mu$, $t > 0$. Let us further write for $n \in \mathbb{N}$ with $n \geq d+1$, ζ_n for a binomial process on $\mathcal{H}$ consisting of $n \in \mathbb{N}$ hyperplanes distributed according to the probability measure $\mu(\mathcal{H})^{-1}\,\mu$.

If $K = \mathbb{R}^d$ in the Poisson case, one obtains a tessellation of the whole $\mathbb{R}^d$ into bounded cells. In this context one is interested in the so-called zero cell Z_0, which is the almost surely uniquely determined cell containing the origin. If $r = 1$, Z_0 has the same distribution as the zero-cell of a rotation- and translation-invariant Poisson hyperplane tessellation. If $r = d$, Z_0 is equal in distribution to the so-called typical cell of a Poisson–Voronoi tessellation as considered in the previous section, see [27]. Thus, the tessellation induced by η_t interpolates in some sense between the translation-invariant Poisson hyperplane and the Poisson–Voronoi tessellation, which explains the recent interest in this model [8, 9, 12]. For more background material about random tessellations (and in particular Poisson hyperplane and Poisson–Voronoi tessellations) we refer to Chap. 10 in [27] and Chap. 9 in [6] and also to [11].

We are interested here in the simplices generated by the hyperplanes of η_t or ζ_n, which are contained in the prescribed convex set K. For a $(d+1)$-tuple $(H_1,\ldots,H_{d+1})$ of distinct hyperplanes of η_t or ζ_n let us write $[H_1,\ldots,H_{d+1}]$ for the simplex generated by $H_1,\ldots,H_{d+1}$ and define the point processes

$$\xi_t := \frac{1}{(d+1)!} \sum_{(H_1,\ldots,H_{d+1})\in\eta^{d+1}_{t,\neq}} \delta_{\ell_d([H_1,\ldots,H_{d+1}])}\,\mathbb{1}\big\{[H_1,\ldots,H_{d+1}]\subset K\big\}$$

and

$$\hat{\xi}_n := \frac{1}{(d+1)!} \sum_{(H_1,\ldots,H_{d+1})\in\zeta^{d+1}_{n,\neq}} \delta_{\ell_d([H_1,\ldots,H_{d+1}])}\,\mathbb{1}\big\{[H_1,\ldots,H_{d+1}]\subset K\big\}\,.$$

By $M_t^{(m)}$ and $\widehat{M}_n^{(m)}$ we mean the mth order statistics associated with ξ_t and $\hat{\xi}_n$, respectively. In particular $M_t^{(1)}$ and $\widehat{M}_n^{(1)}$ are the smallest volume of a simplex included in K. Moreover, for fixed hyperplanes $H_1,\ldots,H_d$ in general position let $z(H_1,\ldots,H_d) := H_1\cap\ldots\cap H_d$ be the intersection point of $H_1,\ldots,H_d$. By $H_{\delta,u}$ we denote the hyperplane with unit normal vector $u \in \mathbb{S}^{d-1}$ and distance $\delta > 0$ to the origin. The following result generalizes [29, Theorem 2.6] from the translation-invariant case $r = 1$ to arbitrary distance parameter $r \geq 1$.

Corollary 4 *Define*

$$\beta := \frac{1}{(d+1)!} \int_{\mathcal{H}^d} \int_{\mathbb{S}^{d-1}} \mathbb{1}\{H_1 \cap \ldots \cap H_d \cap K \neq \emptyset\} \, |u^T z(H_1, \ldots, H_d)|^{r-1} \times \ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}])^{-1/d} \, \mathrm{d}u \, \mu^d(\mathrm{d}(H_1, \ldots, H_d)) \, .$$

Then $t^{d(d+1)}\xi_t$ and $n^{d(d+1)}\hat{\xi}_n$ converge, as $t \to \infty$ or $n \to \infty$, in distribution to a Poisson point process on $\mathbb{R}_+$ with intensity measure given by

$$B \mapsto \frac{\beta}{d} \int_B u^{(1-d)/d} \, \mathrm{d}u$$

for Borel sets $B \subset \mathbb{R}_+$. In particular, for each $m \in \mathbb{N}$, $t^{d(d+1)}M_t^{(m)}$ and $n^{d(d+1)}\widehat{M}_n^{(m)}$ converge towards a random variable with survival function

$$y \mapsto \exp\big(-\beta\, y^{1/d}\big) \sum_{i=0}^{m-1} \frac{(\beta y^{1/d})^i}{i!}, \quad y \geq 0 \, .$$

Proof For $y > 0$ we have

$$\alpha_t(y) = \frac{t^{d+1}}{(d+1)!} \int_{\mathcal{H}^{d+1}} \mathbb{1}\{[H_1, \ldots, H_{d+1}] \subset K\} \times \mathbb{1}\{\ell_d([H_1, \ldots, H_{d+1}]) \leq y t^{-\gamma}\} \, \mu^{d+1}(\mathrm{d}(H_1, \ldots, H_{d+1})) \, .$$

For fixed hyperplanes $H_1, \ldots, H_d$ in general position we parametrize H_{d+1} by a pair $(\delta, u) \in [0, \infty) \times \mathbb{S}^{d-1}$, where δ is the distance of H_{d+1} to the origin. Then $\alpha_t(y)$ can be rewritten as

$$\alpha_t(y) = \frac{1}{2(d+1)!} \int_{\mathcal{H}^d} \int_{\mathbb{S}^{d-1}} \int_{-\infty}^{\infty} t^{d+1} \mathbb{1}\{[H_1, \ldots, H_d, H_{\delta,u}] \subset K\} \times \mathbb{1}\{\ell_d([H_1, \ldots, H_d, H_{\delta,u}]) \leq y t^{-\gamma}\} |\delta|^{r-1} \, \mathrm{d}\delta \, \mathrm{d}u \, \mu^d(\mathrm{d}(H_1, \ldots, H_d)) \, . \tag{6}$$

Since the hyperplane $H_{\delta,u}$ has the distance $|u^T z(H_1, \ldots, H_d) - \delta|$ to $z(H_1, \ldots, H_d)$, we have that

$$\ell_d([H_1, \ldots, H_d, H_{\delta,u}]) = |u^T z(H_1, \ldots, H_d) - \delta|^d \, \ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}]) \, .$$

Let $\gamma = d(d+1)$ and $M := \max\{\|z\|^{r-1} : z \in K\}$. For fixed $H_1, \ldots, H_d \in \mathcal{H}^d$ such that $H_1 \cap \ldots \cap H_d \cap K \neq \emptyset$ and $u \in \mathbb{S}^{d-1}$ we can estimate the inner integral in (6) from above by

$$M \int_{-\infty}^{\infty} t^{d+1} \mathbb{1}\{|u^T z(H_1, \ldots, H_d) - \delta|^d$$
$$\ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}]) \leq y t^{-\gamma}\} \, \mathrm{d}\delta$$
$$\leq 2M \, \ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}])^{-1/d} \, y^{1/d}.$$

The hyperplanes $H_1 - z(H_1, \ldots, H_d), \ldots, H_d - z(H_1, \ldots, H_d)$ partition the unit sphere $\mathbb{S}^{d-1}$ into 2^d spherical caps $S_1, \ldots, S_{2^d}$. For each $u \in S_j$ $(1 \leq j \leq 2^d)$, transformation into spherical coordinates shows that

$$\ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}]) \geq c_d \, \ell_{d-1}(S_j) \, ,$$

where $c_d > 0$ is a dimension dependent constant and $\ell_{d-1}(S_j)$ is the spherical Lebesgue measure of S_j. Consequently, we have

$$\alpha_t(y) \leq \frac{M}{(d+1)!} \int_{\mathcal{H}^d} \mathbb{1}\{H_1 \cap \ldots \cap H_d \cap K \neq \emptyset\}$$
$$\times \sum_{j=1}^{2^d} \int_{S_j} \left(\frac{y}{c_d \, \ell_{d-1}(S_j)} \right)^{1/d} \mathrm{d}u \, \mu^d(\mathrm{d}(H_1, \ldots, H_d))$$
$$\leq \frac{M}{(d+1)!} \int_{\mathcal{H}^d} \mathbb{1}\{H_1 \cap \ldots \cap H_d \cap K \neq \emptyset\}$$
$$\times \sum_{j=1}^{2^d} \ell_{d-1}(S_j) \left(\frac{y}{c_d \, \ell_{d-1}(S_j)} \right)^{1/d} \mu^d(\mathrm{d}(H_1, \ldots, H_d)) \, .$$

Since the last expression is finite, we can apply the dominated convergence theorem in (6). By the same arguments we used to obtain an upper bound for the inner integral in (6), we see that, for $H_1, \ldots, H_d \in \mathcal{H}^d$ and $u \in \mathbb{S}^{d-1}$,

$$\lim_{t \to \infty} \int_{-\infty}^{\infty} t^{d+1} \mathbb{1}\,\{[H_1, \ldots, H_d, H_{\delta,u}] \subset K\} \, \mathbb{1}\,\{\ell_d([H_1, \ldots, H_d, H_{\delta,u}]) \leq y t^{-\gamma}\} \, |\delta|^{r-1} \, \mathrm{d}\delta$$
$$= 2 \mathbb{1}\,\{H_1 \cap \ldots \cap H_d \cap K \neq \emptyset\} \; \ell_d([H_1, \ldots, H_d, z(H_1, \ldots, H_d) + H_{1,u}])^{-1/d}$$
$$\times |u^T z(H_1, \ldots, H_d)|^{r-1} y^{1/d} \, .$$

Altogether, we obtain that

$$\lim_{t\to\infty} \alpha_t(y) = \beta y^{1/d} .$$

By the same estimates as above, we have that, for any $\ell \in \{1, \dots, d\}$,

$$t^\ell \int_{\mathcal{H}^\ell} \Big(t^{d+1-\ell} \int_{\mathcal{H}^{d+1-\ell}} \mathbb{1}\{[H_1, \dots, H_{d+1}] \subset H, V_d([H_1, \dots, H_{d+1}]) \le yt^{-\gamma}\}$$

$$\mu^{d+1-\ell}\big(\mathrm{d}(H_{\ell+1}, \dots, H_{d+1})\big)\Big)^2 \mu^\ell\big(\mathrm{d}(H_1, \dots, H_\ell)\big)$$

$$\le t^\ell \int_{\mathcal{H}^\ell} \Big(M t^{-\ell} \int_{\mathcal{H}^{d-\ell}} \int_{\mathbb{S}^{d-1}} \mathbb{1}\{H_1 \cap \ldots \cap H_d \cap K \ne \emptyset\}\, y^{1/d}$$

$$\ell_d\big([H_1, \dots, H_d, z(H_1, \dots, H_d) + H_{1,u}]\big)^{-1/d} \,\mathrm{d}u\, \mu^{d-\ell}\big(\mathrm{d}(H_{\ell+1}, \dots, H_d)\big)\Big)^2$$

$$\mu^\ell\big(\mathrm{d}(H_1, \dots, H_\ell)\big) .$$

Hence, $r_t(y) \to 0$ as $t \to \infty$ so that application of Corollary 1 completes the proof of the Poisson case. The result for an underlying binomial point process follows from similar estimates and Corollary 2. □

Remark 5 Although Corollary 1 or Corollary 2 deliver a rate of convergence, we cannot provide such rate for this particular example. This is due to the fact that the exact asymptotic behavior of $\alpha_t(y)$ or $\alpha_n(y)$ depends in a delicate way on the smoothness behavior of the boundary of K.

Corollary 4 admits a nice interpretation in the planar case $d = 2$. Namely, the smallest triangle contained in K coincides with the smallest triangular cell included in K of the line tessellation induced by η_t or ζ_n (note that this argument fails in higher dimensions). This way, Corollary 4 also makes a statement about the area of small triangular cells, which generalizes Corollary 2.7 in [29] from the translation-invariant case $r = 1$ to arbitrary distance parameters $r \ge 1$:

Corollary 5 *Denote by A_t or A_n the area of the smallest triangular cell in K of a line tessellation generated by a Poisson line process η_t or a binomial line process ζ_n with distance parameter $r \ge 1$, respectively. Then $t^6 A_t$ and $n^6 A_n$ both converge in distribution, as $t \to \infty$ or $n \to \infty$, to a Weibull random variable with survival function $y \mapsto \exp(-\beta\, y^{1/2})$, $y \ge 0$, where β is as in Corollary 4.*

3.3 Flat Triangles

So-called *ley lines* are expected alignments of a set of locations that are of geographical and/or historical interest, such as ancient monuments, megaliths and natural ridge-tops [4]. For this reason, there is some interest in archaeology, for example, to test a point pattern on spatial randomness against an alternative favoring collinearities. We carry out this program in case of a planar Poisson or binomial point process and follow [31, Sect. 5], where the asymptotic behavior of the number of so-called flat triangles in a binomial point process has been investigated.

Let K be a convex body in the plane and let μ be a probability measure on K which has a continuous density φ with respect to the Lebesgue measure $\ell_2|_K$ restricted to K. By η_t we denote a Poisson point process with intensity measure $\mu_t := t\mu$, $t > 0$, and by ζ_n a binomial process of $n \geq 1$ points which are independent and identically distributed according to μ. For a triple (x_1, x_2, x_3) of distinct points of η_t or ζ_n we let $\theta(x_1, x_2, x_3)$ be the largest angle of the triangle formed by x_1, x_2 and x_3. We can now build the point processes

$$\xi_t := \frac{1}{6} \sum_{(x_1,x_2,x_3)\in\eta^3_{t,\neq}} \delta_{\pi-\theta(x_1,x_2,x_3)}$$

and

$$\hat{\xi}_n := \frac{1}{6} \sum_{(x_1,x_2,x_3)\in\zeta^3_{n,\neq}} \delta_{\pi-\theta(x_1,x_2,x_3)}$$

on the positive real half-line. The interpretation is as follows: if for a triple (x_1, x_2, x_3) in $\eta^3_{t,\neq}$ or $\zeta^3_{n,\neq}$ the value $\pi - \theta(x_1, x_2, x_3)$ is small, then the triangle formed by these points is flat in the sense that its height on the longest side is small.

Corollary 6 *Define*

$$\beta := \int_K \int_K \int_0^1 s(1-s)\,\varphi(sx_1 + (1-s)x_2)\,\|x_1 - x_2\|^2 \,\mathrm{d}s\,\mu(\mathrm{d}x_1)\,\mu(\mathrm{d}x_2).$$

Further assume that the density φ is Lipschitz continuous. Then the rescaled point processes $t^3\xi_t$ and $n^3\hat{\xi}_n$ both converge in distribution to a homogeneous Poisson point process on $\mathbb{R}_+$ with intensity β, as $t \to \infty$ or $n \to \infty$, respectively. In addition, there is a constant $C_y > 0$ depending on K, φ and y such that

$$\left| \mathbb{P}(t^3 M_t^{(m)} > y) - e^{-\beta y} \sum_{i=0}^{m-1} \frac{(\beta y)^i}{i!} \right| \leq C_y\, t^{-1}$$

and

$$\left|\mathbb{P}(n^3 M_n^{(m)} > y) - e^{-\beta y}\sum_{i=0}^{m-1}\frac{(\beta y)^i}{i!}\right| \le C_y\, n^{-1}$$

for all $t \geq 1$, $n \geq 3$ *and* $m \in \mathbb{N}$.

Proof To apply Corollary 1 we have to consider the limit behavior of $\alpha_t(y)$ and $r_t(y)$ for fixed $y > 0$, as $t \to \infty$. For $x_1, x_2 \in K$ and $\varepsilon > 0$ define $A(x_1, x_2, \varepsilon)$ as the set of all $x_3 \in K$ such that $\pi - \theta(x_1, x_2, x_3) \le \varepsilon$. Then we have

$$\alpha_t(y) = \frac{t^3}{6}\int_K\int_K\int_K \mathbb{1}\{x_3 \in A(x_1, x_2, yt^{-\gamma})\}\, \varphi(x_1)\varphi(x_2)\varphi(x_3)\,\mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1\,.$$

Without loss of generality we can assume that x_3 is the vertex adjacent to the largest angle. We indicate this by writing $x_3 = \mathrm{LA}(x_1, x_2, x_3)$. We parametrize x_3 by its distance h to the line through x_1 and x_2 and the projection of x_3 onto that line, which can be represented as $sx_1 + (1-s)x_2$ for some $s \in [0, 1]$. Writing $x_3 = x_3(s, h)$, we obtain that

$$\begin{aligned}\alpha_t(y) = \frac{t^3}{2}\int_K\int_K\int_0^1\int_{-\infty}^{\infty} &\mathbb{1}\{x_3(s,h) \in A(x_1, x_2, yt^{-\gamma}), x_3 = \mathrm{LA}(x_1, x_2, x_3)\}\\ &\times \varphi(x_1)\varphi(x_2)\varphi(x_3(s,h))\|x_1 - x_2\|\,\mathrm{d}h\,\mathrm{d}s\,\mathrm{d}x_2\,\mathrm{d}x_1\,.\end{aligned}$$

The sum of the angles at x_1 and x_2 is given by

$$\arctan(|h|/(s\|x_1 - x_2\|)) + \arctan(|h|/((1-s)\|x_1 - x_2\|))\,.$$

Using, for $x \ge 0$, the elementary inequality $x - x^2 \le \arctan x \le x$, we deduce that

$$\begin{aligned}&\frac{|h|}{s(1-s)\|x_1 - x_2\|} - \frac{h^2}{s^2(1-s)^2\|x_1 - x_2\|^2}\\ &\quad\le \arctan(|h|/(s\|x_1 - x_2\|)) + \arctan(|h|/((1-s)\|x_1 - x_2\|))\\ &\quad\le \frac{|h|}{s(1-s)\|x_1 - x_2\|}\,.\end{aligned}$$

Consequently, $\pi - \theta(x_1, x_2, x_3(s, h)) \le yt^{-\gamma}$ is satisfied if

$$|h| \le s(1-s)\|x_1 - x_2\| yt^{-\gamma}$$

and cannot hold if

$$|h| \geq s(1-s)\|x_1 - x_2\|(yt^{-\gamma} + 2y^2 t^{-2\gamma})$$

and t is sufficiently large. Let $A_{y,t}$ be the set of all $x_1, x_2 \in K$ such that

$$B^d_{\tan(t^{-\gamma}y/2)\|x_1-x_2\|}(x_1),\ B^d_{\tan(t^{-\gamma}y/2)\|x_1-x_2\|}(x_2) \subset K\,.$$

Now the previous considerations yield that, for t sufficiently large and $(x_1, x_2) \in A_{y,t}$,

$$\begin{aligned}
&\frac{t^3}{2}\int_0^1\int_{-\infty}^{\infty} \mathbb{1}\{x_3(s,h) \in A(x_1,x_2,yt^{-\gamma}), x_3(s,h) = \mathrm{LA}(x_1,x_2,x_3)\} \\
&\qquad \times \|x_1 - x_2\|\,\varphi(x_1)\varphi(x_2)\varphi(x_3(s,h))\,\mathrm{d}h\,\mathrm{d}s \\
&\quad = t^3 \int_0^1 \big(s(1-s)\|x_1-x_2\|yt^{-\gamma} + R(x_1,x_2,s)\big)\|x_1 - x_2\| \\
&\qquad \times \varphi(x_1)\varphi(x_2)\varphi(sx_1 + (1-s)x_2)\,\mathrm{d}s \\
&\qquad + \frac{t^3}{2}\int_0^1\int_{-\infty}^{\infty} \mathbb{1}\{x_3(s,h) \in A(x_1,x_2,yt^{-\gamma}), x_3(s,h) = \mathrm{LA}(x_1,x_2,x_3)\}\|x_1 - x_2\| \\
&\qquad \times \varphi(x_1)\varphi(x_2)\big(\varphi(x_3(s,h)) - \varphi(sx_1 + (1-s)x_2)\big)\,\mathrm{d}h\,\mathrm{d}s
\end{aligned}$$

with $R(x_1,x_2,s)$ satisfying the estimate $|R(x_1,x_2,s)| \leq 2s(1-s)\|x_1-x_2\|y^2t^{-2\gamma}$. For $(x_1,x_2) \notin A_{y,t}$ the right hand-side is an upper bound. The choice $\gamma = 3$ leads to

$$\begin{aligned}
&|\alpha_t(y) - \beta y| \\
&\quad \leq \int_{K^2\setminus A_{y,t}}\int_0^1 s(1-s)\|x_1-x_2\|^2 y\,\varphi(x_1)\varphi(x_2)\varphi(sx_1 + (1-s)x_2)\,\mathrm{d}s\,\mathrm{d}(x_1,x_2) \\
&\qquad + 2t^{-3}\int_{K^2}\int_0^1 s(1-s)y^2\|x_1-x_2\|^2\,\varphi(x_1)\varphi(x_2)\varphi(sx_1 + (1-s)x_2)\,\mathrm{d}s\,\mathrm{d}(x_1,x_2) \\
&\qquad + \frac{t^3}{2}\int_{K^2}\int_0^1\int_{-\infty}^{\infty} \mathbb{1}\{x_3(s,h) \in A(x_1,x_2,yt^{-\gamma})\}\|x_1-x_2\|\,\varphi(x_1)\varphi(x_2) \\
&\qquad \times \big|\varphi(x_3(s,h)) - \varphi(sx_1 + (1-s)x_2)\big|\,\mathrm{d}h\,\mathrm{d}s\,\mathrm{d}(x_1,x_2)\,.
\end{aligned}$$

Note that $\ell_2^2(K \setminus A_{y,t})$ is of order t^{-3} so that the first integral on the right-hand side is of the same order. By the Lipschitz continuity of the density φ there is a constant $C_\varphi > 0$ such that

$$|\varphi(x_3(s,h)) - \varphi(sx_1 + (1-s)x_2)| \le C_\varphi h\,.$$

This implies that the third integral is of order t^{-3}. Combined with the fact that also the second integral above is of order t^{-3}, we see that there is a constant $C_{y,1} > 0$ such that

$$|\alpha_t(y) - \beta y| \le C_{y,1} t^{-3}$$

for $t \ge 1$.

For given $x_1, x_2 \in K$, we have that

$$\int_K \mathbb{1}\{x_3 \in A(x_1, x_2, yt^{-\gamma})\}\,\varphi(x_3)\,\mathrm{d}x_3 \le M \int_K \mathbb{1}\{x_3 \in A(x_1, x_2, yt^{-\gamma})\}\,\mathrm{d}x_3$$

with $M = \sup_{z\in K} \varphi(z)$. By the same arguments as above, we see that the integral over all x_3 such that the largest angle is adjacent to x_3 is bounded by

$$\begin{aligned} &M \int_0^1 s(1-s)\|x_1 - x_2\| yt^{-3} + 2s(1-s)\|x_1 - x_2\| y^2 t^{-6} \mathrm{d}s \\ &\quad \le 2M\mathrm{diam}(K)(yt^{-3} + 2y^2 t^{-6})\,, \end{aligned}$$

where $\mathrm{diam}(K)$ stands for the diameter of K. The maximal angle is at x_1 or x_2 if x_3 is contained in the union of two cones with opening angle $2t^{-3}y$ and apices at x_1 and x_2, respectively. The integral over these x_3 is bounded by $2M\mathrm{diam}(K)^2 t^{-3} y$. Altogether, we obtain that

$$\begin{aligned} &\int_K \mathbb{1}\{x_3 \in A(x_1, x_2, yt^{-\gamma})\}\varphi(x_3)\,\mathrm{d}x_3 \\ &\quad \le 2M\mathrm{diam}(K)(yt^{-3} + 2y^2 t^{-6}) + 2M\mathrm{diam}(K)^2 yt^{-3}. \end{aligned}$$

This estimate implies that, for any $\ell \in \{1, 2\}$,

$$\begin{aligned} &t^\ell \int_{K^\ell} \Big(t^{3-\ell} \int_{K^{3-\ell}} \mathbb{1}\{x_3 \in A(x_1, x_2, yt^{-3})\}\,\mu^{3-\ell}(\mathrm{d}(K_{\ell+1}, \ldots, K_3)) \Big)^2 \mu^\ell(\mathrm{d}(K_1, \ldots, K_\ell)) \\ &\quad \le t^{6-\ell}(M\ell_2(K))^{4-\ell}\big(2M\mathrm{diam}(K)(yt^{-3} + 2y^2 t^{-6}) + 2M\mathrm{diam}(K)^2 yt^{-3}\big)^2\,. \end{aligned}$$

Since the upper bound behaves like $t^{-\ell}$ for $t \geq 1$, there is a constant $C_{y,2} > 0$ such that

$$r_t(y) \leq C_{y,2} t^{-1}$$

for $t \geq 1$. Now an application of Corollary 1 concludes the proof in case of an underlying Poisson point process. The binomial case can be handled similarly using Corollary 2. □

Remark 6 We have assumed that the density φ is Lipschitz continuous. If this is not the case, one can still show that the rescaled point processes $t^3 \xi_t$ and $n^3 \hat{\xi}_n$ converge in distribution to a homogeneous Poisson point process on $\mathbb{R}_+$ with intensity β. However, we are then no more able to provide a rate of convergence for the associated order statistics $M_t^{(m)}$.

Remark 7 In [31, Sect. 5] the asymptotic behavior of the number of flat triangles in a binomial point process has been investigated, while our focus here was on the angle statistic of such triangles. However, these two random variables are asymptotically equivalent so that Corollary 6 also delivers an alternative approach to the results in [31]. In addition, it allows to deal with an underlying Poisson point process, where it provides rates of convergence in the case of a Lipschitz density.

3.4 Non-Intersecting k-Flats

Fix a space dimension $d \geq 3$ and let $k \geq 1$ be such that $2k < d$. By $G(d,k)$ let us denote the space of k-dimensional linear subspaces of $\mathbb{R}^d$, which is equipped with a probability measure ς. In what follows we shall assume that ς is absolutely continuous with respect to the Haar probability measure on $G(d,k)$. The space of k-dimensional affine subspaces of $\mathbb{R}^d$ is denoted by $A(d,k)$ and for $t > 0$ a translation-invariant measure μ_t on $A(d,k)$ is defined by the relation

$$\int_{A(d,k)} g(E)\, \mu_t(\mathrm{d}E) = t \int_{G(d,k)} \int_{L^\perp} g(L+x)\, \ell_{d-k}(\mathrm{d}x)\, \varsigma(\mathrm{d}L)\,, \qquad (7)$$

where $g \geq 0$ is a measurable function on $A(d,k)$. We will use E and F to indicate elements of $A(d,k)$, while L and M will stand for linear subspaces in $G(d,k)$, see [11, formula (1)] in this book. We also put $\mu = \mu_1$. For two fixed k-flats $E, F \in A(d,k)$ we denote by $d(E,F) = \inf\{\|x_1 - x_2\| : x_1 \in E, x_2 \in F\}$ the distance of E and F. For almost all E and F it is realized by two uniquely determined points $x_E \in E$ and $x_F \in F$, i.e. $d(E,F) = \|x_E - x_F\|$, and we let $m(E,F) := (x_E + x_F)/2$ be the midpoint of the line segment joining x_E with x_F.

Let $K \subset \mathbb{R}^d$ be a convex body and let η_t be a Poisson point process on $A(d,k)$ with intensity measure μ_t as defined in (7). We will speak about η_t as a Poisson k-flat process and denote, more generally, the elements of $A(d,k)$ or $G(d,k)$ as k-flats. We will not treat the binomial case in what follows since the measures μ_t are not finite. We notice that in view of [27, Theorem 4.4.5 (c)] any two k-flats of η_t are almost surely in general position, a fact which from now on will be used without further comment.

Point processes of k-dimensional flats in $\mathbb{R}^d$ have a long tradition in stochastic geometry and we refer to [6] or [27] as well as to [11] for general background material. Moreover, we mention the works [10, 26], which deal with distance measurements and the so-called proximity of Poisson k-flat processes and are close to what we consider here. While in these papers only mean values are considered, we are interested in the point process ξ_t on $\mathbb{R}_+$ defined by

$$\xi_t := \frac{1}{2} \sum_{(E,F)\in\eta^2_{t,\neq}} \delta_{d(E,F)^a}\, \mathbb{1}\{m(E,F)\in K\}$$

for a fixed parameter $a > 0$. A particular case arises when $a = 1$. Then $M_t^{(1)}$, for example, is the smallest distance between two k-flats from η_t that have their midpoint in K.

Corollary 7 *Define*

$$\beta = \frac{\ell_d(K)}{2}\kappa_{d-2k} \int\limits_{G(d,k)} \int\limits_{G(d,k)} [L,M]\,\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M)\,,$$

where $[L,M]$ is the $2k$-dimensional volume of a parallelepiped spanned by two orthonormal bases in L and M. Then, as $t \to \infty$, $t^{2a/(d-2k)}\xi_t$ converges in distribution to a Poisson point process on $\mathbb{R}_+$ with intensity measure

$$B \mapsto (d-2k)\frac{\beta}{a}\int\limits_B u^{(d-2k-a)/a}\,\mathrm{d}u\,, \qquad B\subset\mathbb{R}_+ \quad \text{Borel}\,.$$

Moreover, there is a constant $C > 0$ depending on K, ς and a such that

$$\left|\mathbb{P}(t^{2a/(d-2k)}M_t^{(m)} > y) - \exp\left(-\beta y^{(d-2k)/a}\right)\sum_{i=1}^{m-1}\frac{\left(\beta y^{(d-2k)/a}\right)^i}{i!}\right|$$
$$\leq C\,(y^{2(d-2k)/a} + y^{d-k+2(d-2k)/a})\,t^{-1}$$

for any $t \geq 1$, $y \geq 0$ and $m \in \mathbb{N}$.

Proof For $y > 0$ and $t > 0$ we have that

$$\alpha_t(y) = \frac{t^2}{2} \int\limits_{A(d,k)} \int\limits_{A(d,k)} \mathbb{1}\{d(E,F) \le y^{1/a} t^{-\gamma/a},\, m(E,F) \in K\}\, \mu(\mathrm{d}E)\mu(\mathrm{d}F)\,.$$

We abbreviate $\delta := y^{1/a} t^{-\gamma/a}$ and evaluate the integral

$$\mathcal{I} := \int\limits_{A(d,k)} \int\limits_{A(d,k)} \mathbb{1}\{d(E,F) \le \delta,\, m(E,F) \in K\}\, \mu(\mathrm{d}E)\mu(\mathrm{d}F)\,.$$

For this, we define $V := E + F$ and $U := V^{\perp}$ and write E and F as $E = L + x_1$ and $F = M + x_2$ with $L, M \in G(d,k)$ and $x_1 \in L^{\perp}$, $x_2 \in M^{\perp}$. Applying now the definition (7) of the measure μ and arguing along the lines of the proof of Theorem 4.4.10 in [27], we arrive at the expression

$$\mathcal{I} = \int\limits_{G(d,k)} \int\limits_{G(d,k)} \int\limits_{U} \int\limits_{U} [L,M]\, \ell_{2k}\left(K \cap \left(V + \left(\frac{x_1 + x_2}{2}\right)\right)\right)$$
$$\times\, \mathbb{1}\{\|x_1 - x_2\| \le \delta\}\, \ell_{d-2k}(\mathrm{d}x_1)\ell_{d-2k}(\mathrm{d}x_2)\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M)\,.$$

Substituting $u = x_1 - x_2$, $v = (x_1 + x_2)/2$ (a transformation having Jacobian equal to 1), we find that

$$\mathcal{I} = \int\limits_{G(d,k)} \int\limits_{G(d,k)} \int\limits_{U} \int\limits_{U} [L,M]\, \ell_{2k}\big(K \cap (V + v)\big)\, \mathbb{1}\{\|u\| \le \delta\} \qquad (8)$$
$$\ell_{d-2k}(\mathrm{d}u)\ell_{d-2k}(\mathrm{d}v)\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M)\,.$$

Since U has dimension $d - 2k$, transformation into spherical coordinates in U gives

$$\int\limits_{U} \mathbf{1}(\|u\| \le \delta)\, \mathrm{d}u = (d - 2k)\kappa_{d-2k} \int\limits_{0}^{\delta} r^{d-2k-1}\, \mathrm{d}r = \kappa_{d-2k}\delta^{d-2k}\,.$$

Moreover,

$$\int\limits_{U} \ell_{2k}\big(K \cap (V + v)\big)\, \ell_{d-2k}(\mathrm{d}v) = \ell_d(K)$$

since $V = U^\perp$. Combining these facts with (8) we find that

$$\mathcal{I} = \delta^{d-2k}\,\ell_d(K)\,\kappa_{d-2k}\int\limits_{G(d,k)}\int\limits_{G(d,k)}[L,M]\,\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M)$$

and that

$$\alpha_t(y) = \frac{1}{2}\,\ell_d(K)\,\kappa_{d-2k}\,y^{(d-2k)/a}\,t^{2-\gamma(d-2k)/a}\int\limits_{G(d,k)}\int\limits_{G(d,k)}[L,M]\,\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M)\,.$$

Consequently, choosing $\gamma = 2a/(d-2k)$ we have that

$$\alpha_t(y) = \beta y^{(d-2k)/a}\,.$$

For the remainder term $r_t(y)$ we write

$$r_t(y) = t\int\limits_{A(d,k)}\Bigg(t\int\limits_{A(d,k)}\mathbb{1}\{d(E,F)^a \le yt^{-\gamma},\, m(E,F)\in K\}\,\mu(\mathrm{d}F)\Bigg)^2\mu(\mathrm{d}E)\,.$$

This can be estimated along the lines of the proof of Theorem 3 in [30]. Namely, using that $[\,\cdot\,,\,\cdot\,] \le 1$ and writing $\mathrm{diam}(K)$ for the diameter of K, we find that

$$\begin{aligned}
r_t(y) &\le t\kappa_{d-k}(\mathrm{diam}(K) + 2t^{-\gamma}y)^{d-k}\int\limits_{G(d,k)}\Bigg(t\int\limits_{G(d,k)}\int\limits_{(L+M)^\perp}\mathbb{1}\{\|x\|^a \le yt^{-\gamma}\}\\
&\qquad\times\kappa_k(\mathrm{diam}(K)/2)^k\,\ell_{d-2k}(\mathrm{d}x)\varsigma(\mathrm{d}M)\Bigg)^2\varsigma(\mathrm{d}L)\\
&\le t\kappa_{d-k}(\mathrm{diam}(K) + 2t^{-\gamma}y)^{d-k}\big(t\kappa_{d-2k}(yt^{-\gamma})^{(d-2k)/a}\kappa_k(\mathrm{diam}(K)/2)^k\big)^2\\
&= \kappa_{d-k}(\mathrm{diam}(K) + 2t^{-2a/(d-2k)}y)^{d-k}\kappa_{d-2k}^2\kappa_k^2(\mathrm{diam}(K)/2)^{2k}\,y^{2(d-2k)/a}\,t^{-1}\,,
\end{aligned}$$

where we have used that $\gamma = 2a/(d-2k)$. This puts us in the position to apply Corollary 1, which completes the proof. □

Remark 8 A particularly interesting case arises when the distribution ς coincides with the Haar probability measure on $G(d,k)$. Then the double integral in the definition of β in Corollary 7 can be evaluated explicitly, namely we have

$$\int\limits_{G(d,k)}\int\limits_{G(d,k)}[L,M]\,\varsigma(\mathrm{d}L)\varsigma(\mathrm{d}M) = \frac{\binom{d-k}{k}\kappa_{d-k}^2}{\binom{d}{k}\kappa_d\kappa_{d-2k}}$$

according to [13, Lemma 4.4].

Remark 9 Corollary 7 generalizes Theorem 4 in [30] (where the case $a = 1$ has been investigated) to general length-powers $a > 0$. However, it should be noticed that the set-up in [30] slightly differs from the one here. In [30] the intensity parameter t was kept fixed, whereas the set K was increased by dilations. But because of the scaling properties of a Poisson k-flat process and the a-homogeneity of $d(E, F)^a$, one can translate one result into the other. Moreover, we refer to [14] for closely related results including directional constraints.

Remark 10 In [29] a similar problem has been addressed in the case where ς coincides with the Haar probability measure on $G(d, k)$. For a pair $(E, F) \in \eta^2_{t,\neq}$ satisfying $E \cap K \neq \emptyset$ and $F \cap K \neq \emptyset$, the distance between E and F was measured by

$$d_K(E, F) = \inf\{\|x_1 - x_2\| : x_1 \in E \cap K,\ x_2 \in F \cap K\},$$

and it has been shown in Theorem 2.1 ibidem that the associated point process

$$\xi_t := \frac{1}{2} \sum_{(E,F)\in\eta^2_{t,\neq}} \delta_{d_K(E,F)}\, \mathbb{1}\{E \cap K \neq \emptyset,\ F \cap K \neq \emptyset\}$$

converges, after rescaling with $t^{2/(d-2k)}$, towards the same Poisson point process as in Corollary 7 when ς is the Haar probability measure on $G(d, k)$ and $a = 1$.

4 Proofs of the Main Results

4.1 Moment Formulas for Poisson U-Statistics

We call a Poisson functional S of the form

$$S = \sum_{(x_1,\ldots,x_k)\in\eta^k_{t,\neq}} f(x_1, \ldots, x_k)$$

with $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $f : \mathbb{X}^k \to \mathbb{R}$ a U-statistic of order k of η_t, or a Poisson U-statistic for short (see [17]). For $k = 0$ we use the convention that f is a constant and $S = f$. In the following, we always assume that f is integrable. Moreover, without loss of generality we assume that f is symmetric since we sum over all permutations of a fixed k-tuple of points in the definition of S.

In order to compute mixed moments of Poisson U-statistics, we use the following notation. For $\ell \in \mathbb{N}$ and $n_1, \ldots, n_\ell \in \mathbb{N}_0$ we define $N_0 = 0$, $N_i = \sum_{j=1}^{i} n_j$, $i \in \{1, \ldots, \ell\}$, and

$$J_i = \begin{cases} \{N_{i-1} + 1, \ldots, N_i\}, & N_{i-1} < N_i \\ \emptyset, & N_{i-1} = N_i \end{cases}, \quad i \in \{1, \ldots, \ell\}.$$

Let $\Pi(n_1,\dots,n_\ell)$ be the set of all partitions σ of $\{1,\dots,N_\ell\}$ such that for any $i \in \{1,\dots,\ell\}$ all elements of J_i are in different blocks of σ. By $|\sigma|$ we denote the number of blocks of σ. We say that two blocks B_1 and B_2 of a partition $\sigma \in \Pi(n_1,\dots,n_\ell)$ intersect if there is an $i \in \{1,\dots,\ell\}$ such that $B_1 \cap J_i \neq \emptyset$ and $B_2 \cap J_i \neq \emptyset$. A partition $\sigma \in \Pi(n_1,\dots,n_\ell)$ with blocks $B_1,\dots,B_{|\sigma|}$ belongs to $\tilde{\Pi}(n_1,\dots,n_\ell)$ if there are no nonempty sets $M_1,M_2 \subset \{1,\dots,|\sigma|\}$ with $M_1 \cap M_2 = \emptyset$ and $M_1 \cup M_2 = \{1,\dots,|\sigma|\}$ such that for any $i \in M_1$ and $j \in M_2$ the blocks B_i and B_j do not intersect. Moreover, we define

$$\Pi_{\neq}(n_1,\dots,n_\ell) = \{\sigma \in \Pi(n_1,\dots,n_\ell) : |\sigma| > \min\{n_1,\dots,n_\ell\}\}.$$

If there are $i,j \in \{1,\dots,\ell\}$ with $n_i \neq n_j$, we have $\Pi_{\neq}(n_1,\dots,n_\ell) = \Pi(n_1,\dots,n_\ell)$.

For $\sigma \in \Pi(n_1,\dots,n_\ell)$ and $f : \mathbb{X}^{N_\ell} \to \mathbb{R}$ we define $f_\sigma : \mathbb{X}^{|\sigma|} \to \mathbb{R}$ as the function which arises by replacing in the arguments of f all variables belonging to the same block of σ by a new common variable. Since we are only interested in the integral of this new function in the sequel, the order of the new variables does not matter. For $f^{(i)} : \mathbb{X}^{n_i} \to \mathbb{R}$, $i \in \{1,\dots,\ell\}$, let $\otimes_{i=1}^{\ell} f^{(i)} : \mathbb{X}^{N_\ell} \to \mathbb{R}$ be given by

$$\big(\otimes_{i=1}^{\ell} f^{(i)} \big)(x_1,\dots,x_{N_\ell}) = \prod_{i=1}^{\ell} f^{(i)}(x_{N_{i-1}+1},\dots,x_{N_i}) \,.$$

The following lemma allows us to compute moments of Poisson U-statistics (see also [23]. Here and in what follows we mean by a Poisson functional $F = F(\eta_t)$ a random variable only depending on the Poisson point process η_t for some fixed $t > 0$.

Lemma 1 *For $\ell \in \mathbb{N}$ and $f^{(i)} \in L_s^1(\mu_t^{k_i})$ with $k_i \in \mathbb{N}_0$, $i = 1,\dots,\ell$, such that*

$$\int_{\mathbb{X}^{|\sigma|}} |\big(\otimes_{i=1}^{\ell} f^{(i)} \big)_\sigma| \, \mathrm{d}\mu_t^{|\sigma|} < \infty \qquad \text{for all} \qquad \sigma \in \Pi(k_1,\dots,k_\ell) \,,$$

let

$$S_i = \sum_{(x_1,\dots,x_{k_i}) \in \eta_{t,\neq}^{k_i}} f^{(i)}(x_1,\dots,x_{k_i}), \quad i = 1,\dots,\ell \,,$$

and let F be a bounded Poisson functional. Then

$$\mathbb{E}\Big[F \prod_{i=1}^{\ell} S_i\Big] = \sum_{\sigma \in \Pi(k_1,\dots,k_\ell)} \int_{\mathbb{X}^{|\sigma|}} \big(\otimes_{i=1}^{\ell} f^{(i)} \big)_\sigma (x_1,\dots,x_{|\sigma|}) \times \mathbb{E}\left[F\left(\eta_t + \sum_{i=1}^{|\sigma|} \delta_{x_i} \right) \right] \mu_t^{|\sigma|}(\mathrm{d}(x_1,\dots,x_{|\sigma|})) \,.$$

Proof We can rewrite the product as

$$F(\eta_t)\prod_{i=1}^{\ell}\sum_{(x_1,\dots,x_{k_i})\in\eta_{t,\neq}^{k_i}} f^{(i)}(x_1,\dots,x_{k_i})$$
$$=\sum_{\sigma\in\Pi(k_1,\dots,k_\ell)}\sum_{(x_1,\dots,x_{|\sigma|})\in\eta_{t,\neq}^{|\sigma|}}\big(\otimes_{i=1}^{\ell} f^{(i)}\big)_\sigma(x_1,\dots,x_{|\sigma|})F(\eta_t)$$

since points occurring in different sums on the left-hand side can be either equal or distinct. Now an application of the multivariate Mecke formula (see [18, formula (1.11)]) completes the proof of the lemma. □

4.2 Poisson Approximation of Poisson U-Statistics

The key argument of the proof of Theorem 1 is a quantitative bound for the Poisson approximation of Poisson U-statistics which is established in this subsection. From now on we consider the Poisson U-statistic

$$S_A=\frac{1}{k!}\sum_{(x_1,\dots,x_k)\in\eta_{t,\neq}^{k}}\mathbb{1}\{f(x_1,\dots,x_k)\in A\}\,,$$

where f is as in Sect. 2 and $A\subset\mathbb{R}$ is measurable and bounded. We assume that $k\geq 2$ since S_A follows a Poisson distribution for $k=1$ (see Sect. 2.3 in [16], for example). In the sequel, we use the abbreviation

$$h(x_1,\dots,x_k):=\frac{1}{k!}\mathbb{1}\{f(x_1,\dots,x_k)\in A\},\quad x_1,\dots,x_k\in\mathbb{X}\,.$$

It follows from the multivariate Mecke formula (see [18, formula (1.11)]) that

$$s_A:=\mathbb{E}[S_A]=\int\limits_{\mathbb{X}^k} h(x_1,\dots,x_k)\,\mu_t^k(\mathrm{d}(x_1,\dots,x_k))\,.$$

In order to compare the distributions of two integer-valued random variables Y and Z, we use the so-called total variation distance d_{TV} defined by

$$d_{\mathrm{TV}}(Y,Z)=\sup_{B\subset\mathbb{Z}}\big|\mathbb{P}(Y\in B)-\mathbb{P}(Z\in B)\big|\,.$$

Proposition 1 *Let S_A be as above, let Y be a Poisson distributed random variable with mean $s > 0$ and define*

$$\varrho_A := \max_{1\le \ell \le k-1} \int_{\mathbb{X}^\ell} \bigg(\int_{\mathbb{X}^{k-\ell}} h(x_1,\dots,x_k)\, \mu_t^{k-\ell}\big(\mathrm{d}(x_{\ell+1},\dots,x_k)\big)\bigg)^2 \mu_t^\ell\big(\mathrm{d}(x_1,\dots,x_\ell)\big).$$

Then there is a constant $C \ge 1$ only depending on k such that

$$d_{\mathrm{TV}}(S_A, Y) \le |s_A - s| + C \min\left\{1, \frac{1}{s_A}\right\} \varrho_A\,. \tag{9}$$

Remark 11 The inequality (9) still holds if Y is almost surely zero (such a Y can be interpreted as a Poisson distributed random variable with mean $s = 0$). In this case, we obtain by Markov's inequality that

$$d_{\mathrm{TV}}(S_A, Y) = \mathbb{P}(S_A \ge 1) \le \mathbb{E} S_A = s_A.$$

Our proof of Proposition 1 is a modification of the proof of Theorem 3.1 in [21]. It makes use of the special structure of S_A and improves of the bound in [21] in case of Poisson U-statistics. To prepare for what follows, we need to introduce some facts around the Chen–Stein method for Poisson approximation (compare with [3]). For a function $f : \mathbb{N}_0 \to \mathbb{R}$ let us define $\Delta f(k) := f(k+1) - f(k)$, $k \in \mathbb{N}_0$, and $\Delta^2 f(k) := f(k+2) - 2f(k+1) + f(k)$, $k \in \mathbb{N}_0$. For $B \subset \mathbb{N}_0$ let f_B be the solution of the Chen–Stein equation

$$\mathbb{1}\{k \in B\} - \mathbb{P}(Y \in B) = s f(k+1) - k f(k), \quad k \in \mathbb{N}_0\,. \tag{10}$$

It is known (see Lemma 1.1.1 in [2]) that f_B satisfies

$$\|f_B\|_\infty \le 1 \quad \text{and} \quad \|\Delta f_B\|_\infty \le \min\left\{1, \frac{1}{s}\right\} =: \varepsilon_1\,, \tag{11}$$

where $\|\cdot\|_\infty$ is the usual supremum norm.

Besides the Chen–Stein method we need some facts concerning the Malliavin calculus of variations on the Poisson space (see [18]). First, the so-called integration by parts formula implies that

$$\mathbb{E}[f_B(S_A)(S_A - \mathbb{E}[S_A])] = \mathbb{E} \int_{\mathbb{X}} D_x f_B(S_A)(-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x)\,, \tag{12}$$

where D stands for the difference operator and L^{-1} is the inverse of the Ornstein–Uhlenbeck generator (this step requires that $\mathbb{E}\int_{\mathbb{X}} (D_x S_A)^2\, \mu_t(\mathrm{d}x) < \infty$, which is a consequence of the calculations in the proof of Proposition 1). The following lemma (see Lemma 3.3 in [24]) implies that the difference operator applied to a Poisson U-statistic leads again to a Poisson U-statistic.

Lemma 2 *Let* $k \in \mathbb{N}$, $f \in L_s^1(\mu_t^k)$ *and*

$$S = \sum_{(x_1,\dots,x_k)\in\eta_{t,\neq}^k} f(x_1,\dots,x_k)\,.$$

Then

$$D_xS = k \sum_{(x_1,\dots,x_{k-1})\in\eta_{t,\neq}^{k-1}} f(x,x_1,\dots,x_{k-1})\,, \quad x \in \mathbb{X}\,.$$

Proof It follows from the definition of the difference operator and the assumption that f is a symmetric function that

$$\begin{aligned} D_xS &= \sum_{(x_1,\dots,x_k)\in(\eta_t+\delta_x)_{\neq}^k} f(x_1,\dots,x_k) - \sum_{(x_1,\dots,x_k)\in\eta_{t,\neq}^k} f(x_1,\dots,x_k) \\ &= \sum_{(x_1,\dots,x_{k-1})\in\eta_{t,\neq}^{k-1}} \big(f(x,x_1,\dots,x_{k-1}) + \dots + f(x_1,\dots,x_{k-1},x)\big) \\ &= k \sum_{(x_1,\dots,x_{k-1})\in\eta_{t,\neq}^{k-1}} f(x,x_1,\dots,x_{k-1}) \end{aligned}$$

for $x \in \mathbb{X}$. This completes the proof. □

In order to derive an explicit formula for the combination of the difference operator and the inverse of the Ornstein–Uhlenbeck generator of S_A, we define $h_\ell : \mathbb{X}^\ell \to \mathbb{R}$, $\ell \in \{1,\dots,k\}$, by

$$h_\ell(x_1,\dots,x_\ell) := \int_{\mathbb{X}^{k-\ell}} h(x_1,\dots,x_\ell,\hat{x}_1,\dots,\hat{x}_{k-\ell})\,\mu_t^{k-\ell}(\mathrm{d}(\hat{x}_1,\dots,\hat{x}_{k-\ell}))\,.$$

We shall see now that the operator $-DL^{-1}$ applied to S_A can be expressed as a sum of Poisson U-statistics (see also Lemma 5.1 in [28]).

Lemma 3 *For* $x \in \mathbb{X}$,

$$-D_xL^{-1}S_A = \sum_{\ell=1}^{k} \sum_{(x_1,\dots,x_{\ell-1})\in\eta_{t,\neq}^{\ell-1}} h_\ell(x,x_1,\dots,x_{\ell-1})\,.$$

Proof By Mehler's formula (see Theorem 3.2 in [20] and also [18, Sect. 1.7]) we have

$$-L^{-1}S_A = \int_0^1 \int \frac{1}{s}\mathbb{E}\left[\sum_{(x_1,\dots,x_k)\in(\eta_t^{(s)}+\chi)^k_{\neq}} h(x_1,\dots,x_k) - s_A \Big| \eta_t\right] \mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s$$

where $\eta_t^{(s)}$, $s \in [0,1]$, is an s-thinning of η_t and $\mathbb{P}_{(1-s)\mu_t}$ is the distribution of a Poisson point process with intensity measure $(1-s)\mu_t$. Note in particular that $\eta_t^{(s)}+\chi$ is a Poisson point process with intensity measure $s\mu_t + (1-s)\mu_t = \mu_t$. The last expression can be rewritten as

$$\begin{aligned}-L^{-1}S_A &= \int_0^1 \int \frac{1}{s}\mathbb{E}\Big[\sum_{(\hat{x}_1,\dots,\hat{x}_k)\in\chi^k_{\neq}} h(\hat{x}_1,\dots,\hat{x}_k) - s_A\Big|\eta_t\Big]\mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s\\ &\quad+\sum_{\ell=1}^{k}\binom{k}{\ell}\int_0^1\int\frac{1}{s}\mathbb{E}\Big[\sum_{(x_1,\dots,x_\ell)\in(\eta_t^{(s)})^\ell_{\neq}}\\ &\qquad\sum_{(\hat{x}_1,\dots,\hat{x}_{k-\ell})\in\chi^{k-\ell}_{\neq}} h(x_1,\dots,x_\ell,\hat{x}_1,\dots,\hat{x}_{k-\ell})\Big|\eta_t\Big]\,\mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s\,.\end{aligned}$$

By the multivariate Mecke formula (see [18, formula (1.11)]), we obtain for the first term that

$$\begin{aligned}&\int_0^1\int\frac{1}{s}\mathbb{E}\left[\sum_{(\hat{x}_1,\dots,\hat{x}_k)\in\chi^k_{\neq}} h(\hat{x}_1,\dots,\hat{x}_k) - s_A\Big|\eta_t\right]\mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s\\ &=\int_0^1\int\frac{1}{s}\Big(\sum_{(\hat{x}_1,\dots,\hat{x}_k)\in\chi^k_{\neq}} h(\hat{x}_1,\dots,\hat{x}_k) - s_A\Big)\mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s = \int_0^1\frac{(1-s)^k-1}{s}\,\mathrm{d}s\; s_A\,.\end{aligned}$$

To evaluate the second term further, we notice that for an ℓ-tuple $(x_1,\dots,x_\ell)\in\eta^\ell_{t,\neq}$ the probability of surviving the s-thinning procedure is s^ℓ. Thus

$$\begin{aligned}&\mathbb{E}\left[\sum_{(x_1,\dots,x_\ell)\in(\eta_t^{(s)})^\ell_{\neq}}\ \sum_{(\hat{x}_1,\dots,\hat{x}_{k-\ell})\in\chi^{k-\ell}_{\neq}} h(x_1,\dots,x_\ell,\hat{x}_1,\dots,\hat{x}_{k-\ell})\Big|\eta_t\right]\\ &= s^\ell \sum_{(x_1,\dots,x_\ell)\in\eta^\ell_{t,\neq}}\ \sum_{(\hat{x}_1,\dots,\hat{x}_{k-\ell})\in\chi^{k-\ell}_{\neq}} h(x_1,\dots,x_\ell,\hat{x}_1,\dots,\hat{x}_{k-\ell})\end{aligned}$$

for $\ell \in \{1, \ldots, k\}$. This leads to

$$-L^{-1}S_A = \int_0^1 \frac{(1-s)^k - 1}{s} \,\mathrm{d}s\, s_A + \sum_{\ell=1}^{k} \binom{k}{\ell} \int_0^1 \int s^{\ell-1} \sum_{(x_1,\ldots,x_\ell)\in\eta_{t,\neq}^{\ell}} \sum_{(\hat{x}_1,\ldots,\hat{x}_{k-\ell})\in\chi_{\neq}^{k-\ell}} h(x_1,\ldots,x_\ell,\hat{x}_1,\ldots,\hat{x}_{k-\ell}) \, \mathbb{P}_{(1-s)\mu_t}(\mathrm{d}\chi)\,\mathrm{d}s\,.$$

Finally, we may interpret χ as $(1-s)$-thinning of an independent copy of η_t, in which each point has survival probability $(1-s)$. Then the multivariate Mecke formula ([18, formula (1.11)]) implies that

$$-L^{-1}S_A = \int_0^1 \frac{(1-s)^k - 1}{s} \,\mathrm{d}s\, s_A + \sum_{\ell=1}^{k} \binom{k}{\ell} \int_0^1 s^{\ell-1}(1-s)^{k-\ell}\,\mathrm{d}s \sum_{(x_1,\ldots,x_\ell)\in\eta_{t,\neq}^{\ell}} h_\ell(x_1,\ldots,x_\ell)\,.$$

Together with

$$\int_0^1 s^{\ell-1}(1-s)^{k-\ell}\,\mathrm{d}s = \frac{(\ell-1)!(k-\ell)!}{k!}, \quad \ell \in \{1,\ldots,k\},$$

we see that

$$-L^{-1}S_A = s_A \int_0^1 \frac{(1-s)^k - 1}{s}\,\mathrm{d}s + \sum_{\ell=1}^{k} \frac{1}{\ell} \sum_{(x_1,\ldots,x_\ell)\in\eta_{t,\neq}^{\ell}} h_\ell(x_1,\ldots,x_\ell)\,.$$

Applying now the difference operator to the last equation, we see that the first term does not contribute, whereas the second term can be handled by using Lemma 2. □

Now we are prepared for the proof of Proposition 1.

Proof (of Proposition 1) Let Y_A be a Poisson distributed random variable with mean $s_A > 0$. The triangle inequality for the total variation distance implies that

$$d_{\mathrm{TV}}(S_A, Y) \leq d_{\mathrm{TV}}(Y, Y_A) + d_{\mathrm{TV}}(Y_A, S_A)\,.$$

A standard calculation shows that

$$d_{\mathrm{TV}}(Y, Y_A) \leq |s - s_A|$$

so that it remains to bound

$$d_{\mathrm{TV}}(Y_A, S_A) = \sup_{B \subset \mathbb{N}_0} |\mathbb{P}(S_A \in B) - \mathbb{P}(Y_A \in B)| \,.$$

For a fixed $B \subset \mathbb{N}_0$ it follows from (10) and (12) that

$$\begin{aligned}\mathbb{P}(S_A \in B) - \mathbb{P}(Y_A \in B) &= \mathbb{E}\left[s_A \Delta f_B(S_A) - (S_A - s_A) f_B(S_A)\right] \\ &= \mathbb{E}\left[s_A \Delta f_B(S_A) - \int_{\mathbb{X}} D_x f_B(S_A)(-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x)\right].\end{aligned} \tag{13}$$

Now a straightforward computation using a discrete Taylor-type expansion as in [21] shows that

$$\begin{aligned}D_x f_B(S_A) &= f_B(S_A + D_x S_A) - f_B(S_A) \\ &= \sum_{k=1}^{D_x S_A} \big(f_B(S_A + k) - f_B(S_A + k - 1)\big) \\ &= \sum_{k=1}^{D_x S_A} \Delta f_B(S_A + k - 1) \\ &= \Delta f_B(S_A) D_x S_A + \sum_{k=2}^{D_x S_A} \big(\Delta f_B(S_A + k - 1) - \Delta f_B(S_A)\big).\end{aligned}$$

Together with (11), we obtain that

$$\begin{aligned}\left|\sum_{k=2}^{D_x S_A} \big(\Delta f_B(S_A + k - 1) - \Delta f_B(S_A)\big)\right| &\leq 2\|\Delta f_B\|_\infty \max\{0, D_x S_A - 1\} \\ &\leq 2\varepsilon_{1,A} \max\{0, D_x S_A - 1\}\end{aligned}$$

with

$$\varepsilon_{1,A} := \min\left\{1, \frac{1}{s_A}\right\}.$$

Hence, we have

$$D_x f_B(S_A) = \Delta f_B(S_A) D_x S_A + R_x,$$

where the remainder term satisfies $|R_x| \leq 2\varepsilon_{1,A} \max\{0, D_x S_A - 1\}$. Together with (13) and $-D_x L^{-1} S_A \geq 0$, which follows from Lemma 3, we obtain that

$$\begin{aligned} &|\mathbb{P}(S_A \in B) - \mathbb{P}(Y_A \in B)| \\ &\leq \left| \mathbb{E}\left[s_A \Delta f_B(S_A) - \Delta f_B(S_A) \int\limits_{\mathbb{X}} D_x S_A (-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x) \right] \right| \\ &\quad + 2\varepsilon_{1,A} \int\limits_{\mathbb{X}} \mathbb{E}[\max\{0, D_x S_A - 1\}(-D_x L^{-1} S_A)]\, \mu_t(\mathrm{d}x)\,. \end{aligned} \tag{14}$$

It follows from Lemmas 2 and 3 that

$$\begin{aligned} &\mathbb{E}\left[\Delta f_B(S_A) \int\limits_{\mathbb{X}} D_x S_A (-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x) \right] \\ &= \mathbb{E}\left[\Delta f_B(S_A(\eta_t)) \int\limits_{\mathbb{X}} \left(k \sum_{(x_1,\ldots,x_{k-1}) \in \eta_{t,\neq}^{k-1}} h(x, x_1, \ldots, x_{k-1}) \right) \right. \\ &\qquad \left. \times \left(\sum_{\ell=1}^{k} \sum_{(x_1,\ldots,x_{\ell-1}) \in \eta_{t,\neq}^{\ell-1}} h_\ell(x, x_1, \ldots, x_{\ell-1}) \right) \mu_t(\mathrm{d}x) \right]. \end{aligned}$$

Consequently, we can deduce from Lemma 1 that

$$\begin{aligned} &\mathbb{E}\left[\Delta f_B(S_A) \int\limits_{\mathbb{X}} D_x S_A (-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x) \right] \\ &= k \sum_{\ell=1}^{k} \sum_{\sigma \in \Pi(k-1,\ell-1)} \int\limits_{\mathbb{X}^{|\sigma|+1}} \mathbb{E}\left[\Delta f_B \left(S_A \left(\eta_t + \sum_{i=1}^{|\sigma|} \delta_{x_i} \right) \right) \right] \\ &\qquad (h(x,\cdot) \otimes h_\ell(x,\cdot))_\sigma\, (x_1, \ldots, x_{|\sigma|})\, \mu_t^{|\sigma|+1}(\mathrm{d}(x, x_1, \ldots, x_{|\sigma|}))\,. \end{aligned}$$

For the particular choice $\ell = k$ and $|\sigma| = k - 1$ we have

$$\int\limits_{\mathbb{X}^{|\sigma|+1}} \mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{|\sigma|} \delta_{x_i}\right)\right)\right] \big(h(x,\cdot) \otimes h_\ell(x,\cdot)\big)_\sigma (x_1,\dots,x_{|\sigma|})$$
$$\mu_t^{|\sigma|+1}(\mathrm{d}(x,x_1,\dots,x_{|\sigma|}))$$
$$= \frac{1}{k!}\int\limits_{\mathbb{X}^k} \mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i}\right)\right)\right] h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$
$$= \frac{1}{k!}\int\limits_{\mathbb{X}^k} \mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i}\right)\right) - \Delta f_B(S_A(\eta_t))\right] h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$
$$+ \frac{1}{k!}\int\limits_{\mathbb{X}^k} \mathbb{E}\left[\Delta f_B(S_A(\eta_t))\right]\, h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$
$$= \frac{1}{k!}\int\limits_{\mathbb{X}^k} \mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i}\right)\right) - \Delta f_B(S_A(\eta_t))\right] h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$
$$+ \frac{1}{k!}\mathbb{E}[\Delta f_B(S_A)]s_A\,.$$

Since there are $(k-1)!$ partitions $\sigma \in \Pi(k-1,k-1)$ with $|\sigma| = k-1$, we obtain that

$$\left|\mathbb{E}\left[s_A \Delta f_B(S_A) - \Delta f_B(S_A)\int\limits_{\mathbb{X}} D_x S_A(-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x)\right]\right|$$
$$\le k \sum_{\ell=1}^{k} \sum_{\sigma \in \Pi_{\neq}(k-1,\ell-1)} \int\limits_{\mathbb{X}^{|\sigma|+1}} \left|\mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{|\sigma|} \delta_{x_i}\right)\right)\right]\right|$$
$$(h(x,\cdot) \otimes h_\ell(x,\cdot))_\sigma\, (x_1,\dots,x_{|\sigma|})\, \mu_t^{|\sigma|+1}(\mathrm{d}(x,x_1,\dots,x_{|\sigma|}))$$
$$+ \int\limits_{\mathbb{X}^k} \left|\mathbb{E}\left[\Delta f_B\left(S_A\left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i}\right)\right) - \Delta f_B(S_A(\eta_t))\right]\right|$$
$$h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))\,.$$

Now (11) and the definition of ϱ_A imply that, for $\ell \in \{1,\dots,k\}$,

$$\sum_{\sigma \in \Pi_{\neq}(k-1,\ell-1)} \int_{\mathbb{X}^{|\sigma|+1}} \left| \mathbb{E}\left[\Delta f_B \left(S_A \left(\eta_t + \sum_{i=1}^{|\sigma|} \delta_{x_i} \right)\right)\right] \right|$$
$$\big(h(x,\cdot) \otimes h_\ell(x,\cdot)\big)_\sigma (x_1,\dots,x_{|\sigma|})\, \mu_t^{|\sigma|+1}(\mathrm{d}(x,x_1,\dots,x_{|\sigma|}))$$
$$\leq \varepsilon_{1,A}\, |\Pi_{\neq}(k-1,\ell-1)| \varrho_A\,.$$

Hence, the first summand above is bounded by

$$k\varepsilon_{1,A} \sum_{\ell=1}^{k} |\Pi_{\neq}(k-1,\ell-1)| \varrho_A\,.$$

By (11) we see that

$$\left| \mathbb{E}\left[\Delta f_B \left(S_A \left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i} \right)\right) - \Delta f_B(S_A(\eta_t)) \right] \right|$$
$$\leq 2\varepsilon_{1,A} \mathbb{E}\left[S_A \left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i} \right) - S_A(\eta_t) \right],$$

and the multivariate Mecke formula for Poisson point processes (see [18, formula (1.11)]) leads to

$$\mathbb{E}\left[\left(S_A \Big(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i} \Big)\right) - S_A(\eta_t) \right]$$
$$= \sum_{\emptyset \neq I \subset \{1,\dots,k-1\}} \frac{k!}{(k-|I|)!}\, \mathbb{E} \sum_{(y_1,\dots,y_{k-|I|}) \in \eta_{t,\neq}^{k-|I|}} h(x_I, y_1,\dots,y_{k-|I|})$$
$$= \sum_{\emptyset \neq I \subset \{1,\dots,k-1\}} \frac{k!}{(k-|I|)!}\, h_{|I|}(x_I)\,,$$

where for a subset $I = \{i_1,\dots,i_j\} \subset \{1,\dots,k-1\}$ we use the shorthand notation x_I for $(x_{i_1},\dots,x_{i_j})$. Hence,

$$\int_{\mathbb{X}^k} \left| \mathbb{E}\left[\Delta f_B \left(S_A \left(\eta_t + \sum_{i=1}^{k-1} \delta_{x_i} \right)\right) - \Delta f_B(S_A(\eta_t)) \right] \right| h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$

$$\leq 2\varepsilon_{1,A} \int_{\mathbb{X}^k} \sum_{\emptyset \neq I \subset \{1,\dots,k-1\}} h_{|I|}(x_I) \frac{k!}{(k-|I|)!} h(x_1,\dots,x_k)\, \mu_t^k(\mathrm{d}(x_1,\dots,x_k))$$

$$\leq 2\varepsilon_{1,A}\, k!(2^{k-1}-1)\varrho_A\,.$$

This implies that

$$\begin{aligned} &\left| \mathbb{E}\left[s_A \Delta f_B(S_A) - \Delta f_B(S_A) \int_{\mathbb{X}} D_x S_A(-D_x L^{-1} S_A)\, \mu_t(\mathrm{d}x) \right] \right| \\ &\quad \leq \varepsilon_{1,A}\left(k \sum_{\ell=1}^{k} |\Pi_{\neq}(k-1,\ell-1)| + 2k!(2^{k-1}-1) \right) \varrho_A =: C_1\, \varepsilon_{1,A}\varrho_A\,. \end{aligned} \tag{15}$$

For the second term in (14) we have

$$\begin{aligned} &2\int_{\mathbb{X}} \mathbb{E}\big[\max\{0, D_x S_A - 1\}(-D_x L^{-1} S_A)\big]\, \mu_t(\mathrm{d}x) \\ &\leq \frac{2}{k}\int_{\mathbb{X}} \mathbb{E}\big[\max\{0, D_x S_A - 1\} D_x S_A\big]\, \mu_t(\mathrm{d}x) \\ &\quad + 2\int_{\mathbb{X}} \mathbb{E}\big[\max\{0, D_x S_A - 1\}\, |D_x L^{-1} S_A + D_x S_A/k|\big]\, \mu_t(\mathrm{d}x) \\ &\leq \frac{2}{k}\int_{\mathbb{X}} \mathbb{E}\big[(D_x S_A - 1) D_x S_A\big]\, \mu_t(\mathrm{d}x) \\ &\quad + 2\int_{\mathbb{X}} \mathbb{E}\Big[\sqrt{D_x S_A(D_x S_A - 1)}\, |D_x L^{-1} S_A + D_x S_A/k|\Big]\, \mu_t(\mathrm{d}x) \\ &\leq 3\int_{\mathbb{X}} \mathbb{E}\big[(D_x S_A - 1) D_x S_A\big]\, \mu_t(\mathrm{d}x) + \int_{\mathbb{X}} \mathbb{E}\big[|D_x L^{-1} S_A + D_x S_A/k|^2\big]\, \mu_t(\mathrm{d}x)\,. \end{aligned}$$

It follows from Lemmas 2 and 1 that

$$\begin{aligned} &\int_{\mathbb{X}} \mathbb{E}\big[(D_x S_A - 1) D_x S_A\big]\, \mu_t(\mathrm{d}x) \\ &\quad = \int_{\mathbb{X}} k^2 \sum_{\sigma \in \Pi(k-1,k-1)} \int_{\mathbb{X}^{|\sigma|}} (h(x,\cdot) \otimes h(x,\cdot))_\sigma\, \mathrm{d}\mu_t^{|\sigma|}\, \mu_t(\mathrm{d}x) - k \int_{\mathbb{X}^k} h\, \mathrm{d}\mu_t^k\,. \end{aligned}$$

Since there are $(k-1)!$ partitions with $|\sigma| = k-1$ and for each of them

$$(h(x,\cdot) \otimes h(x,\cdot))_\sigma(x_1,\dots,x_{|\sigma|}) = \frac{1}{k!} h(x,x_1,\dots,x_{|\sigma|}),$$

this leads to

$$\begin{aligned}
&\int_{\mathbb{X}} \mathbb{E}\big[(D_xS_A - 1)D_xS_A\big]\,\mu_t(\mathrm{d}x) \\
&\quad = k^2 \sum_{\sigma \in \Pi_{\neq}(k-1,k-1)} \int_{\mathbb{X}} \int_{\mathbb{X}^{|\sigma|}} (h(x,\cdot)\otimes h(x,\cdot))_\sigma \,\mathrm{d}\mu_t^{|\sigma|}\,\mu_t(\mathrm{d}x) \\
&\quad \le k^2 |\Pi_{\neq}(k-1,k-1)| \varrho_A .
\end{aligned}$$

Lemmas 2 and 3 imply that

$$D_xL^{-1}S_A + D_xS_A/k = -\sum_{\ell=1}^{k-1} \sum_{(x_1,\dots,x_{\ell-1}) \in \eta_{t,\neq}^{\ell-1}} h_\ell(x,x_1,\dots,x_{\ell-1})$$

so that Lemma 1 yields

$$\begin{aligned}
&\int_{\mathbb{X}} \mathbb{E}\big[|D_xL^{-1}S_A + D_xS_A/k|^2\big]\,\mu_t(\mathrm{d}x) \\
&\quad = \int_{\mathbb{X}} \sum_{i,j=1}^{k-1} \sum_{\sigma \in \Pi(i-1,j-1)} \int_{\mathbb{X}^{|\sigma|}} (h_i(x,\cdot)\otimes h_j(x,\cdot))_\sigma \,\mathrm{d}\mu_t^{|\sigma|}\,\mu_t(\mathrm{d}x) \\
&\quad \le \sum_{i,j=1}^{k-1} |\Pi(i-1,j-1)| \varrho_A .
\end{aligned}$$

From the previous estimates, we can deduce that

$$\begin{aligned}
&2\varepsilon_{1,A} \int_{\mathbb{X}} \mathbb{E}\big[\max\{0, D_xS_A - 1\}(-D_xL^{-1}S_A)\big]\,\mu_t(\mathrm{d}x) \\
&\quad \le \varepsilon_1 \Big(3k^2|\Pi_{\neq}(k-1,k-1)| + \sum_{i,j=1}^{k-1} |\Pi(i-1,j-1)|\Big)\varrho_A =: C_2\,\varepsilon_{1,A}\varrho_A .
\end{aligned} \tag{16}$$

Combining (14) with (15) and (16) shows that

$$d_{\mathrm{TV}}(S_A, Y) \le |s_A - s| + (C_1 + C_2)\varepsilon_{1,A}\varrho_A ,$$

which concludes the proof. □

Remark 12 As already discussed in the introduction, the proof of Proposition 1—the main tool for the proof of Theorem 1—is different from that given in [29]. One of the differences is Lemma 3, which provides an explicit representation for $-D_x L^{-1} S_A$ based on Mehler's formula. We took considerable advantage of this in the proof of Proposition 1 and remark that the proof of the corresponding result in [29] uses the chaotic decomposition of U-statistics and the product formula for multiple stochastic integrals (see [18]). Another difference is that our proof here does not make use of the estimates established by the Malliavin–Chen–Stein method in [21]. Instead, we directly manipulate the Chen–Stein equation for Poisson approximation and this way improve the rate of convergence compared to [29] . A different method to show Theorems 1 and 2 is the content of the recent paper [7].

4.3 Poisson Approximation of Classical U-Statistics

In this section we consider U-statistics based on a binomial point process ζ_n defined as

$$S_A = \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\zeta^k_{n,\neq}} \mathbb{1}\{f(x_1,\dots,x_k) \in A\} ,$$

where f is as in Sect. 2 and $A \subset \mathbb{R}$ is bounded and measurable. Recall that in the context of a binomial point process ζ_n we assume that $\mu(\mathbb{X}) = 1$. Denote as in the previous section by $s_A := \mathbb{E}[S_A]$ the expectation of S_A. Notice that

$$s_A = (n)_k \int\limits_{\mathbb{X}^k} h(x_1,\dots,x_k)\, \mu^k(\mathrm{d}(x_1,\dots,x_k)) \tag{17}$$

with $h(x_1 \dots, x_k) = (k!)^{-1}\mathbb{1}\{f(x_1,\dots,x_k) \in A\}$.

Proposition 2 *Let S_A be as above and let Y be a Poisson distributed random variable with mean $s > 0$ and define*

$$\varrho_A := \max_{1\le\ell\le k-1} (n)_{2k-\ell} \int\limits_{\mathbb{X}^\ell} \Big(\int\limits_{\mathbb{X}^{k-\ell}} h(x_1,\dots,x_k)\, \mu^{k-\ell}(\mathrm{d}(x_{\ell+1},\dots,x_k)) \Big)^2 \mu^\ell\big(\mathrm{d}(x_1,\dots,x_\ell)\big) .$$

Then there is a constant $C \geq 1$ *only depending on* k *such that*

$$d_{\mathrm{TV}}(S_A, Y) \leq |s_A - s| + C \min\left\{1, \frac{1}{s_A}\right\}\left(\varrho_A + \frac{s_A^2}{n}\right).$$

Proof By the same arguments as at the beginning of the proof of Proposition 1 it is sufficient to assume that $s = s_A$ in what follows. To simplify the presentation we put $N := \{I \subset \{1, \ldots, n\} : |I| = k\}$ and rewrite S_A as

$$S_A = \sum_{I \in N} \mathbb{1}\{f(X_I) \in A\},$$

where $X_1, \ldots, X_n$ are i.i.d. random elements in $\mathbb{X}$ with distribution μ and where X_I is shorthand for $(X_{i_1}, \ldots, X_{i_k})$ if $I = \{i_1, \ldots, i_k\}$. In this situation it follows from Theorem 2 in [1] that

$$d_{\mathrm{TV}}(S, Y) \leq \min\left\{1, \frac{1}{s_A}\right\} \sum_{I \in N} \left(\mathbb{P}(f(X_I) \in A)^2 + \sum_{r=1}^{k-1} \sum_{\substack{J \in N \\ |I \cap J| = r}} \mathbb{P}(f(X_I) \in A)\mathbb{P}(f(X_J) \in A)\right)$$
$$+ \min\left\{1, \frac{1}{s_A}\right\} \sum_{I \in N} \sum_{r=1}^{k-1} \sum_{\substack{J \in N \\ |I \cap J| = r}} \mathbb{P}(f(X_I) \in A, f(X_J) \in A).$$

Since $s_A = \mathbb{E}[S_A] = \frac{(n)_k}{k!}\mathbb{P}(f(X_1, \ldots, X_k) \in A)$, we have that

$$\begin{aligned}
&\sum_{I \in N} \left(\mathbb{P}(f(X_I) \in A)^2 + \sum_{r=1}^{k-1} \sum_{\substack{J \in N \\ |I \cap J| = r}} \mathbb{P}(f(X_I) \in A)\mathbb{P}(f(X_J) \in A)\right) \\
&= \frac{(n)_k}{k!} \left(\left(\frac{k!}{(n)_k} s_A\right)^2 + \sum_{r=1}^{k-1} \sum_{\substack{J \in N \\ |I \cap J| = r}} \left(\frac{k!}{(n)_k} s_A\right)^2\right) \\
&= \frac{k!}{(n)_k} s_A^2 \left(1 + \sum_{r=1}^{k-1} \binom{k}{r}\binom{n-k}{k-r}\right) \\
&\leq \frac{k!}{(n)_k} s_A^2\, 2^k (n-1)_{k-1} \\
&\leq \frac{2^k k! s_A^2}{n}.
\end{aligned}$$

For the second term we find that

$$\sum_{I\in N}\sum_{r=1}^{k-1}\sum_{\substack{J\in N\\ |I\cap J|=r}} \mathbb{P}(f(X_I)\in A, f(X_J)\in A)$$
$$= \frac{(n)_k}{k!}\sum_{r=1}^{k-1}\binom{k}{r}\binom{n-k}{k-r}\mathbb{P}(f(X_1,\dots,X_k)\in A, f(X_1,\dots,X_r,X_{k+1},\dots,X_{2k-r})\in A)$$
$$\leq \frac{(n)_k}{k!}\sum_{r=1}^{k-1}\binom{k}{r}\binom{n-k}{k-r}\frac{(k!)^2}{(n)_{2k-r}}\varrho_A$$
$$\leq 2^k k!\,\varrho_A\,.$$

Putting $C := 2^k k!$ proves the claim. □

4.4 Proofs of Theorems 1 and 2 and Corollaries 1 and 2

Proof (of Theorem 1) We define the set classes

$$\mathbf{I} = \{I = (a,b] : a,b\in\mathbb{R}, a<b\}$$

and

$$\mathbf{V} = \{V = \bigcup_{i=1}^{n} I_i : n\in\mathbb{N}, I_i\in\mathbf{I}, i=1,\dots,n\}.$$

From [15, Theorem 16.29] it follows that $(t^\gamma\xi_t)_{t>0}$ converges in distribution to a Poisson point process ξ with intensity measure ν if

$$\lim_{t\to\infty}\mathbb{P}(\xi_t(t^{-\gamma}V)=0) = \mathbb{P}(\xi(V)=0) = \exp(-\nu(V)),\quad V\in\mathbf{V}, \tag{18}$$

and

$$\lim_{t\to\infty}\mathbb{P}(\xi_t(t^{-\gamma}I)>1) = \mathbb{P}(\xi(I)>1) = 1-(1+\nu(I))\exp(-\nu(I)),\quad I\in\mathbf{I}. \tag{19}$$

Note that $\mathbf{I}\subset\mathbf{V}$ and that every set $V\in\mathbf{V}$ can be represented in the form

$$V = \bigcup_{i=1}^{n}(a_i,b_i]\quad\text{with}\quad a_1<b_1<\dots<a_n<b_n\quad\text{and}\quad n\in\mathbb{N}\,.$$

For $V \in \mathbf{V}$ we define the Poisson U-statistic

$$S_{V,t} = \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\eta^k_{t,\neq}} \mathbb{1}\{f(x_1,\dots,x_k) \in t^{-\gamma} V\}\,,$$

which has expectation

$$\begin{aligned}\mathbb{E}[S_{V,t}] &= \frac{1}{k!}\mathbb{E} \sum_{(x_1,\dots,x_k)\in\eta^k_{t,\neq}} \mathbb{1}\{f(x_1,\dots,x_k) \in t^{-\gamma} V\}\\ &= \sum_{i=1}^{n} \frac{1}{k!}\mathbb{E} \sum_{(x_1,\dots,x_k)\in\eta^k_{t,\neq}} \mathbb{1}\{f(x_1,\dots,x_k) \in t^{-\gamma} (a_i,b_i]\} = \sum_{i=1}^{n} \alpha_t(a_i,b_i).\end{aligned}$$

Since $\xi(V)$ is Poisson distributed with mean $\nu(V) = \sum_{i=1}^{n} \nu((a_i,b_i])$, it follows from Proposition 1 that

$$d_{\mathrm{TV}}(S_{V,t},\xi(V)) \le \left| \sum_{i=1}^{n} \alpha_t(a_i,b_i) - \sum_{i=1}^{n} \nu((a_i,b_i]) \right| + C\, r_t(y_{\max})$$

with $y_{\max} := \max\{|a_1|,|b_n|\}$ and $C \ge 1$. Now, assumptions (2) and (3) yield that

$$\lim_{t\to\infty} d_{\mathrm{TV}}(S_{V,t},\xi(V)) = 0\,.$$

Consequently, the conditions (18) and (19) are satisfied so that $(t^{\gamma}\xi_t)_{t>0}$ converges in distribution to ξ. Choosing $V = (0,y]$ and using the fact that $t^{\gamma} M_t^{(m)} > y$ is equivalent to $S_{(0,y],t} < m$ lead to the first inequality in Theorem 1. The second one follows analogously from $V = (-y,0]$ and by using the equivalence of $t^{\gamma} M_t^{(-m)} \ge y$ and $S_{(-y,0],t} < m$. □

Proof (of Corollary 1) Theorem 1 with ν defined as in (5) yields the assertions of Corollary 1. □

Proof (of Theorem 2 and Corollary 2) Since the proofs are similar to those of Theorem 1 and Corollary 1, we skip the details. □

References

1. Barbour, A.D., Eagleson, G.K.: Poisson convergence for dissociated statistics. J. R. Stat. Soc. B **46**, 397–402 (1984)
2. Barbour, A.D., Holst, L., Janson, S.: Poisson Approximation. Oxford University Press, Oxford (1992)

3. Bourguin, S., Peccati, G.: The Malliavin-Stein method on the Poisson space. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 185–228. Springer, Cham (2016)
4. Broadbent, S.: Simulating the ley-hunter. J. R. Stat. Soc. Ser. A (Gen) **143**, 109–140 (1980)
5. Calka, P., Chenavier, N.: Extreme values for characteristic radii of a Poisson-Voronoi tessellation. Extremes **17**, 359–385 (2014)
6. Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J.: Stochastic Geometry and its Applications, 3rd edn. Wiley, Chichester (2013)
7. Decreusefond, L., Schulte, M., Thäle, C.: Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. Accepted for publication in Ann. Probab. (2015)
8. Hörrmann, J., Hug, D.: On the volume of the zero cell of a class of isotropic Poisson hyperplane tessellations. Adv. Appl. Probab. **46**, 622–642 (2014)
9. Hörrmann, J., Hug, D., Reitzner, M., Thäle, C.: Poisson polyhedra in high dimensions. Adv. Math. **281**, 1–39 (2015)
10. Hug, D., Last, G., Weil, W.: Distance measurements on processes of flats. Adv. Appl. Probab. **35**, 70–95 (2003)
11. Hug, D., Reitzner, M.: Introduction to stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi and Springer Series, vol. 6, pp. 145–184. Springer/Bocconi University Press, Milan (2016)
12. Hug, D., Schneider, R.: Asymptotic shapes of large cells in random tessellations. Geom. Funct. Anal. **17**, 156–191 (2007)
13. Hug, D., Schneider, R., Schuster, R.: Integral geometry of tensor valuations. Adv. Appl. Math. **41**, 482–509 (2008)
14. Hug, D., Thäle, C., Weil, W.: Intersection and proximity for processes of flats. J. Math. Anal. Appl. **426**, 1–42 (2015)
15. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2001)
16. Kingman, J.F.C.: Poisson Processes. Oxford University Press, Oxford (1993)
17. Lachièze-Rey, R., Reitzner, M.: U-statistics in stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic geometry. Bocconi & Springer Series, vol. 7, pp. 229–253. Springer, Cham (2016)
18. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Milan (2016)
19. Last, G., Penrose, M., Schulte, M., Thäle, C.: Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**, 348–364 (2014)
20. Last, G., Peccati, G., Schulte, M.: Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequality and stabilization. Accepted for Publication in Probab. Theory Relat. Fields (2015)
21. Peccati, G.: The Chen-Stein method for Poisson functionals. arXiv: 1112.5051 (2011)
22. Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
23. Privault, N.: Combinatorics of Poisson stochastic integrals with random integrands. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 37–80. Springer, Cham (2016)
24. Reitzner, M., Schulte, M.: Central limit theorems for U-statistics of Poisson point processes. Ann. Probab. **41**, 3879–3909 (2013)
25. Reitzner, M., Schulte, M., Thäle, C.: Limit theory for the Gilbert graph. arXiv: 1312.4861 (2013)
26. Schneider, R.: A duality for Poisson flats. Adv. Appl. Probab. **31**, 63–68 (1999)
27. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Springer, Berlin (2008)

28. Schulte, M.: Normal approximation of Poisson functionals in Kolmogorov distance. Accepted for publication in J. Theoret. Probab. **29**, 96–117 (2016)
29. Schulte, M., Thäle, C.: The scaling limit of Poisson-driven order statistics with applications in geometric probability. Stoch. Proc. Appl. **122**, 4096–4120 (2012)
30. Schulte, M., Thäle, C.: Distances between Poisson *k*-flats. Methodol. Comput. Appl. Probab. **16**, 311–329 (2014)
31. Silverman, B., Brown, T.: Short distances, flat triangles and Poisson limits. J. Appl. Probab. **15**, 815–825 (1978)
32. Surgailis, D.: On multiple Poisson stochastic integrals and associated Markov semigroups. Probab. Math. Statist. **3**, 217–239 (1984)

U-Statistics on the Spherical Poisson Space

Solesne Bourguin, Claudio Durastanti, Domenico Marinucci, and Giovanni Peccati

Abstract We review a recent stream of research on normal approximations for linear functionals and more general U-statistics of wavelets/needlets coefficients evaluated on a homogeneous spherical Poisson field. We show how, by exploiting results from Peccati and Zheng (Electron J Probab 15(48):1487–1527, 2010) based on Malliavin calculus and Stein's method, it is possible to assess the rate of convergence to Gaussianity for a triangular array of statistics with growing dimensions. These results can be exploited in a number of statistical applications, such as spherical density estimations, searching for point sources, estimation of variance, and the spherical two-sample problem.

1 Introduction

1.1 Overview

The purpose of this chapter is to review some recent developments concerning computations of Berry–Esseen bounds in two classical statistical frameworks, that is: linear functionals and U-statistics associated with wavelets coefficients evaluated on

S. Bourguin (✉)
Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA 02215, USA
e-mail: solesne.bourguin@gmail.com

C. Durastanti
Faculty of Mathematics, Ruhr University Bochum, 44780 Bochum, Germany
e-mail: claudio.durastanti@gmail.com

D. Marinucci
Department of Mathematics, University of Rome Tor Vergata, Via della Ricerca Scientifica, 1, 00133 Rome, Italy
e-mail: marinucc@mat.uniroma2.it

G. Peccati
Unité de Recherche en Mathématiques, Université du Luxembourg, 6 rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Luxembourg
e-mail: giovanni.peccati@gmail.com

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_9

spherical Poisson fields. These statistics are motivated by some standard problems in statistical inference, such as: (1) testing for the functional form of an unknown density function $f(\cdot)$; (2) estimation of the variance; (3) comparison between two unknown density functions $f(\cdot)$ and $g(\cdot)$ (the so-called *two sample problem*). While the former are indeed among the most common (and basic) problems in statistical inference, we shall investigate their solution under circumstances which are not standard, for a number of reasons. Firstly, we shall consider the case of directional data, e.g., we shall assume that the domain of the density functions $f(\cdot)$ and $g(\cdot)$ is a compact manifold, which we shall take for definiteness (and for practical relevance) to be the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Note that all the arguments we review can easily be extended to $\mathbb{S}^d,\ d \geq 2$ (or, with more work, to other compact manifolds), but we shall not pursue these generalizations here for brevity and simplicity. Secondly, as opposed to most of the existing statistical procedures, we shall focus on "local" tests, e.g., we shall allow for the possibility that not all the manifold (the sphere) is observable, but possibly only strict subsets. Finally, and most importantly, we shall consider classes of "high-frequency" tests, where the number of procedures to be implemented is itself a function of the number of observations available, in a manner to be made rigorous later. For all these objectives, but especially for the latter, the Malliavin–Stein techniques that we shall adopt and describe turn out to be of the greatest practical importance, as they allow, for instance, to determine how many joint procedures can be run while maintaining an acceptable Gaussian approximation for the resulting statistics.

Malliavin–Stein techniques for Poisson processes are discussed in detail in [4] of this volume. Our specific purpose, in view of the previous considerations, is to apply and extend the now well-known results of [23, 24] (see also [21]) in order to deduce bounds that are well adapted to the applications we mentioned, e.g., those where the dimension of a given statistic increases with the number of observations. Our principal motivation originates from the implementation of wavelet systems on the sphere in the framework of statistical analysis for cosmic rays data. As noted in [7], in these circumstances, when more and more data become available, a higher number of wavelet coefficients is evaluated, as it is customarily the case when considering, for instance, thresholding nonparametric estimators. We shall hence be concerned with sequences of Poisson fields, whose intensity grows monotonically; it is then possible to exploit local normal approximations, where the rate of convergence to the asymptotic Gaussian distribution is related to the scale parameter of the corresponding wavelet transform in a natural and intuitive way. Moreover, in a multivariate setting the wavelets localization properties are exploited to establish bounds that grow linearly with the number of functionals considered; it is then possible to provide explicit recipes, for instance, for the number of joint testing procedures that can be simultaneously entertained ensuring that the Gaussian approximation may still be shown to hold, in a suitable sense. These arguments are presented for both linear and U-statistics; proof for the latter (which are provided in [5]) are considerably more complicated from the technical point of view, but remarkably the main qualitative conclusions go through unaltered.

In Sect. 2 we review some background material and notation on Poisson processes, Malliavin–Stein approximations, and spherical wavelets/needlets. In Sect. 3 we review results on linear functionals, while in Sect. 4 we focus on nonlinear U-statistics. Throughout the chapter, we discuss motivating applications and highlight avenues for further research; for the proofs, at most main ideas and quick sketches are provided, but we provide full references to the existing literature whenever needed.

2 Background

2.1 Poisson Random Measures and Malliavin–Stein Bounds

Throughout this chapter, we take for granted that we are working on a suitable probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We work within the general framework outlined in [15], namely: $(\mathbb{X}, \mathscr{X}, \mu)$ is a σ-finite measure space, and η is a *proper* Poisson random measure on $(\mathbb{X}, \mathscr{X})$ with intensity measure μ. For $p \geq 1$, we denote by L^p_η the class of those random variables F such that $\mathbb{E}|F|^p < \infty$ and $F = f(\eta)$, $\mathbb{P}$-a.s., where f is a representative of F. Recall that f is a measurable function from $\mathbf{N}_\sigma$ (the class of all measures on $(\mathbb{X}, \mathscr{X})$ taking values in $\mathbb{Z}_+ \cup \{+\infty\}$)—see [15, Sect. 1.2] for more details. We will, however, specialize this general framework for the purpose of the present chapter: we take $\mathbb{X} = \mathbb{R}_+ \times \mathbb{S}^2$, with $\mathscr{X} = \mathcal{B}(\mathbb{X})$, the class of Borel subsets of $\mathbb{X}$. The control measure μ of η is taken to have the form $\mu = \rho \times \nu$, where ρ is some measure on $\mathbb{R}_+$ and ν is a probability measure on $\mathbb{S}^2$ of the form $\nu(dx) = f(x)dx$, where f is a density on the sphere. We shall assume that $\rho(\{0\}) = 0$ and that the mapping $\rho \mapsto \rho([0,t])$ is strictly increasing and diverging to infinity as $t \to \infty$. We also adopt the notation

$$R_t := \rho([0,t]), \quad t \geq 0,$$

that is, $t \mapsto R_t$ is the distribution function of ρ. We shall also assume $f(x)$ to be bounded and bounded away from zero, e.g.,

$$\zeta_1 \leq f(x) \leq \zeta_2 \text{ , some } \zeta_1, \zeta_2 > 0, \quad \text{for all } x \in \mathbb{S}^2. \tag{1}$$

To simplify the discussion, we shall take $\rho(ds) = R \cdot \ell(ds)$, where ℓ is the Lebesgue measure and $R > 0$, in such a way that $R_t = R \cdot t$.

We now state two Stein bounds for random variables living in the *first chaos* associated with the Poisson measure η. These statements are specializations of the general results presented in [4]. In what follows, we shall use the symbols $\eta(f)$ and $\hat{\eta}(f)$, respectively, to denote the Wiener–Itô integrals of f with respect to η and with respect to the *compensated Poisson measure*

$$\hat{\eta}(A) = \eta(A) - \mu(A), \quad A \in \mathcal{B}(\mathbb{X}),$$

where we use the convention $\eta(A) - \mu(A) = \infty$ whenever $\mu(A) = \infty$ (recall that μ is σ-finite). We shall consider Wiener–Itô integrals of functions f having the form $f = 1_{[0,t]} \times h$, where $t > 0$ and $h \in L^2(\mathbb{S}^2, \nu) \cap L^1(\mathbb{S}^2, \nu)$. For a function f of this type we simply write

$$\eta(f) = \eta(1_{[0,t]} \times h) := \eta_t(h), \quad \text{and} \quad \hat{\eta}(f) = \hat{\eta}(1_{[0,t]} \times h) := \hat{\eta}_t(h).$$

Observe that, for fixed t, the mapping $A \mapsto \hat{\eta}_t(1_A)$ defines a Poisson measure on $(\mathbb{S}^2, \mathscr{B}(\mathbb{S}^2))$, with control $R_t \cdot \nu := \nu_t$.

Theorem 1 *Let the notation and assumptions of this section prevail.*

1. *Let $h \in L^2(\mathbb{S}^2, \nu) := L^2(\nu)$, let $N \sim \mathcal{N}(0, 1)$ and fix $t > 0$. Then, the following bound holds:*

$$d_W(\hat{\eta}_t(h), N) \leq \left|1 - \|h\|^2_{L^2(\mathbb{S}^2, \nu_t)}\right| + \int_{\mathbb{S}^2} |h(z)|^3 \nu_t(\mathrm{d}z),$$

where d_W denotes the Wasserstein distance as defined in [4, Sect. 6.2.1].

2. *For a fixed integer $d \geq 1$, let $N \sim \mathcal{N}_d(0, \Sigma)$, with Σ a positive definite covariance matrix and let*

$$F_t = (F_{t,1}, \ldots, F_{t,d}) = (\hat{\eta}_t(h_{t,1}), \ldots \hat{\eta}_t(h_{t,d}))$$

be a collection of d-dimensional random vectors such that $h_{t,a} \in L^2(\nu)$. If we call Γ_t the covariance matrix of F_t, then:

$$\begin{aligned} d_2(F_t, N) &\leq \left\|\Sigma^{-1}\right\|_{\mathrm{op}} \|\Sigma\|_{\mathrm{op}}^{\frac{1}{2}} \|\Sigma - \Gamma_t\|_{H.S.} \\ &\quad + \frac{\sqrt{2\pi}}{8} \left\|\Sigma^{-1}\right\|_{\mathrm{op}}^{\frac{3}{2}} \|\Sigma\|_{\mathrm{op}} \sum_{i,j,k=1}^{d} \int_{\mathbb{S}^2} |h_{t,i}(x)| \left|h_{t,j}(x)\right| |h_{t,k}(x)| \, \nu_t(\mathrm{d}x), \\ &\leq \left\|\Sigma^{-1}\right\|_{\mathrm{op}} \|\Sigma\|_{\mathrm{op}}^{\frac{1}{2}} \|\Sigma - \Gamma_t\|_{H.S.} \\ &\quad + \frac{d^2\sqrt{2\pi}}{8} \left\|\Sigma^{-1}\right\|_{\mathrm{op}}^{\frac{3}{2}} \|\Sigma\|_{\mathrm{op}} \sum_{i=1}^{d} \int_{\mathbb{S}^2} |h_{t,i}(x)|^3 \, \nu_t(\mathrm{d}x), \end{aligned}$$

where $\|\cdot\|_{\mathrm{op}}$ and $\|\cdot\|_{H.S.}$ stand, respectively, for the operator and Hilbert–Schmidt norms and where the distance d_2 is the one defined in [4, Sect. 6.2.1].

Remark 1 As pointed out earlier, Theorem 1 is a specialized restatement of [4, Theorems 7 and 14].

It should be noted that the convergence in law implied by Theorem 1 is in fact *stable*, as defined, e.g., in the classic reference [10, Chap. 4].

2.2 Needlets

In this subsection, we review very quickly some simple and basic facts on the construction of spherical wavelets. Recall first that the set of spherical harmonics

$$\{Y_{lm} : l \geq 0, m = -l, \ldots, l\}$$

provides an orthonormal basis for the space of square-integrable functions on the unit sphere $L^2\left(\mathbb{S}^2, \mathrm{d}x\right) := L^2\left(\mathbb{S}^2\right)$, where d$x$ stands for the Lebesgue measure on $\mathbb{S}^2$ (see for instance [1, 11, 17, 32]). Spherical harmonics are eigenfunctions of the spherical Laplacian $\Delta_{\mathbb{S}^2}$ corresponding to eigenvalues $-l\,(l+1)$, e.g. $\Delta_{\mathbb{S}^2} Y_{lm} = -l(l+1)Y_{lm}$. For every $l \geq 0$, we define as usual $\mathcal{K}_l$ as the linear space given by the restriction to the sphere of the polynomials with degree at most l, e.g.,

$$\mathcal{K}_l = \bigoplus_{k=0}^{l} \operatorname{span}\left\{Y_{km} : m = -k, \ldots, k\right\},$$

where the direct sum is in the sense of $L^2\left(\mathbb{S}^2\right)$. It is well known that for every integer $l = 1, 2, \ldots$, there exists a finite set of *cubature points* $\mathcal{Q}_l \subset \mathbb{S}^2$, as well as a collection of *weights* $\{\lambda_\eta\}$, indexed by the elements of $\mathcal{Q}_l$, such that

$$\forall f \in \mathcal{K}_l, \quad \int_{\mathbb{S}^2} f(x)\mathrm{d}x = \sum_{\eta \in \mathcal{Q}_l} \lambda_\eta f(\eta).$$

Now fix $B > 1$, and write $[x]$ to indicate the integer part of a given real x. In what follows, we shall denote by $\mathcal{X}_j = \{\xi_{jk}\}$ and $\{\lambda_{jk}\}$, respectively, the set $\mathcal{Q}_{[2B^j+1]}$ and the associated class of weights. We also write $K_j = card\{\mathcal{X}_j\}$. As proved in [19, 20] (see also e.g. [3, 25, 26] and [17, Chap. 10]), cubature points and weights can be chosen to satisfy

$$\lambda_{jk} \approx B^{-2j}\,,\ K_j \approx B^{2j},$$

where by $a \approx b$, we mean that there exists $c_1, c_2 > 0$ such that $c_1 a \leq b \leq c_2 a$.

Fix $B > 1$ as before, as well as a real-valued mapping b on $(0, \infty)$. We assume that b verifies the following properties: (1) the function $b\,(\cdot)$ has compact support in $\left[B^{-1}, B\right]$ (in such a way that the mapping $l \mapsto b\left(\frac{l}{B^j}\right)$ has compact support in $l \in \left[B^{j-1}, B^{j+1}\right]$) (2) for every $\xi \geq 1$, $\sum_{j=0}^{\infty} b^2(\xi B^{-j}) = 1$ (*partition of unit property*), and (3) $b(\cdot) \in C^{\infty}(0,\infty)$. The collection of spherical needlets $\{\psi_{jk}\}$, associated with B and $b(\cdot)$, are then defined as a weighted convolution of the projection operator $L_l(\langle x, y\rangle) = \sum_{m=-l}^{l} \overline{Y}_{lm}(x) Y_{lm}(y)$, that is

$$\psi_{jk}(x) := \sqrt{\lambda_{jk}} \sum_{l} b\left(\frac{l}{B^j}\right) L_l\left(\langle x, \xi_{jk}\rangle\right). \tag{2}$$

The properties of b entail the following quasi-exponential *localization property* (see [19] or [17, Sect. 13.3]): for any $\tau = 1, 2, \ldots$, there exists $\kappa_\tau > 0$ such that for any $x \in \mathbb{S}^2$,

$$\left|\psi_{jk}(x)\right| \leq \frac{\kappa_\tau B^j}{\left(1 + B^j \arccos\left(\langle x, \xi_{jk}\rangle\right)\right)^{\tau}}.$$

Note that $d(x, y) := \arccos(\langle x, y\rangle)$ is indeed the spherical distance. From localization, the following bound can be established on the $L^p\left(\mathbb{S}^2\right)$ norms: for all $1 \leq p \leq +\infty$, there exist two positive constants q_p and q_p' such that

$$q_p B^{j\left(1-\frac{2}{p}\right)} \leq \left\|\psi_{jk}\right\|_{L^p\left(\mathbb{S}^2\right)} \leq q_p' B^{j\left(1-\frac{2}{p}\right)}. \tag{3}$$

In the sequel, we shall write

$$\beta_{jk} := \left\langle f, \psi_{jk}\right\rangle_{L^2\left(\mathbb{S}^2\right)} = \int_{\mathbb{S}^2} f(x)\, \psi_{jk}(x)\, \mathrm{d}x$$

for the so-called *needlet coefficient* of index j, k. Our aim in this chapter is to review asymptotic results for linear and nonlinear functionals of needlet coefficients given by

$$\hat{\beta}_{jk} = \sum_{i=1}^{N_t} \psi_{jk}(X_i), \quad j = 1, 2, \ldots, \; k = 1, \ldots, K_j, \tag{4}$$

where the function ψ_{jk} is defined according to (2), and where $\{X_1, \ldots, X_{N_t}\}$ represent the points in the support of some random Poisson measure.

2.3 Needlet Coefficients as Wiener–Itô Integrals

The first step to be able to exploit the existing results on Malliavin–Stein approximations is to express our needlet coefficients as Wiener–Itô integrals with respect to a Poisson random measure. For every $j \geq 1$ and every $k = 1, \ldots, N_j$, consider the function ψ_{jk} defined in (2), and observe that ψ_{jk} is trivially an element of $L^3(\mathbb{S}^2, \nu) \cap L^2(\mathbb{S}^2, \nu) \cap L^1(\mathbb{S}^2, \nu)$. We write

$$\sigma_{jk}^2 := \int_{\mathbb{S}^2} \psi_{jk}^2(x) f(x) \mathrm{d}x, \quad b_{jk} := \int_{\mathbb{S}^2} \psi_{jk}(x) f(x) \mathrm{d}x.$$

Observe that, if $f(x) = \frac{1}{4\pi}$ (that is, the uniform density on the sphere), then $b_{jk} = 0$ for every $j > 1$. On the other hand, under (3),

$$\zeta_1 \left\| \psi_{jk} \right\|_{L^2(\mathbb{S}^2)}^2 \leq \sigma_{jk}^2 \leq \zeta_2 \left\| \psi_{jk} \right\|_{L^2(\mathbb{S}^2)}^2 .$$

Note that (see (3)) the L^2-norm of $\{\psi_{jk}\}$ is uniformly bounded above and below, and therefore the same is true for $\left\{\sigma_{jk}^2\right\}$. For every $t > 0$ and every j, k, introduce the kernel

$$h_{jk}^{(R_t)}(x) = \frac{\psi_{jk}(x)}{\sqrt{R_t}\sigma_{jk}}, \quad x \in \mathbb{S}^2,$$

and write

$$\tilde{\beta}_{jk}^{(R_t)} := \hat{\eta}_t\left(h_{jk}^{(R_t)}\right) = \int_{\mathbb{S}^2} h_{jk}^{(R_t)}(x)\, \hat{\eta}_t(\mathrm{d}x) = \sum_{x \in \mathrm{supp}(\eta_t)} h_{jk}^{(R_t)}(x) - R_t \cdot \int_{\mathbb{S}^2} h_{jk}^{(R_t)}(x) \nu(\mathrm{d}x), \tag{5}$$

The random variable $\tilde{\beta}_{jk}^{(R_t)}$ can always be represented in the form

$$\tilde{\beta}_{jk}^{(R_t)} = \frac{\left(\sum_{i=1}^{\eta_t(\mathbb{S}^2)} \psi_{jk}(X_i) - R_t b_{jk}\right)}{\sqrt{R_t}\sigma_{jk}},$$

where $\{X_i : i \geq 1\}$ is a sequence of i.i.d. random variables with common distribution ν, and independent of the Poisson random variable $\hat{\eta}_t(\mathbb{S}^2)$. Moreover, the following relations are immediately checked:

$$\mathbb{E}\left[\tilde{\beta}_{jk}^{(R_t)}\right] = 0, \quad \mathbb{E}\left[\left(\tilde{\beta}_{jk}^{(R_t)}\right)^2\right] = 1.$$

This representation will provide the main working tool for deducing Malliavin–Stein bounds for linear and U-statistics on spherical Poisson fields. We start first with the linear case, which is discussed in [7].

3 Linear Statistics

The joint distribution of the coefficients $\{\hat{\beta}_{jk}\}$ (as defined in (4)) is required in statistical procedures devised for the research of so-called *point sources*, again for instance in an astrophysical context (see, e.g., [31]). The astrophysical motivation can be summarized as follows: we assume that under the null hypothesis, we are observing a background of cosmic rays governed by a Poisson random measure on the sphere $\mathbb{S}^2$, with the form of the measure $\eta_t(\cdot)$ defined earlier. In particular, η_t is built from a measure η verifying the stated regularity conditions, and the intensity of $\mu_t(\mathrm{d}x) = \mathbb{E}[\eta_t(\mathrm{d}x)]$ is given by the absolutely continuous measure $R_t \cdot f(x)\mathrm{d}x$, where $R_t > 0$ and f is a density on the sphere. This situation corresponds, for instance, to the presence of a diffuse background of cosmological emissions. Under the alternative hypothesis, the background of cosmic rays is generated by a Poisson random measure of the type:

$$\eta_t^*(A) = \eta_t(A) + \sum_{p=1}^{P} \eta_t^{(p)} \int_A \delta_{\xi_p}(x)\mathrm{d}x \,,$$

where $\{\xi_1, \ldots, \xi_P\} \subset \mathbb{S}^2$, each mapping $t \mapsto \eta_t^{(p)}$ is an independent Poisson point process over $[0, \infty)$ with intensity λ_p, and

$$\left\{\int_A \delta_{\xi_p}(x)\mathrm{d}x = 1\right\} \Longleftrightarrow \{\xi_p \in A\} \;.$$

In this case, one has that η_t^* is a Poisson measure with atomic intensity

$$\mu_t^*(A) := \mathbb{E}[\eta_t^*(A)] = R_t \int_A f(x)\mathrm{d}x + \sum_{p=1}^{P} \lambda_p t \cdot \int_A \delta_{\xi_p}(x)\mathrm{d}x \,.$$

In this context, the informal expression "searching for point sources" can then be translated into "testing for $P = 0$" or "jointly testing for $\lambda_p > 0$ at $p = 1, \ldots, P$." The number P and the locations $\{\xi_1, \ldots \xi_P\}$ can be in general known or unknown. We refer for instance to [9, 30] for astrophysical applications of these ideas.

3.1 Bounds in Dimension One

The following result is proved in [7], and it is established applying the content of Theorem 1, Part 1, to the random variables $\tilde{\beta}_{jk}^{(R_t)}$ introduced in the previous subsection. In the next statement, we write $N \sim \mathcal{N}(0, 1)$ to denote a centered

Gaussian random variable with unit variance. Recall that $\zeta_2 := \sup_{x \in \mathbb{S}^2} |f(x)|$, $p \geq 1$, and that the constants q_p, q'_p have been defined in (3).

Proposition 1 (See [7]) *For every j, k and every $t > 0$, one has that*

$$d_W\left(\tilde{\beta}_{jk}^{(R_t)}, N\right) \leq \frac{(q'_3)^3 \zeta_2 B^j}{\sqrt{R_t} \sigma_{jk}^3}.$$

It follows that for any sequence $(j(n), k(n), t(n))$, $\tilde{\beta}_{j(n)k(n)}^{(R_{t(n)})}$ converges in distribution to N, as $n \to \infty$, provided $B^{2j(n)} = o(R_{t(n)})$ (remember that the family $\{\sigma_{jk}\}$ is bounded from above and below) .

The Gaussian approximation that we reported can be given the following heuristic interpretation. It is natural to view the factor B^{-j} as the "effective scale" of the wavelet, i.e., the radius of the region centered at ξ_{jk} where the wavelet function concentrates its energy. Because needlets are isotropic, this "effective area" is of order B^{-2j}. For governing measures with bounded densities which are bounded away from zero, the expected number of observations on a spherical cap of radius B^{-j} around ξ_{jk} is hence given by

$$\text{card}\left\{X_i : d(X_i, \xi_{jk}) \leq B^{-j}\right\} \simeq R_t \int_{d(x,\xi_{jk}) \leq B^{-j}} f(\xi_{jk}) \mathrm{d}x,$$

where

$$\zeta_1 B^{-2j} R_t \leq \eta_t \int_{d(x) \leq B^j} f(\xi_{jk}) \mathrm{d}x \leq \zeta_2 B^{-2j} R_t.$$

Because the Central Limit Theorem can only hold when the effective number of observations grows to infinity, the condition $B^{-2j} R_t \to \infty$ is consequently rather natural.

3.2 Multidimensional Bounds

A natural further step is to exploit Part 2 of Theorem 1 for the computation of multidimensional Berry–Esseen bounds involving vectors of needlet coefficients of the type (5). We stress that it is possible here to allow for a growing number of coefficients to be evaluated simultaneously, and investigate the bounds that can be obtained under these circumstances. More precisely, it is possible to focus on

$$\tilde{\beta}_{j(t)\cdot}^{(R_t)} := \left(\tilde{\beta}_{j(t)k_1}^{(R_t)}, \ldots, \tilde{\beta}_{j(t)k_{d_t}}^{(R_t)}\right),$$

where $d_t \to \infty$, as $t \to \infty$. Throughout the sequel, we shall assume that the points at which these coefficients are evaluated satisfy the condition:

$$\inf_{k_1 \neq k_2 = 1,\ldots,d_t} d\left(\xi_{j(t)k_1}, \xi_{j(t)k_2}\right) \approx \frac{1}{\sqrt{d_t}}, \tag{6}$$

where the symbol $\approx$ indicates in general that the ratio of two positive numerical sequences is bounded from above and below. Condition (6) is rather minimal; in fact, the cubature points for a standard needlet/wavelet construction can be taken to form a $(d_t)^{-1/2}$-net [2, 8, 19, 25], so that the following, stronger condition holds:

$$\inf_{k_1 \neq k_2 = 1,\ldots,d_t} d\left(\xi_{j(t)k_1}, \xi_{j(t)k_2}\right) \approx \sup_{k_1 \neq k_2 = 1,\ldots,d_t} d\left(\xi_{j(t)k_1}, \xi_{j(t)k_2}\right) \approx \frac{1}{\sqrt{d_t}} .$$

The following result is the main achievement in [7].

Theorem 2 *Let the previous assumptions and notation prevail. Then for all $\tau = 2, 3 \ldots$, there exist positive constants c and c', (depending on τ, ζ_1, ζ_2 but not from $t, j(t), d(t)$) such that we have*

$$d_2\left(\tilde{\beta}^{(R_t)}_{j(t).}, N\right) \leq \frac{c d_t}{\left(1 + B^{j(t)} \inf_{k_1 \neq k_2 = 1,\ldots,d_t} d\left(\xi_{j(t)k_1}, \xi_{j(t)k_2}\right)\right)^{\tau}} + \frac{\sqrt{2\pi}}{8} \frac{c' d_t B^{j(t)}}{\sqrt{R_t \sigma^2_{j(t)k_1} \sigma^2_{j(t)k_2} \sigma^2_{j(t)k_3}}} .$$

Under tighter conditions on the rate of growth of $d_t, B^{j(t)}$ with respect to R_t, it is possible to obtain a much more explicit bound, as follows:

Corollary 1 *Let the previous assumptions and notation prevail, and assume moreover that there exist α, β such that, as $t \to \infty$*

$$B^{2j(t)} \approx R_t^{\alpha}, \quad 0 < \alpha < 1, \quad d_t \approx R_t^{\beta}, \quad 0 < \beta < 2\alpha.$$

Then, there exists a constant κ (depending on ζ_1, ζ_2, but not on j, d_j, B) such that

$$d_2\left(\tilde{\beta}^{(R_t)}_{j(t).}, Z\right) \leq \kappa \frac{d_t B^{j(t)}}{\sqrt{R_t}} , \tag{7}$$

for all vectors $\left(\tilde{\beta}^{(R_t)}_{jk_1}, \ldots, \tilde{\beta}^{(R_t)}_{jk_{d_t}}\right)$, such that (6) holds.

To make the previous results more explicit, assume that d_t scales as $B^{2j(t)}$; loosely speaking, this corresponds to the cardinality of cubature points at scale j, so in a sense we would be focussing on the "whole" set of coefficients needed for exact reconstruction of a bandlimited function at that scale. In these circumstances, however, the "covariance" term $A(t)$, i.e., the first element on the right-hand side of (7), is no longer asymptotically negligible and the approximation with Gaussian independent variables cannot be expected to hold (the approximation may however be implemented in terms of a Gaussian vector with dependent components).

On the other hand, for the second term, convergence to zero when $d_{j(t)} \approx B^{2j(t)}$ requires $B^{3j(t)} = o(\sqrt{R_t})$. In terms of astrophysical applications, for $R_t \simeq 10^{12}$ this implies that one can focus on scales until $180°/B^j \simeq 180°/10^2 \simeq 2°$; this is close to the resolution level considered for ground-based cosmic rays experiments such as ARGO-YBJ (see [9]). Of course, this value is much lower than the factor $B^j = o(\sqrt{R_t}) = o(10^6)$ required for the Gaussian approximation to hold in the one-dimensional case (e.g., on a univariate sequence of coefficients, for instance corresponding to a single location on the sphere).

As mentioned earlier, in this chapter we presented the specific framework of spherical Poisson fields, which we believe is of interest from the theoretical and the applied point of view. It is readily verified, however, how these results continue to hold with trivial modifications in a much greater span of circumstances, indeed in some cases with simpler proofs. For instance, assume that one observes a sample of i.i.d. random variables $\{X_t\}$, with probability density function $f(\cdot)$ which is bounded and has support in $[a, b] \subset \mathbb{R}$. Consider the kernel estimates

$$\hat{f}_n(x_{nk}) := \frac{1}{nB^{-j}} \sum_{t=1}^{n} K\left(\frac{X_t - x_{nk}}{B^{-j}}\right) , \tag{8}$$

where $K(\cdot)$ denotes a compactly supported and bounded kernel satisfying standard regularity conditions, and for each j the evaluation points $(x_{n0}, \ldots, x_{nB^j})$ form a B^{-j}-net; for instance

$$a = x_{n0} < x_{n1} < \cdots < x_{nB^j} = b, \quad x_{nk} = a + k\frac{b-a}{B^j}, \quad k = 0, 1, \ldots, B^j.$$

Conditionally on $\eta_t([a, b]) = n$, (8) has the same distribution as

$$\hat{f}_{\eta_t}(x_{nk}) := \frac{1}{\eta_t[a, b]B^{-j}} \int_a^b K\left(\frac{u - x_{nk}}{B^{-j}}\right) \eta_t(\mathrm{d}u) ,$$

where η_t is a Poisson measure governed by $R_t \times \int_A f(x)\mathrm{d}x$ for all $A \subset [a, b]$. Considering that $\frac{\eta_t}{R_t} \to 1$ a.s., a bound analogous to (7) can be established with little efforts for the vector $\hat{f}_n(x_{n.}) := \left\{\hat{f}_n(x_{n1}), \ldots, \hat{f}_n(x_{nB^j})\right\}$. Rather than discussing these developments, though, we move to more general nonlinear transforms of wavelets coefficients, as considered by Bourguin et al. [5].

4 Nonlinear Transforms and U-Statistics

In the sequel, it is convenient to rewrite the needlet transforms in the following, equivalent version:

$$\psi_j(x,\xi) = \sqrt{\lambda_{j\xi}} \sum_{l\in\Lambda_j} b\left(\frac{l}{B^j}\right) L_l\left(\langle x,\xi\rangle\right)\ ,\ \Lambda_j := \left[2B^{j-1}, 2B^{j+1}\right]\ .$$

We shall repeatedly use the following bounds on their L^p-norms

$$c_p B^{j(p-2)} \leq \left\|\psi_j\right\|^p_{L^p(\mathbb{S}^2)} \leq C_p B^{j(p-2)}. \tag{9}$$

We shall now review some results on U-statistics of order 2 (see [14]) on the sphere, so we have as before $\mathbb{X} = [0,t]\times\mathbb{S}^2$, $\mathscr{X}$ being the corresponding Borel σ-field and $\mu \equiv \mu_t(x) := \rho([0,t]) \times \nu(x)$, $x \in \mathbb{S}^2$ such that ν is absolutely continuous with respect to the Lebesgue measure over the sphere, allowing f to be its corresponding density function such that

$$\nu(\mathrm{d}x) = f(x)\,\mathrm{d}x.$$

Writing as before $\hat{\eta}_t(\mathrm{d}x) = \eta_t(\mathrm{d}x) - \mu_t(\mathrm{d}x)$ to be the compensated Poisson random measure, we obtain

$$U_{j\xi}(t) = \sum_{(x,y)\in\eta^2_{\neq}} h_{j\xi}(x,y)\ ; \tag{10}$$

we shall focus in particular on the kernel

$$h_{j\xi}(x,y) = (\psi_j(x;\xi) - \psi_j(y;\xi))^2. \tag{11}$$

It is readily seen that statistics such as (10) provide a natural estimator of

$$M_{j\xi,t} := \mathbb{E}\left[U_{j\xi}(t)\right] = \int_{\mathbb{S}^2\times\mathbb{S}^2} h_{j\xi}(x,y)\mu_t^2(\mathrm{d}x,\mathrm{d}y) = 2R_t^2\left(G_{(2)} - G^2_{(1)}\right),$$

where

$$G_{j\xi}(n) := \mathbb{E}_f\left[\psi_j^n(\cdot,\xi)\right] = \int_{\mathbb{S}^2} \psi^n_{j\xi}(x) f(x)\,\mathrm{d}x \tag{12}$$

and

$$G_{j\xi}(2) - G^2_{j\xi}(1) = \mathbb{E}_f\left[\psi_j^2(\cdot,\xi)\right] - \mathbb{E}_f\left[\psi_j(\cdot,\xi)\right]^2 := \sigma_f^2(\xi)$$

is the conditional variance of $\psi_j(X_t;\xi)$, e.g., the variance of $\psi_j(X;\xi)$ for X a single random variable with density $f(\cdot)$. As such, it can be used as a goodness-of-fit testing procedure, or to check uniformity ($f(\cdot) = (4\pi)^{-1}$).
Applying [14, Theorem 1], we get

$$\begin{aligned} U_{j\xi}(t) &= I_2(h_{j\xi}(x,y)) + I_1\left(2\int_{\mathbb{S}^2} h_{j\xi}(x,y)\mu_t(\mathrm{d}y)\right) + \int_{\mathbb{S}^2\times\mathbb{S}^2} h_{j\xi}(x,y)\nu^2(\mathrm{d}x,\mathrm{d}y) \\ &= I_2(h_{j\xi}(x,y)) + I_1\left(h_{j\xi}^{(1,t)}(x)\right) + \int_{\mathbb{S}^2\times\mathbb{S}^2} h(x,y)\nu^2(\mathrm{d}x,\mathrm{d}y), \end{aligned}$$

where

$$h_{j\xi}^{(1,t)} := \int_{\mathbb{S}^2} h_{j\xi}(x,y)\mu_t(\mathrm{d}y).$$

As for the linear case, we shall then introduce a normalized process

$$\tilde{U}_{j\xi}(t) := \frac{U_{j\xi}(t) - M_{j\xi,t}}{V_{j\xi,t}^{\frac{1}{2}}},$$

where

$$M_{j\xi,t} := \mathbb{E}\left[U_{j\xi}(t)\right] = \int_{\mathbb{S}^2\times\mathbb{S}^2} h_{j\xi}(x,y)\mu_t^2(\mathrm{d}x,\mathrm{d}y) = 2R_t^2\left(G_{(2)} - G_{(1)}^2\right),$$

and

$$V_{j\xi,t} := \mathrm{Var}\left[U_{j\xi}(t)\right] = \mathrm{Var}\left[\sum_{(x,y)\in\eta^2_{\neq}} h_{j\xi}(x,y)\right].$$

Here and in the sequel, we write for brevity $G_{(n)}$ rather than $G_{j\xi}(n)$ when no confusion is possible. It can be shown that (see [5] for details)

$$V_{j\xi,t} = R_t^3\Sigma_{j\xi} + o\left(R_t^3\right),$$

where

$$\Sigma_{j\xi} := \left[4\left(\left(G_{j\xi}(4) - G_{j\xi}^2(2)\right) + 4\left(G_{j\xi}(2) - G_{j\xi}^2(1)\right)^2\right)\right].$$

We can hence focus on the (Hoeffding) decomposition

$$\tilde{U}_{j\xi;t} = I_1\left(\tilde{h}^{(1,t)}_{j\xi}(x)\right) + I_2(\tilde{h}_{j\xi}(x,y)), \tag{13}$$

where

$$\tilde{h}_{j\xi}(x,y) = \frac{h_{j\xi}(x,y)}{\sqrt{R_t^3 \Sigma_{j\xi}}}$$

and

$$\tilde{h}^{(1,t)}_{j\xi}(x) = \frac{h^{(1,t)}_{j\xi}(x)}{\sqrt{R_t^3 \Sigma_{j\xi}}} = \frac{\int_{\mathbb{S}^2} h_{j\xi}(x,y)\mu_t(\mathrm{d}y)}{\sqrt{R_t^3 \Sigma_{j\xi}}} .$$

It is then possible to establish the following result

Theorem 3 *As $R_t \to \infty$, for any $j > 0$ and ξ belonging to set of cubature points $\mathcal{Q}_l \subset \mathbb{S}^2$,*

$$d_W\left(\tilde{U}_{j\xi;t}, N\right) \leq \frac{CB^j}{\sqrt{R_t}} ,$$

where N denotes a standard Gaussian variable.

For details of the proof, we refer to [5]. The main ideas can be sketched as follows; for any j, ξ, in the decomposition (13), it is possible to show that

$$\mathbb{E}\left[\left(I_2(\tilde{h}_{j\xi}(x,y))\right)^2\right] = O\left(\frac{1}{R_t}\right),$$

so that $I_1\left(\tilde{h}^{(1,t)}_{j\xi}(x)\right)$ is the dominant term as $R_t \to \infty$. Therefore we apply again Theorem 1, Part 1, with the help of the following technical result.

Lemma 1 *Let $\Gamma = \left\{k_1, \ldots, k_D : \sum_{i=1}^D k_i = K, k_i \neq k_j \ \forall i \neq j\right\}$,*

$$L_K(x_1, x_2, .., x_D) = \sum_{\{k_1,\ldots,k_D\}\in\Gamma} c_{k_1,\ldots,k_D} \prod_{i=1}^{D} \psi_j(x_i,\xi)^{k_i} ,$$

and C_{k_i} as defined in (9). Hence

$$\left| \int_{(\mathbb{S}^2)^{\otimes D}} L_K(x_1, x_2, \ldots, x_D)\, \nu^{\otimes D}(\mathrm{d}x_1 \ldots \mathrm{d}x_D) \right| \leq \sum_{\{k_1,\ldots,k_D\}\in\Gamma} c_{k_1,\ldots,k_D} \left(\prod_{i=1}^{D} C_{k_i} \right) B^{j\left(K-2D+2\sum_{i=1}^{D}\delta_{k_i}^{0}\right)},$$

where $\delta_0^{k_i}$ is the Kronecker delta function.

Indeed, the previous Lemma yields the following useful bound:

$$\left| \int_{(\mathbb{S}^2)^{\otimes D}} L_K(x_1, x_2, \ldots, x_D)\, \nu^{\otimes D}(\mathrm{d}x_1 \ldots \mathrm{d}x_D) \right| \leq M B^{j(K-2D+2N_0)}.$$

From this bound, it is for instance easy to see that

$$m_n B^{(n-2)j} \leq G_{j\xi}(n) \leq M_n B^{(n-2)j},$$

where $G_{j\xi}(n)$ is defined in (12); it follows also that

$$m_4 B^{2j} \leq \Sigma_{j\xi} \leq M_4 B^{2j}.$$

As for the linear case, (3) can be extended to growing arrays of statistics, and then applied, for instance, to high-frequency local estimates of variances/dispersion or to the classical two-sample problem. Details and further discussion are provided in [5].

References

1. Adler, R.J., Jonathan, E.T.: Random Fields and Geometry. Springer Monographs in Mathematics. Springer, New York (2007)
2. Baldi, P., Kerkyacharian, G., Marinucci, D., Picard, D.: Subsampling needlet coefficients on the sphere. Bernoulli **15**, 438–463 (2009)
3. Baldi, P., Kerkyacharian, G., Marinucci, D., Picard, D.: Adaptive density estimation for directional data using needlets. Ann. Stat. **37**(6A), 3362–3395 (2009)
4. Bourguin, S., Peccati, G.: The Malliavin-Stein method on the Poisson space. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 185–228. Springer, Cham (2016)
5. Bourguin, S., Durastanti, C., Marinucci, D., Peccati, G.: Gaussian approximations of non-linear statistics on the sphere. J. Math. Anal. Appl. **436**(2), 1121–1148 (2016)

6. Chen, L.H.Y., Goldstein, L., Shao, Q.-M.: Normal Approximation by Stein's Method. Springer, Berlin (2011)
7. Durastanti, C., Marinucci, D., Peccati, G.: Normal approximations for wavelet coefficients on spherical Poisson fields. J. Math. Anal. Appl. **409**(1), 212–227 (2014)
8. Geller, D., Mayeli, A.: Continuous wavelets on manifolds. Math. Z. **262**, 895–927 (2009)
9. Iuppa, R.: A needlet-based approach to the full-sky data analysis. arXiv: 1112.0248 (2011)
10. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, Berlin (1987)
11. Koo, J.-Y., Kim, P.T.: Sharp adaptation for spherical inverse problems with applications to medical imaging. J. Multivariate Anal. **99**, 165–190 (2008)
12. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space I: contractions, cumulants and geometric random graphs. Electron. J. Probab. **18**(32), 1–32 (2011)
13. Lachièze-Rey, R., Peccati, G.: Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics. Stoch. Process. Appl. **123**(12), 4186–4218 (2012)
14. Lachièze-Rey, R., Reitzner, M.: U-statistics in stochastic geometry. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 229–253. Springer, Cham (2016)
15. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
16. Last, G., Penrose, M., Schulte, M., Thäle, C.: Moments and central limit theorems for some multivariate Poisson functionals. Adv. Appl. Probab. **46**(2), 348–364 (2014)
17. Marinucci, D., Peccati, G.: Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications. Lecture Notes of the London Mathematical Society, vol. 389. Cambridge University Press, Cambridge (2011)
18. Minh, N.T.: Malliavin-Stein method for multi-dimensional U-statistics of Poisson point processes. Manuscript (2011)
19. Narcowich, F.J., Petrushev, P., Ward, J.D.: Localized tight frames on spheres. SIAM J. Math. Anal. **38**, 574–594 (2006)
20. Narcowich, F.J., Petrushev, P., Ward, J.D.: Decomposition of besov and Triebel-Lizorkin spaces on the sphere. J. Funct. Anal. **238**(2), 530–564 (2006)
21. Nourdin, I., Peccati, G.: Normal Approximations Using Malliavin Calculus: From Stein's Method to Universality. Cambridge University Press, Cambridge (2012)
22. Peccati, G., Taqqu, M.S.: Wiener Chaos: Moments, Cumulants and Diagrams. Bocconi & Springer Series, vol. 1, Springer, Milan (2011)
23. Peccati, G., Zheng, C.: Multi-dimensional Gaussian fluctuations on the Poisson space. Electron. J. Probab. **15**(48), 1487–1527 (2010)
24. Peccati, G., Solé, J.-L., Taqqu, M.S., Utzet, F.: Stein's method and normal approximation of Poisson functionals. Ann. Probab. **38**(2), 443–478 (2010)
25. Pesenson, I.: Sampling of band-limited vectors. J. Fourier Anal. Appl. **7**(1), 93–100 (2001)
26. Pesenson, I.: Poincaré-type inequalities and reconstruction of Paley-Wiener functions on manifolds. J. Geom. Anal. **14**(1), 101–121 (2004)
27. Privault, N.: Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales. Springer, New York (2009)
28. Reitzner, M., Schulte, M.: Central limit theorems for U-statistics of Poisson point processes. Ann. Probab. **41**(6), 3879–3909 (2013)
29. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Springer, New York (2008)
30. Scodeller, S., Hansen, F.K., Marinucci, D.: Detection of new point sources in WMAP 7 year data using internal templates and needlets. Astrophys. J. **573**(27), 1–12 (2012)
31. Starck, J.-L., Fadili, J.M., Digel, S., Zhang, B., Chiang, J.: Source detection using a 3D sparse representation: application to the Fermi gamma-ray space telescope. arXiv: 0904.3299 (2009)
32. Stein, E.M., Weiss, G.: Introduction to Fourier Analysis on Euclidean Spaces. Princeton University Press, Princeton (1971)

Determinantal Point Processes

Laurent Decreusefond, Ian Flint, Nicolas Privault, and Giovanni Luca Torrisi

Abstract In this survey we review two topics concerning determinantal (or fermion) point processes. First, we provide the construction of diffusion processes on the space of configurations whose invariant measure is the law of a determinantal point process. Second, we present some algorithms to sample from the law of a determinantal point process on a finite window. Related open problems are listed.

1 Introduction

Determinantal (or fermion) point processes have been introduced in [27] to represent configurations of fermions. Determinantal point processes play a fundamental role in the theory of random matrices as the eigenvalues of many ensembles of random matrices form a determinantal point process, see, e.g., [18]. The full existence theorem for these processes was proved in [34], in which many examples occurring in mathematics and physics were discussed. The construction of [34] has been extended in [32] with the introduction of the family of α-determinantal point processes.

Determinantal point processes have notable mathematical properties, e.g., their Laplace transforms, Janossy densities, and Papangelou conditional intensities admit closed form expressions. Due to their repulsive nature, determinantal point processes have been recently proposed as models for nodes' locations in wireless communication, see [28, 37].

L. Decreusefond (✉) • I. Flint
Département INFRES, Telecom Paristech, 23 avenue d'Italie, 75013 Paris, France
e-mail: laurent.decreusefond@mines-telecom.fr; iflint@ntu.edu.sg

N. Privault
Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link, Singapore 637371, Singapore
e-mail: nprivault@ntu.edu.sg

G.L. Torrisi
CNR-Istituto per le Applicazioni del Calcolo "Mauro Picone", Via dei Taurini 19, 00185 Roma, Italy
e-mail: torrisi@iac.rm.cnr.it

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*, Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5_10

This paper is structured as follows. In Sect. 2 we give some preliminaries on point processes, including the definition of determinantal point processes, the expression of their Laplace transform (Theorem 1), Janossy densities (Proposition 3), and Papangelou intensity (Theorems 2 and 3), cf. [9, 10, 15, 15, 18, 23, 29, 32]. We also refer to [7, 13, 33] for the required background on functional analysis.

In Sect. 3 we review the integration by parts formula for determinantal point processes and its extension by closability, cf. [8, 11].

In Sect. 4 we report a result in [11] on the construction of a diffusion on the space of configurations which has the law of a determinantal point process as invariant measure. To this aim we use arguments based on the theory of Dirichlet forms, cf. [14, 25] and the appendix. It has to be noticed that the construction of the diffusion provided in [11] differs from that one given in [36], where alternative techniques are used.

Section 5 deals with the simulation of determinantal point processes. We provide two different simulation algorithms to sample from the law of a determinantal point processes on a compact. In particular, we describe the (standard) sampling algorithm given in [17] (see Algorithm 1 below) and an alternative simulation algorithm obtained by specializing the well-known routine to sample from the law of a finite point process with bounded Papangelou conditional intensity (see, e.g., [19, 20, 23] and Algorithm 2 below). We show that the number of steps in the latter algorithm grows logarithmically with the size of the initial dominating point process, which gives a rough idea of the simulation time required by this algorithm. Finally, we propose a new approximate simulation algorithm for the Ginibre point process, which presents advantages in terms of complexity and CPU time.

Finally, some open problems are listed in Sect. 6.

2 Preliminaries

2.1 Locally Finite Point Processes, Correlation Functions, Janossy Density, and Papangelou Intensity

Let $\mathbb{X}$ be a locally compact second countable Hausdorff space, and $\mathscr{X}$ be the Borel σ-algebra on $\mathbb{X}$. For any subset $A \subseteq \mathbb{X}$, let $|A|$ denote the cardinality of A, setting $|A| = \infty$ if A is not finite. We denote by $\mathbf{N}_s$ the set of locally finite point configurations on $\mathbb{X}$:

$$\mathbf{N}_s := \{\xi \subseteq \mathbb{X} \ : \ |\xi \cap \Lambda| < \infty \quad \text{for all relatively compact sets } \Lambda \subset \mathbb{X}\}.$$

In fact, $\mathbf{N}_s$ can be identified with the set of all simple nonnegative integer-valued Radon measures on $\mathbb{X}$ (an integer-valued Radon measure ν is said to be simple if for all $x \in \mathbb{X}$, $\nu(\{x\}) \in \{0, 1\}$). Hence, it is naturally topologized by the vague topology, which is the weakest topology such that for any continuous and compactly supported function f on $\mathbb{X}$, the mapping

$$\xi \mapsto \langle f, \xi \rangle := \sum_{y \in \xi} f(y)$$

is continuous. We denote by $\mathcal{N}_s$ the corresponding Borel σ-field. For $\xi \in \mathbf{N}_s$, we write $\xi \cup y_0 = \xi \cup \{y_0\}$ for the addition of a particle at y_0 and $\xi \setminus y_0 = \xi \setminus \{y_0\}$ for the removal of a particle at y_0. We define the set of finite point configurations on $\mathbb{X}$ by

$$\mathbf{N}_s^f := \{\xi \subseteq \mathbb{X} \colon \ |\xi| < \infty\},$$

which is equipped with the trace σ-algebra $\mathcal{N}_s^f = \mathcal{N}_s|_{\mathbf{N}_s^f}$. For any relatively compact subset $\Lambda \subseteq \mathbb{X}$, let $\mathbf{N}_s(\Lambda)$ be the space of finite configurations on Λ, and $\mathcal{N}_s(\Lambda)$ the associated (trace-) σ-algebra. As in [15], we define for any Radon measure μ on $\mathbb{X}$ the (μ-)sample measure L^μ on $(\mathbf{N}_s^f, \mathcal{N}_s^f)$ by

$$\int_{\mathbf{N}_s^f} f(\alpha)\, L^\mu(\mathrm{d}\alpha) := \sum_{n \geq 0} \frac{1}{n!} \int_{\mathbb{X}^n} f(\{x_1, \ldots, x_n\})\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n), \tag{1}$$

for any measurable $f : \mathbf{N}_s^f \to \mathbb{R}_+$. Similarly, we define its restriction to the relatively compact set $\Lambda \subseteq \mathbb{X}$ by

$$\int_{\mathbf{N}_s(\Lambda)} f(\alpha)\, L_\Lambda^\nu(\mathrm{d}\alpha) := \sum_{n \geq 0} \frac{1}{n!} \int_{\Lambda^n} f(\{x_1, \ldots, x_n\})\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n),$$

for any measurable $f : \mathbf{N}_s(\Lambda) \to \mathbb{R}_+$. A simple and locally finite point process η is defined as a random element on a probability space $(\Omega, \mathcal{A})$ with values in $\mathbf{N}_s$. We denote its distribution by $\mathbb{P}$. It is characterized by its Laplace transform $\mathcal{L}_\eta$, which is defined, for any measurable nonnegative function f on $\mathbb{X}$, by

$$\mathcal{L}_\eta(f) = \int_{\mathbf{N}_s} e^{-\langle f, \xi \rangle}\, \mathbb{P}(\mathrm{d}\xi).$$

We denote the expectation of an integrable random variable F defined on $(\mathbf{N}_s, \mathcal{N}_s, \mathbb{P})$ by

$$\mathbb{E}[F(\eta)] := \int_{\mathbf{N}_s} F(\xi)\, \mathbb{P}(\mathrm{d}\xi).$$

For ease of notation, we define by

$$\xi_A := \xi \cap A,$$

the restriction of $\xi \in \mathbf{N}_s$ to a set $A \subset \mathbb{X}$. The restriction of $\mathbb{P}$ to $\mathscr{N}_s(A)$ is denoted by $\mathbb{P}_A$ and the number of points of ξ_A, i.e., $\xi(A) := |\xi \cap A|$, is denoted by $\xi(A)$. A point process η is said to have a correlation function $\rho : \mathbf{N}_s^f \to [0, \infty)$ with respect to (*w.r.t.*) a Radon measure μ on $(\mathbb{X}, \mathscr{X})$ if ρ is measurable and

$$\int \sum_{\alpha\subset\xi,\ \alpha\in\mathbf{N}_s^f} f(\alpha)\,\mathbb{P}(\mathrm{d}\xi) = \int_{\mathbf{N}_s^f} f(\alpha)\,\rho(\alpha)\,L^{\mu}(\mathrm{d}\alpha),$$

for all measurable nonnegative functions f on $\mathbf{N}_s^f$. When such a measure μ exists, it is known as the intensity measure of η. For $\alpha = \{x_1, \dots, x_k\}$, where $k \geq 1$, we will sometimes write $\rho(\alpha) = \rho_k(x_1, \dots, x_k)$ and call ρ_k the k-th correlation function. Here ρ_k is a symmetric function on $\mathbb{X}^k$. Similarly, the correlation functions of η, *w.r.t.* a Radon measure μ on $\mathbb{X}$, are (if they exist) measurable symmetric functions $\rho_k : \mathbb{X}^k \longrightarrow [0, \infty)$ such that

$$\mathbb{E}\left[\prod_{i=1}^{k} \eta(B_i)\right] = \int_{B_1\times\cdots\times B_k} \rho_k(x_1, \dots, x_k)\,\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k),$$

for any family of mutually disjoint bounded subsets $B_1, \dots, B_k$ of $\mathbb{X}$, $k \geq 1$. The previous formula can be generalized as follows:

Proposition 1 *Let $B_1, \dots, B_n$ be disjoint bounded Borel subsets of $\mathbb{X}$. Let $k_1, \dots, k_n$ be integers such that $\sum_{i=1}^{n} k_i = N$. Then,*

$$\mathbb{E}\left[\prod_{i=1}^{n} \frac{\eta(B_i)!}{(\eta(B_i) - k_i)!}\right] = \int_{B_1^{k_1}\times\ldots\times B_n^{k_n}} \rho(\{x_1, \dots, x_N\})\,\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_N).$$

We require in addition that $\rho_n(x_1, \dots, x_n) = 0$ whenever $x_i = x_j$ for some $1 \leq i \neq j \leq n$. Heuristically, ρ_1 is the particle density with respect to μ, and

$$\rho_n(x_1, \dots, x_n)\,\mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n)$$

is the probability of finding a particle in the vicinity of each x_i, $i = 1, \dots, n$. For any relatively compact subset $\Lambda \subseteq \mathbb{X}$, the *Janossy densities* of η, *w.r.t.* a Radon measure μ on $\mathbb{X}$, are (if they exist) measurable functions $j_\Lambda^n : \Lambda^n \to [0, \infty)$ satisfying for all measurable functions $f : \mathbf{N}_s(\Lambda) \to [0, \infty)$,

$$\mathbb{E}\left[f(\boldsymbol{\eta}_\Lambda)\right] = \sum_{n\geq 0} \frac{1}{n!} \int_{\Lambda^n} f(\{x_1, \dots, x_n\}) j_\Lambda^n(x_1, \dots, x_n)\ \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n). \tag{2}$$

Using the simplified notation $j_\Lambda(\alpha) := j_\Lambda^n(x_1, \dots, x_n)$, for $\alpha = \{x_1, \dots, x_n\}$, where $n \geq 1$, by (2) it follows that j_Λ is the density of $\mathbb{P}_\Lambda$ with respect to L_Λ^μ, when $\mathbb{P}_\Lambda \ll L_\Lambda^\mu$. Now we list some properties of the Janossy densities.

- Symmetry:

$$j_\Lambda^n\left(x_{\sigma(1)}, \dots, x_{\sigma(n)}\right) = j_\Lambda^n\left(x_1, \dots, x_n\right),$$

  for every permutation σ of $\{1, \dots, n\}$.
- Normalization constraint: for each relatively compact subset $\Lambda \subseteq \mathbb{X}$,

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} j_\Lambda^n(x_1, \dots, x_n)\ \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_n) = 1.$$

For $n \geq 1$, the Janossy density $j_\Lambda^n(x_1, \dots, x_n)$ is in fact the joint density (multiplied by a constant) of the n points given that the point process has exactly n points. For $n = 0$, $j_\Lambda^0(\emptyset)$ is the probability that there are no points in Λ. We also recall that the Janossy densities can be recovered from the correlation functions via the relation

$$j_\Lambda^n(x_1, \dots, x_n) = \sum_{m\geq 0} \frac{(-1)^m}{m!} \int_{\Lambda^m} \rho_{n+m}(x_1, \dots, x_n, y_1, \dots, y_m)\, \mu(\mathrm{d}y_1)\cdots\mu(\mathrm{d}y_m),$$

and vice versa using the equality

$$\rho_n(x_1, \dots, x_n) = \sum_{m\geq 0} \frac{1}{m!} \int_{\Lambda^m} j_\Lambda^{m+n}(x_1, \dots, x_n, y_1, \dots, y_m)\, \mu(\mathrm{d}y_1)\cdots\mu(\mathrm{d}y_m),$$

see [9, Theorem 5.4.II].

Following [15], we now recall the definition of the so-called reduced and reduced compound Campbell measures. The reduced Campbell measure of a point process η is the measure C_η on the product space $(\mathbb{X} \times \mathbf{N}_s, \mathscr{X} \otimes \mathscr{N}_s)$ defined by

$$C_\eta(A \times B) = \int_{\mathbf{N}_s} \sum_{x\in\xi} \mathbb{1}_A(x)\mathbb{1}_B(\xi \setminus x)\, \mathbb{P}(\mathrm{d}\xi).$$

The reduced compound Campbell measure of a point process η is the measure $\hat{C}_\eta$ on the product space $(\mathbf{N}_s^f \times \mathbf{N}_s, \mathscr{N}_s^f \otimes \mathscr{N}_s)$ defined by

$$\hat{C}_\eta(A \times B) = \int_{\mathbf{N}_s} \sum_{\alpha \subset \xi,\ \alpha \in \mathbf{N}_s^f} \mathbb{1}_A(\alpha) \mathbb{1}_B(\xi \setminus \alpha)\, \mathbb{P}(\mathrm{d}\xi).$$

The integral versions of the equations above can be written respectively as

$$\int h(x, \xi)\, C_\eta(\mathrm{d}x \times \mathrm{d}\xi) = \int \sum_{x \in \xi} h(x, \xi \setminus x)\, \mathbb{P}(\mathrm{d}\xi), \tag{3}$$

for all nonnegative measurable functions $h : \mathbb{X} \times \mathbf{N}_s \to \mathbb{R}_+$, and

$$\int h(\alpha, \xi)\, \hat{C}_\eta(\mathrm{d}\alpha \times \mathrm{d}\xi) = \int \sum_{\alpha \subset \xi,\ \alpha \in \mathbf{N}_s^f} h(\alpha, \xi \setminus \alpha)\, \mathbb{P}(\mathrm{d}\xi),$$

for all nonnegative measurable functions $h : \mathbf{N}_s^f \times \mathbf{N}_s \to \mathbb{R}_+$. Comparing (3) with the well-known Mecke formula (see (7) in [21]) leads us to introduce the following condition:

(Σ): $\quad C_\eta \ll \mu \otimes \mathbb{P}$.

The Radon–Nikodym derivative c of C_η w.r.t. $\mu \otimes \mathbb{P}$ is called (a version of) the Papangelou intensity of η. Assumption (Σ) implies that $\hat{C}_\eta \ll L^\mu \otimes \mathbb{P}$ and we denote the Radon–Nikodym derivative of $\hat{C}_\eta$ w.r.t. $L^\mu \otimes \mathbb{P}$ by $\hat{c}$, and call $\hat{c}$ the compound *Papangelou intensity* of η. One then has for any $\xi \in \mathbf{N}_s$, $\hat{c}(\emptyset, \xi) = 1$, as well as for all $x \in \mathbb{X}$, $\hat{c}(x, \xi) = c(x, \xi)$. The Papangelou intensity c has the following interpretation:

$$c(x, \xi)\, \mu(\mathrm{d}x)$$

is the probability of finding a particle in the vicinity of $x \in \mathbb{X}$ conditional on the configuration ξ.
The compound Papangelou intensity verifies the following commutation relation:

$$\hat{c}(\nu, \eta \cup \xi)\, \hat{c}(\eta, \xi) = \hat{c}(\nu \cup \eta, \xi), \tag{4}$$

for all $\eta, \nu \in \mathbf{N}_s^f$ and $\xi \in \mathbf{N}_s$. The recursive application of the previous relation also yields

$$\hat{c}(\{x_1, \ldots, x_n\}, \xi) = \prod_{k=1}^{n} c(x_k, \xi \cup x_1 \cup \cdots \cup x_{k-1}),$$

for all $x_1, \ldots, x_n \in \mathbb{X}$ and $\xi \in \mathbf{N}_s$, where we have used the convention $x_0 := \emptyset$.

The assumption (Σ), along with the definition of the reduced Campbell measure, allows us to write the following identity, known as the Georgii–Nguyen–Zessin identity:

$$\int_{\mathbf{N}_s} \sum_{y \in \xi} u(y, \xi \setminus y) \, \mathbb{P}(\mathrm{d}\xi) = \int_{\mathbf{N}_s} \int_{\mathbb{X}} u(z, \xi) \, c(z, \xi) \, \mu(\mathrm{d}z) \mathbb{P}(\mathrm{d}\xi), \tag{5}$$

for all measurable nonnegative functions $u : \mathbb{X} \times \mathbf{N}_s \to \mathbb{R}_+$. We also have a similar identity for the compound Papangelou intensity:

$$\int_{\mathbf{N}_s} \sum_{\alpha \subset \xi, \, \alpha \in \mathbf{N}_s^f} u(\alpha, \xi \setminus \alpha) \, \mathbb{P}(\mathrm{d}\xi) = \int_{\mathbf{N}_s} \int_{\mathbf{N}_s^f} u(\alpha, \xi) \, \hat{c}(\alpha, \xi) \, L^{\mu}(\mathrm{d}\alpha) \mathbb{P}(\mathrm{d}\xi), \tag{6}$$

for all measurable functions $u : \mathbf{N}_s^f \times \mathbf{N}_s \to \mathbb{R}_+$.
Note that Eqs. (5) and (6) are generalizations of Eqs. (1.7) and (1.8) of [21]. Indeed, in the case of the Poisson point process, $c(z, \xi) = 1$ and $c(\alpha, \xi) = 1$.
Combining relation (5) and the definition of the correlation functions, we find

$$\mathbb{E}[c(x, \eta)] = \rho_1(x),$$

for μ-a.e. $x \in \mathbb{X}$. More generally, using (6), we also have

$$\mathbb{E}[\hat{c}(\alpha, \eta)] = \rho(\alpha), \tag{7}$$

for $\mathbb{P}$-a.e. $\alpha \in \mathbf{N}_s^f$.

2.2 Kernels and Integral Operators

As usual, we denote by $\mathbb{X}$ a locally compact second countable Hausdorff space and by μ a Radon measure on $\mathbb{X}$. For any compact set $\Lambda \subseteq \mathbb{X}$, we denote by $L^2(\Lambda, \mu)$ the Hilbert space of complex-valued square integrable functions *w.r.t.* the restriction of the Radon measure μ on Λ, equipped with the inner product

$$\langle f, g \rangle_{L^2(\Lambda,\mu)} := \int_{\Lambda} f(x)\overline{g(x)} \, \mu(\mathrm{d}x), \quad f, g \in L^2(\Lambda, \mu),$$

where $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$. By definition, an integral operator $\mathcal{K} : L^2(\mathbb{X}, \mu) \to L^2(\mathbb{X}, \mu)$ with kernel $K : \mathbb{X}^2 \to \mathbb{C}$ is a bounded operator defined by

$$\mathcal{K}f(x) := \int_{\mathbb{X}} K(x,y)f(y)\,\mu(\mathrm{d}y), \quad \text{for } \mu\text{-almost all } x \in \mathbb{X}.$$

We denote by $\mathcal{P}_\Lambda$ the projection operator from $L^2(\mathbb{X},\mu)$ to $L^2(\Lambda,\mu)$ and define the operator $\mathcal{K}_\Lambda = \mathcal{P}_\Lambda \mathcal{K} \mathcal{P}_\Lambda$. We note that the kernel of $\mathcal{K}_\Lambda$ is given by $K_\Lambda(x,y) := \mathbb{1}_\Lambda(x)K(x,y)\mathbb{1}_\Lambda(y)$, for $x, y \in \mathbb{X}$. It can be shown that $\mathcal{K}_\Lambda$ is a compact operator.
The operator $\mathcal{K}$ is said to be Hermitian or self-adjoint if its kernel verifies

$$K(x,y) = \overline{K(y,x)}, \quad \text{for } \mu^{\otimes 2}\text{-almost all } (x,y) \in \mathbb{X}^2. \tag{8}$$

Equivalently, this means that the integral operators $\mathcal{K}_\Lambda$ are self-adjoint for any compact set $\Lambda \subseteq \mathbb{X}$. If $\mathcal{K}_\Lambda$ is self-adjoint, by the spectral theorem for self-adjoint and compact operators we have that $L^2(\Lambda,\mu)$ has an orthonormal basis $\{\varphi_j^\Lambda\}_{j\geq 1}$ of eigenfunctions of $\mathcal{K}_\Lambda$. The corresponding eigenvalues $\{\mu_j^\Lambda\}_{j\geq 1}$ have finite multiplicity (except possibly the zero eigenvalue) and the only possible accumulation point of the eigenvalues is the zero eigenvalue. In that case, the kernel K_Λ of $\mathcal{K}_\Lambda$ can be written as

$$K_\Lambda(x,y) = \sum_{n\geq 1} \mu_n^\Lambda \varphi_n^\Lambda(x)\overline{\varphi_n^\Lambda(y)}, \tag{9}$$

for $x, y \in \Lambda$. We say that an operator $\mathcal{K}$ is positive (respectively nonnegative) if its spectrum is included in $(0,+\infty)$ (respectively $[0,+\infty)$). For two operators $\mathcal{K}$ and $\mathcal{I}$, we say that $\mathcal{K} > \mathcal{I}$ (respectively $\mathcal{K} \geq \mathcal{I}$) in the operator ordering if $\mathcal{K} - \mathcal{I}$ is a positive operator (respectively nonnegative operator).
We say that a self-adjoint integral operator $\mathcal{K}_\Lambda$ is of trace class if

$$\sum_{n\geq 1} |\mu_n^\Lambda| < \infty,$$

and define the trace of $\mathcal{K}_\Lambda$ as $\operatorname{Tr}\mathcal{K}_\Lambda = \sum_{n\geq 1} \mu_n^\Lambda$. If $\mathcal{K}_\Lambda$ is of trace class for every compact subset $\Lambda \subseteq \mathbb{X}$, then we say that $\mathcal{K}$ is locally of trace class. It is easily seen that if a Hermitian integral operator $\mathcal{K} : L^2(\mathbb{X},\mu) \to L^2(\mathbb{X},\mu)$ is of trace class, then $\mathcal{K}^n$ is also of trace class for all $n \geq 2$. Indeed, $\operatorname{Tr}(\mathcal{K}^n) \leq \|\mathcal{K}\|_{op}^{n-1}\operatorname{Tr}(\mathcal{K})$, where $\|\mathcal{K}\|_{op}$ is the operator norm of $\mathcal{K}$.
Let Id denote the identity operator on $L^2(\mathbb{X},\mu)$ and let $\mathcal{K}$ be a trace class operator on $L^2(\mathbb{X},\mu)$. We define the Fredholm determinant of $\mathrm{Id} + \mathcal{K}$ as

$$\operatorname{Det}(\mathrm{Id} + \mathcal{K}) = \exp\left(\sum_{n\geq 1} \frac{(-1)^{n-1}}{n}\operatorname{Tr}(\mathcal{K}^n)\text{ă}\right). \tag{10}$$

It turns out that

$$\mathrm{Det}(\mathrm{Id} + \mathcal{K}) = \sum_{n\geq 0} \frac{1}{n!} \int_{\mathbb{X}^n} \det(K(x_i, x_j))_{1\leq i,j\leq n}\, \mu(\mathrm{d}x_1) \ldots \mu(\mathrm{d}x_n), \tag{11}$$

where K is the kernel of $\mathcal{K}$ and $\det(K(x_i, x_j))_{1\leq i,j\leq n}$ is the determinant of the $n \times n$ matrix $(K(x_i, x_j))_{1\leq i,j\leq n}$. Equation (11) was obtained in Theorem 2.4 of [32], see also [7] for more details on the Fredholm determinant.
We end this section by recalling the following result from [15, Lemma A.4]:

Proposition 2 *Let $\mathcal{K}$ be a nonnegative and locally of trace class integral operator on $L^2(\mathbb{X}, \mu)$. Then one can choose its kernel K (defined everywhere) such that the following properties hold:*

(i) *K is nonnegative, in the sense that for any $c_1, \ldots, c_n \in \mathbb{C}$ and μ-a.e. $x_1, \ldots, x_n \in \mathbb{X}$, we have $\sum_{i,j=1}^{n} \bar{c}_i K(x_i, x_j) c_j \geq 0$.*
(ii) *K is a Carleman kernel, i.e., $K_x = K(\cdot, x) \in L^2(\mathbb{X}, \mu)$ for μ-a.e. $x \in \mathbb{X}$.*
(iii) *For any compact subset $\Lambda \subseteq \mathbb{X}$, $\mathrm{Tr}\, \mathcal{K}_\Lambda = \int_\Lambda K(x, x)\, \mu(\mathrm{d}x)$ and*

$$\mathrm{Tr}\, (\mathcal{P}_\Lambda \mathcal{K}^k \mathcal{P}_\Lambda) = \int_\Lambda \langle K_x, \mathcal{K}^{k-2} K_x \rangle_{L^2(\Lambda,\mu)}\, \mu(\mathrm{d}x),$$

for $k \geq 2$.

Henceforth, the kernel of a nonnegative and locally of trace class integral operator $\mathcal{K}$ will be chosen according to the previous proposition.

2.3 Determinantal Point Processes

A locally finite and simple point process η on $\mathbb{X}$ is called *determinantal point process* if its correlation functions *w.r.t.* the Radon measure μ on $(\mathbb{X}, \mathscr{X})$ exist and are of the form

$$\rho_k(x_1, \ldots, x_k) = \det(K(x_i, x_j))_{1\leq i,j\leq k},$$

for any $k \geq 1$ and $x_1, \ldots, x_k \in \mathbb{X}$, where $K(\cdot, \cdot)$ is a measurable function. Throughout this paper we shall consider the following hypothesis:

(H1): *The operator $\mathcal{K}$ is locally of trace class, satisfies (8), and its spectrum is contained in $[0, 1)$, i.e., $0 \leq \mathcal{K} < \mathrm{Id}$ in the operator ordering. We denote by K the kernel of $\mathcal{K}$.*

By the results in [27, 34] (see also Lemma 4.2.6 and Theorem 4.5.5 in [18]), it follows that under **(H1)**, there exists a unique (in law) determinantal point process with integral operator $\mathcal{K}$. In this survey, we shall only consider determinantal point processes with Hermitian kernel. However, we mention that many important examples of determinantal point processes exhibit a non-Hermitian kernel, see [2–6, 24, 35].

Let us now recall the following result from, e.g., [32] (see Theorem 3.6 therein) that gives the Laplace transform of η.

Theorem 1 *Let $\mathcal{K}$ be an operator satisfying* **(H1)** *and η the determinantal point process with kernel K. Then η has Laplace transform*

$$\mathcal{L}_\eta(f) = \operatorname{Det}\left(\mathrm{Id} - \mathcal{K}_\varphi\right),$$

for each nonnegative f on $\mathbb{X}$ with compact support, where $\varphi = 1 - e^{-f}$ and $\mathcal{K}_\varphi$ is the trace class integral operator with kernel

$$K_\varphi(x, y) = \sqrt{\varphi(x)}K(x, y)\sqrt{\varphi(y)}, \quad x, y \in \mathbb{X}.$$

Let $\mathcal{K}$ be an operator satisfying assumption **(H1)**. We define the operators on $L^2(\mathbb{X}, \mu)$:

$$\mathcal{J} := (\mathrm{Id} - \mathcal{K})^{-1}\mathcal{K}, \tag{12}$$

and

$$\mathcal{J}[\Lambda] := (\mathrm{Id} - \mathcal{K}_\Lambda)^{-1}\mathcal{K}_\Lambda, \tag{13}$$

where Λ is a compact subset of $\mathbb{X}$. The operator $\mathcal{J}$ is called *global interaction operator*, and the operator $\mathcal{J}[\Lambda]$ is called *local interaction operator*. We emphasize that, unlike $\mathcal{K}_\Lambda$, $\mathcal{J}[\Lambda]$ is not a projection operator, i.e., in general $\mathcal{J}[\Lambda] \neq \mathcal{P}_\Lambda \mathcal{J} \mathcal{P}_\Lambda$. In any case, $\mathcal{J}[\Lambda]$ has some notable properties, as proved in [15]. First, it is easily seen that $\mathcal{J}[\Lambda]$ exists as a bounded operator and its spectrum is included in $[0, +\infty)$. Second, $\mathcal{J}[\Lambda]$ is also an integral operator, and we denote by $J[\Lambda]$ its kernel (in fact, one can even show that $\mathcal{J}[\Lambda]$ is a Carleman operator, cfr. the beginning of Sect. 3 in [15]). Third, $\mathcal{J}[\Lambda]$ is a trace class operator. Finally, by (9) we have

$$J[\Lambda](x, y) = \sum_{n \geq 1} \frac{\mu_n^\Lambda}{1 - \mu_n^\Lambda} \varphi_n^\Lambda(x)\overline{\varphi_n^\Lambda(y)},$$

for $x, y \in \Lambda$.
For $\alpha = \{x_1, \ldots, x_k\} \in \mathbf{N}_s(\Lambda)$, we denote by $\det J[\Lambda](\alpha)$ the determinant

$$\det\left(J[\Lambda](x_i, x_j)\right)_{1 \leq i,j \leq k}.$$

Note that for all $k \in \mathbb{N}^*$, the function

$$(x_1, \ldots, x_k) \mapsto \det J[\Lambda](\{x_1, \ldots, x_k\})$$

is $\mu^{\otimes k}$-a.e. nonnegative (thanks to Proposition 2) and symmetric in $x_1, \ldots, x_k$ (see, e.g., the Appendix of [15]), and we simply write $\det J[\Lambda](\{x_1, \ldots, x_k\}) = \det J[\Lambda](x_1, \ldots, x_k)$. The relevance of the local interaction operator becomes clear when computing the Janossy densities of the determinantal point process. More precisely, the following proposition holds.

Proposition 3 (Lemma 3.3 of [32]) *Let $\mathcal{K}$ be an operator satisfying* (**H1**) *and η the determinantal point process with kernel K. Then, for a compact subset $\Lambda \subseteq \mathbb{X}$ and $n \in \mathbb{N}^*$, the determinantal process η admits Janossy densities*

$$j^n_\Lambda(x_1, \ldots, x_n) = \mathrm{Det}(\mathrm{Id} - \mathcal{K}_\Lambda) \det J[\Lambda](x_1, \ldots, x_k), \tag{14}$$

for $x_1, \ldots, x_k \in \Lambda$. The void probability is equal to $j^0_\Lambda(\emptyset) = \mathrm{Det}(\mathrm{Id} - \mathcal{K}_\Lambda)$.

We emphasize that (14) still makes sense if $\|\mathcal{K}_\Lambda\|_{op} = 1$; indeed the zeros of $\mathrm{Det}(\mathrm{Id} - \mathcal{K}_\Lambda)$ are of the same order of the poles of $\det J[\Lambda](x_1, \ldots, x_k)$, see Lemma 3.4 of [32] for a more formal proof.
We now give some properties linking the rank of $\mathcal{K}$, $\mathrm{Rank}(\mathcal{K})$, and the number of points of the determinantal point process with integral operator $\mathcal{K}$.

Proposition 4 (Theorem 4 in [34], See also [18]) *Let $\mathcal{K}$ be an operator satisfying* (**H1**) *and η the determinantal point process with kernel K. We have:*

(a) The probability of the event that the number of points is finite is either 0 *or* 1*, depending on whether* $\mathrm{Tr}(\mathcal{K})$ *is finite or infinite. The number of points in a compact subset $\Lambda \subseteq \mathbb{X}$ is finite since* $\mathrm{Tr}(\mathcal{K}_\Lambda) < \infty$.
(b) The number of points is less than or equal to $n \in \mathbb{N}^*$ with probability 1 *if and only if $\mathcal{K}$ is a finite rank operator satisfying* $\mathrm{Rank}(\mathcal{K}) \leq n$.
(c) The number of points is $n \in \mathbb{N}^*$ with probability 1 *if and only if $\mathcal{K}$ is an orthogonal projection satisfying* $\mathrm{Rank}(\mathcal{K}) = n$.

We now give the Papangelou intensity of determinantal point processes.

Theorem 2 (Theorem 3.1 of [15]) *Let $\mathcal{K}$ be an operator satisfying* (**H1**) *and η the determinantal point process with kernel K. Then, for each compact set $\Lambda \subseteq \mathbb{X}$, η_Λ satisfies condition* (Σ) *(with μ_Λ in place of μ). A version of its compound Papangelou intensity $\hat{c}_\Lambda$ is given by*

$$\hat{c}_\Lambda(\alpha, \xi) = \frac{\det J[\Lambda](\alpha \cup \xi)}{\det J[\Lambda](\xi)}, \quad \alpha \in \mathbf{N}_s^f, \ \xi \in \mathbf{N}_s,$$

where the ratio is defined to be zero whenever the denominator vanishes. This version also satisfies the inequalities

$$\hat{c}_\Lambda(\alpha,\xi) \geq \hat{c}_\Lambda(\alpha,\xi'), \quad \text{and} \quad 0 \leq \hat{c}_\Lambda(\alpha,\xi) \leq \det J[\Lambda](\alpha) \leq \prod_{x\in\alpha} J[\Lambda](x,x), \tag{15}$$

whenever $\xi \subset \xi' \in \mathbf{N}_s(\Lambda)$ *and* $\alpha \in \mathbf{N}_s(\Lambda) \setminus \omega$.

Let $\mathcal{K}$ be an operator satisfying **(H1)** and let η be the determinantal point process with kernel K. Let $\mathcal{J}$ be the operator defined in (12). As proved in [15], $\mathcal{J}$ satisfies the following properties: it is locally of trace class and its kernel $(x, y) \mapsto J(x, y)$ can be chosen to satisfy Proposition 2. Moreover, η is stochastically dominated by a Poisson point process with mean measure $J(x, x)\,\mu(\mathrm{d}x)$ i.e., denoting by $\tilde{\mathbb{P}}$ the law of the Poisson process,

$$\int f\,\mathrm{d}\mathbb{P} \leq \int f\,\mathrm{d}\tilde{\mathbb{P}},$$

for all increasing measurable f. Here, we say that f is increasing if $f(\xi) \leq f(\xi')$ whenever $\xi \subset \xi' \in \mathbf{N}_s$.
We finally report the following theorem.

Theorem 3 (Theorem 3.6 in [15]) *Let* $\mathcal{K}$ *be an operator satisfying* **(H1)** *and* η *the determinantal point process with kernel K. Then* η *satisfies condition* (Σ)*, and its compound Papangelou intensity is given by*

$$\hat{c}(\alpha,\xi) = \lim_{n\to\infty} \hat{c}_{\Delta_n}(\alpha, \xi_{\Delta_n}), \quad \text{for } L^\mu \otimes \mathbb{P} - \text{almost every}(\alpha,\xi), \tag{16}$$

where $(\Delta_n)_{n\in\mathbb{N}}$ *is an increasing sequence of compact sets in* $\mathbb{X}$ *converging to* $\mathbb{X}$.

In general (16) does not give a closed form for the compound Papangelou intensity. In order to write $\hat{c}$ in closed form, additional hypotheses have to be assumed, see Proposition 3.9 in [15].

3 Integration by Parts

Hereafter we assume that $\mathbb{X}$ is a subset of $\mathbb{R}^d$, equipped with the Euclidean distance, μ is a Radon measure on $\mathbb{X}$ and $\Lambda \subseteq \mathbb{X}$ is a fixed compact set. We denote by $x^{(i)}$ the ith component of $x \in \mathbb{R}^d$.

3.1 Differential Calculus

We denote by $\mathcal{C}_c^\infty(\Lambda, \mathbb{R}^d)$ the set of all $\mathcal{C}^\infty$-vector fields $v : \Lambda \longrightarrow \mathbb{R}^d$ (with compact support) and by $\mathcal{C}_b^\infty(\Lambda^k)$ the set of all $\mathcal{C}^\infty$-functions on Λ^k whose derivatives are bounded.

Definition 1 A function $F : \mathbf{N}_s(\Lambda) \to \mathbb{R}$ is said to be in $\mathcal{S}_\Lambda$ if

$$F(\xi_\Lambda) = f_0 \mathbb{1}_{\{\xi(\Lambda)=0\}} + \sum_{k=1}^{n} \mathbb{1}_{\{\xi(\Lambda)=k\}} f_k(\xi_\Lambda), \tag{17}$$

for some integer $n \geq 1$, where for $k = 1, \ldots, n$, $f_k \in \mathcal{C}_b^\infty(\Lambda^k)$ is a symmetric function and $f_0 \in \mathbb{R}$ is a constant.

The gradient of $F \in \mathcal{S}_\Lambda$ of the form (17) is defined by

$$\nabla_x^{\mathbf{N}_s} F(\xi_\Lambda) := \sum_{k=1}^{n} \mathbb{1}_{\{\xi(\Lambda)=k\}} \sum_{y \in \xi_\Lambda} \mathbb{1}_{\{x=y\}} \nabla_x f_k(\xi_\Lambda), \qquad x \in \Lambda, \tag{18}$$

where ∇_x denotes the usual gradient on $\mathbb{R}^d$ with respect to the variable $x \in \Lambda$.
For $v \in \mathcal{C}_c^\infty(\Lambda, \mathbb{R}^d)$, we also let

$$\nabla_v^{\mathbf{N}_s} F(\xi_\Lambda) := \sum_{y \in \xi_\Lambda} \nabla_y^{\mathbf{N}_s} F(\xi_\Lambda) \cdot v(y) = \sum_{k=1}^{n} \mathbb{1}_{\{\xi(\Lambda)=k\}} \sum_{y \in \xi_\Lambda} \nabla_y f_k(\xi_\Lambda) \cdot v(y), \tag{19}$$

where $\cdot$ denotes the inner product on $\mathbb{R}^d$.
Next, we recall some notation from [1, 11]. Let $\mathrm{Diff}_0(\mathbb{X})$ be the set of all diffeomorphisms from $\mathbb{X}$ into itself with compact support, i.e., for any $\varphi \in \mathrm{Diff}_0(\mathbb{X})$, there exists a compact set outside of which φ is the identity map. In particular, note that $\mathrm{Diff}_0(\Lambda)$ is the set of diffeomorphisms from Λ into itself. In the following, μ_φ denotes the image measure of μ by φ.
Henceforth, we assume the following technical condition.

(H2) : *The Radon measure μ is absolutely continuous* w.r.t. *the Lebesgue measure ℓ on $\mathbb{X}$, with Radon–Nikodym derivative $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\ell}$ which is strictly positive and continuously differentiable on Λ.*

Then for any $\varphi \in \mathrm{Diff}_0(\Lambda)$, μ_φ is absolutely continuous with respect to μ with density given by

$$p_\varphi^\mu(x) = \frac{\mathrm{d}\mu_\varphi(x)}{\mathrm{d}\mu(x)} = \frac{\rho(\varphi^{-1}(x))}{\rho(x)} \mathrm{Jac}(\varphi^{-1})(x), \tag{20}$$

where $\mathrm{Jac}(\varphi^{-1})(x)$ is the Jacobian of φ^{-1} at point $x \in \mathbb{X}$. We are now in a position to give the quasi-invariance result, see [8, 11, 35].

Proposition 5 *Assume* **(H1)** *and* **(H2)** *and let η be the determinantal point process with kernel K. Then, for any measurable nonnegative f on Λ and any $\varphi \in \mathrm{Diff}_0(\Lambda)$,*

$$\mathbb{E}\left[\mathrm{e}^{-\langle f\circ\varphi,\,\eta\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{-\langle f-\ln(p^{\mu}_{\varphi}),\,\eta\rangle}\,\frac{\det J^{\varphi}[\Lambda](\eta)}{\det J[\Lambda](\eta)}\right]. \tag{21}$$

We point out that the right-hand side of (21) is well defined since $\det J[\Lambda] > 0$, $\mathbb{P}_{\Lambda}$-a.e.

3.2 Integration by Parts

Here we give an integration by parts formula on the set of test functionals $\mathscr{S}_{\Lambda}$ and an extension to closed gradients and divergence operators.
We start by introducing a further condition.

(H3) : *For any $n \geq 1$, the function*

$$(x_1,\ldots,x_n) \longmapsto \det J[\Lambda](x_1,\ldots,x_n)$$

is continuously differentiable on Λ^n.

Assuming **(H1)** and **(H3)**, we define the potential energy $U : \mathbf{N}_s(\Lambda) \longrightarrow \mathbb{R}$

$$U[\Lambda](\alpha) := -\log\det J[\Lambda](\alpha)$$

and its directional derivative along $v \in \mathcal{C}^{\infty}_c(\Lambda, \mathbb{R}^d)$

$$\begin{aligned}\nabla^{\mathbf{N}_s}_v U[\Lambda](\xi_\Lambda) &:= -\sum_{k=1}^{\infty} \mathbb{1}_{\{\xi(\Lambda)=k\}} \sum_{y\in\xi_\Lambda} \frac{\nabla_y \det J[\Lambda](\xi_\Lambda)}{\det J[\Lambda](\xi_\Lambda)} \cdot v(y) \\ &= \sum_{k=1}^{\infty} \mathbb{1}_{\{\xi(\Lambda)=k\}} \sum_{y\in\xi_\Lambda} U_{y,k}(\xi_\Lambda)\cdot v(y). \end{aligned} \tag{22}$$

The term $U_{y,k}$ in the previous definition is given by

$$U_{y,k}(\xi_\Lambda) := -\frac{\nabla_y \det J[\Lambda](\xi_\Lambda)}{\det J[\Lambda](\xi_\Lambda)} \quad \text{on } \{\xi(\Lambda) = k\}.$$

Under Condition (**H2**) we define

$$\beta^{\mu}(x) := \frac{\nabla \rho(x)}{\rho(x)},$$

and

$$B_v^{\mu}(\xi_\Lambda) := \sum_{y \in \xi_\Lambda} (-\beta^{\mu}(y) \cdot v(y) + \mathrm{div} v(y)), \qquad v \in \mathcal{C}_c^{\infty}(\Lambda, \mathbb{R}^d),$$

where div denotes the adjoint of the gradient ∇ on Λ, i.e., div verifies

$$\int_\Lambda g(x)\, \mathrm{div} \nabla f(x)\, \mathrm{d}x = \int_\Lambda \nabla f(x) \cdot \nabla g(x)\, \mathrm{d}x, \quad f, g \in \mathcal{C}^{\infty}(\Lambda).$$

The following integration by parts formula holds, see [11].

Lemma 1 *Assume* (**H1**)*,* (**H2**) *and* (**H3**)*, and let* η *be the determinantal point process with kernel* K*. Then, for any compact subset* $\Lambda \subseteq \mathbb{X}$*, any* $F, G \in \mathscr{S}_\Lambda$ *and vector field* $v \in \mathcal{C}_c^{\infty}(\Lambda, \mathbb{R}^d)$*, we have*

$$\mathbb{E}\left[G(\eta_\Lambda) \nabla_v^{\mathbf{N}_s} F(\eta_\Lambda)\right] = \mathbb{E}\left[F(\eta_\Lambda) \mathrm{div}_v^{\mathbf{N}_s} G(\eta_\Lambda)\right], \tag{23}$$

where

$$\mathrm{div}_v^{\mathbf{N}_s} G(\eta_\Lambda) := -\nabla_v^{\mathbf{N}_s} G(\eta_\Lambda) + G(\eta_\Lambda) \left(-B_v^{\mu}(\eta_\Lambda) + \nabla_v^{\mathbf{N}_s} U[\Lambda](\eta_\Lambda)\right).$$

Next, we extend the integration by parts formula by closability to a larger class of functionals. We refer to the appendix for the notion of closability. Let

$$L_\Lambda^2 := L^2(\mathbf{N}_s(\Lambda), \mathbb{P}_\Lambda)$$

be the space of square-integrable functions with respect to $\mathbb{P}_\Lambda$. It may be checked that $\mathscr{S}_\Lambda$ is dense in L_Λ^2.
For $v \in \mathcal{C}_c^{\infty}(\Lambda, \mathbb{R}^d)$, we consider the linear operators $\nabla_v^{\mathbf{N}_s} : \mathscr{S}_\Lambda \longrightarrow L_\Lambda^2$ and $\mathrm{div}_v^{\mathbf{N}_s} : \mathscr{S}_\Lambda \longrightarrow L_\Lambda^2$ defined, respectively, by $F \mapsto \nabla_v^{\mathbf{N}_s} F$ and $F \mapsto \mathrm{div}_v^{\mathbf{N}_s} F$. The following theorem is proved in [11].

Theorem 4 *Assume* (**H1**)*,* (**H2**)*,* (**H3**) *and*

$$\int_{\Lambda^n} \left| \frac{\partial_{x_i^{(h)}} \det J[\Lambda](x_1, \ldots, x_n) \partial_{x_j^{(k)}} \det J[\Lambda](x_1, \ldots, x_n)}{\det J[\Lambda](x_1, \ldots, x_n)} \right| \mathbb{1}_{\{\det J[\Lambda](x_1, \ldots, x_n) > 0\}}\, \mu(\mathrm{d}x_1) \cdots \mu(\mathrm{d}x_n) < \infty \tag{24}$$

for any $n \geq 1$*,* $1 \leq i, j \leq n$*, and* $1 \leq h, k \leq d$*. Then* (*i*) *For any vector field* $v \in \mathcal{C}_c^\infty(\Lambda, \mathbb{R}^d)$*, the linear operators* $\nabla_v^{\mathbf{N}_s}$ *and* $\mathrm{div}_v^{\mathbf{N}_s}$ *are well defined and closable. In particular, we have*

$$\nabla_v^{\mathbf{N}_s}(\mathscr{S}_\Lambda) \subset L_\Lambda^2 \qquad \text{and} \qquad \mathrm{div}_v^{\mathbf{N}_s}(\mathscr{S}_\Lambda) \subset L_\Lambda^2.$$

(*ii*) *Denoting by* $\overline{\nabla_v^{\mathbf{N}_s}}$ *(respectively* $\overline{\mathrm{div}_v^{\mathbf{N}_s}}$*) the minimal closed extension of* $\nabla_v^{\mathbf{N}_s}$ *(respectively* $\mathrm{div}_v^{\mathbf{N}_s}$*), for any vector field* $v \in \mathcal{C}_c^\infty(\Lambda, \mathbb{R}^d)$*, we have*

$$\mathbb{E}\left[G(\eta_\Lambda)\overline{\nabla_v^{\mathbf{N}_s}}F(\eta_\Lambda)\right] = \mathbb{E}\left[F(\eta_\Lambda)\overline{\mathrm{div}_v^{\mathbf{N}_s}}G(\eta_\Lambda)\right],$$

for all $F \in \mathrm{dom}\left(\overline{\nabla_v^{\mathbf{N}_s}}\right)$*,* $G \in \mathrm{dom}\left(\overline{\mathrm{div}_v^{\mathbf{N}_s}}\right)$.

Note that under the assumptions **(H1)**, **(H2)** and **(H3)**, condition (24) is satisfied if, for any $n \geq 1$, the function

$$(x_1, \ldots, x_n) \longmapsto \det J[\Lambda](x_1, \ldots, x_n),$$

is strictly positive on the compact Λ^n.

4 Stochastic Dynamics

4.1 Dirichlet Forms

Assume **(H1)**, and let η be the determinantal point process with kernel K. We consider the bilinear map $\mathcal{E}$ defined on $\mathscr{S}_\Lambda \times \mathscr{S}_\Lambda$ by

$$\mathcal{E}(F, G) := \mathbb{E}\left[\sum_{y \in \eta_\Lambda} \nabla_y^{\mathbf{N}_s} F(\eta_\Lambda) \cdot \nabla_y^{\mathbf{N}_s} G(\eta_\Lambda)\right]. \tag{25}$$

For $F \in \mathscr{S}_\Lambda$ of the form (17), i.e.,

$$F(\xi_\Lambda) = f_0 \mathbb{1}_{\{\xi(\Lambda)=0\}} + \sum_{k=1}^n \mathbb{1}_{\{\xi(\Lambda)=k\}} f_k(\xi_\Lambda),$$

we also define the Laplacian $\mathcal{H}$ by

$$\mathcal{H}F(\xi_\Lambda) = \sum_{k=1}^n \mathbb{1}_{\{\xi(\Lambda)=k\}} \sum_{y \in \xi_\Lambda} \left(-\beta^\mu(y) \cdot \nabla_y f_k(\xi_\Lambda) - \Delta_y f_k(\xi_\Lambda) + U_{y,k}(\xi_\Lambda) \cdot \nabla_y f_k(\xi_\Lambda)\right),$$

where $\Delta = -\mathrm{div}\nabla$ denotes the Laplacian operator on $\mathbb{R}^d$.

In the following, we consider the subspace $\tilde{\mathscr{S}}_\Lambda$ of $\mathscr{S}_\Lambda$ consisting of functions $F \in \mathscr{S}_\Lambda$ of the form

$$F(\xi_\Lambda) = f\left(\langle\varphi_1, \xi_\Lambda\rangle, \ldots, \langle\varphi_M, \xi_\Lambda\rangle\right) \mathbb{1}_{\{\xi(\Lambda)\leq K\}},$$

for some integers $M, K \geq 1$, $\varphi_1, \ldots, \varphi_M \in \mathcal{C}^\infty(\Lambda)$, $f \in \mathcal{C}_b^\infty(\mathbb{R}^M)$. Note that $\tilde{\mathscr{S}}_\Lambda$ is dense in L^2_Λ (see, e.g., [25, p. 54]).
Theorem 5 below is proved in [11]. We refer the reader to the appendix for the required notions of Dirichlet forms theory.

Theorem 5 *Under the assumptions of Theorem 4, we have*

(i) *The linear operator* $\mathcal{H} : \tilde{\mathscr{S}}_\Lambda \longrightarrow L^2_\Lambda$ *is symmetric, nonnegative definite, and well defined, i.e.,* $\mathcal{H}(\tilde{\mathscr{S}}_\Lambda) \subset L^2_\Lambda$*. In particular the operator square root* $\mathcal{H}^{1/2}$ *of* $\mathcal{H}$ *exists.*
(ii) *The bilinear form* $\mathcal{E} : \tilde{\mathscr{S}}_\Lambda \times \tilde{\mathscr{S}}_\Lambda \longrightarrow \mathbb{R}$ *is symmetric, nonnegative definite, and well defined, i.e.,* $\mathcal{E}(\tilde{\mathscr{S}}_\Lambda \times \tilde{\mathscr{S}}_\Lambda) \subset \mathbb{R}$.
(iii) $\mathcal{H}^{1/2}$ *and* $\mathcal{E}$ *are closable and the following relation holds:*

$$\overline{\mathcal{E}}(F, G) = \mathbb{E}[\overline{\mathcal{H}^{1/2}}\, F(\eta_\Lambda)\, \overline{\mathcal{H}^{1/2}}\, G(\eta_\Lambda)], \quad \forall\, F, G \in \mathrm{dom}(\overline{\mathcal{H}^{1/2}}). \tag{26}$$

(iv) *The bilinear form* $(\overline{\mathcal{E}}, \mathrm{dom}(\overline{\mathcal{H}^{1/2}}))$ *is a symmetric Dirichlet form.*

4.2 Associated Diffusion Processes

We start recalling some notions, see Chaps. IV and V in [25]. We call $\mathbf{N}$ the space of $\mathbb{N}$-valued Radon measures on $\mathbb{X}$, as opposed to $\mathbf{N}_s$ the space of *simple* $\mathbb{N}$-valued Radon measures on $\mathbb{X}$. We denote by $\mathbf{N}(\Lambda)$ the space of $\mathbb{N}$-valued Radon measures supported on a compact $\Lambda \subseteq \mathbb{X}$. We equip $\mathbf{N}$ with the vague topology and denote by $\mathscr{N}$ the corresponding Borel σ-algebra and by $\mathscr{N}(\Lambda)$ the corresponding trace-σ-algebra. Given π in the set $\mathrm{P}(\mathbf{N}(\Lambda))$ of the probability measures on $(\mathbf{N}(\Lambda), \mathscr{N}(\Lambda))$, we call a π-stochastic process with state space $\mathbf{N}(\Lambda)$ the collection

$$\mathbf{M}_{\Lambda,\pi} = (\boldsymbol{\Omega}, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, (\mathbf{M}_t)_{t\geq 0}, (\mathbf{P}_\xi)_{\xi\in\mathbf{N}(\Lambda)}, \mathbf{P}_\pi),$$

where $\mathcal{A} := \bigvee_{t\geq 0} \mathcal{A}_t$ is a σ-algebra on the set $\boldsymbol{\Omega}$, $(\mathcal{A}_t)_{t\geq 0}$ is the $\mathbf{P}_\pi$-completed filtration generated by the process $\mathbf{M}_t : \boldsymbol{\Omega} \longrightarrow \mathbf{N}(\Lambda)$, $\mathbf{P}_\xi$ is a probability measure on $(\boldsymbol{\Omega}, \mathcal{A})$ for all $\xi \in \mathbf{N}(\Lambda)$, and $\mathbf{P}_\pi$ is the probability measure on $(\boldsymbol{\Omega}, \mathcal{A})$ defined by

$$\mathbf{P}_\pi(A) := \int_{\mathbf{N}(\Lambda)} \mathbf{P}_\xi(A)\, \pi(\mathrm{d}\xi), \quad A \in \mathcal{A}.$$

A collection $(\mathbf{M}_{\Lambda,\pi}, (\theta_t)_{t\geq 0})$ is called a π-time homogeneous Markov process with state space $\mathbf{N}(\Lambda)$ if $\theta_t : \boldsymbol{\Omega} \longrightarrow \boldsymbol{\Omega}$ is a shift operator, i.e., $\mathbf{M}_s \circ \theta_t = \mathbf{M}_{s+t}$, $s, t \geq 0$, the map $\xi \mapsto \mathbf{P}_\xi(A)$ is measurable for all $A \in \mathcal{A}$, and the time homogeneous Markov property

$$\mathbf{P}_\xi(\mathbf{M}_t \in A \,|\, \mathcal{A}_s) = \mathbf{P}_{\mathbf{M}_s}(\mathbf{M}_{t-s} \in A),\ \mathbf{P}_\xi\text{-a.s.},\ A \in \mathcal{A},\ 0 \leq s \leq t,\ \xi \in \mathbf{N}(\Lambda),$$

holds. Recall that a π-time homogeneous Markov process $(\mathbf{M}_{\Lambda,\pi}, (\theta_t)_{t\geq 0})$ with state space $\mathbf{N}(\Lambda)$ is said to be π-tight on $\mathbf{N}(\Lambda)$ if $(\mathbf{M}_t)_{t\geq 0}$ is right-continuous with left limits $\mathbf{P}_\pi$-almost surely; $\mathbf{P}_\xi(\mathbf{M}_0 = \xi) = 1\ \forall \xi \in \mathbf{N}(\Lambda)$; the filtration $(\mathcal{A}_t)_{t\geq 0}$ is right continuous; the following strong Markov property holds:

$$\mathbf{P}_{\pi'}(\mathbf{M}_{t+\tau} \in A \,|\, \mathcal{A}_\tau) = \mathbf{P}_{\mathbf{M}_\tau}(\mathbf{M}_t \in A)$$

$\mathbf{P}_{\pi'}$-almost surely for all $\mathcal{A}_t$-stopping time τ, $\pi' \in \mathrm{P}(\mathbf{N}(\Lambda))$, $A \in \mathcal{A}$ and $t \geq 0$, cfr. Theorem IV.1.15 in [25]. In addition, a π-tight process on $\mathbf{N}(\Lambda)$ is said to be a π-special standard process on $\mathbf{N}(\Lambda)$ if for any $\pi' \in \mathrm{P}(\mathbf{N}(\Lambda))$ which is equivalent to π and all $\mathcal{A}_t$-stopping times τ, $(\tau_n)_{n\geq 1}$ such that $\tau_n \uparrow \tau$ we have that $\mathbf{M}_{\tau_n}$ converges to $\mathbf{M}_\tau$, $\mathbf{P}_{\pi'}$-almost surely.
The following theorem is proved in [11]. Therein $\mathbf{E}_\xi$ denotes the expectation under $\mathbf{P}_\xi$, $\xi \in \mathbf{N}(\Lambda)$. Here again, we refer the reader to the appendix for the required notions of Dirichlet forms theory.

Theorem 6 *Assume the hypotheses of Theorem 4, let $\mathbb{P}$ be the law of a determinantal point process η with kernel K, and $\overline{\mathcal{E}}$ be the Dirichlet form constructed in Theorem 5. Then there exists a $\mathbb{P}_\Lambda$-tight special standard process $(\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}, (\theta_t)_{t\geq 0})$ on $\mathbf{N}(\Lambda)$ such that:*

1. *$\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}$ is a diffusion, in the sense that:*

$$\mathbf{P}_\xi(\{\omega :\ t \mapsto \mathbf{M}_t(\omega) \text{ is continuous on } [0,+\infty)\}) = 1, \qquad \overline{\mathcal{E}}\text{-a.e. } \xi \in \mathbf{N}(\Lambda). \tag{27}$$

2. *The transition semigroup of $\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}$ is given by*

$$p_t F(\xi) := \mathbf{E}_\xi[F(\mathbf{M}_t)], \quad \xi \in \mathbf{N}(\Lambda), \quad F : \mathbf{N}(\Lambda) \longrightarrow \mathbb{R} \quad \text{square integrable},$$

   and it is properly associated with the Dirichlet form $(\overline{\mathcal{E}}, \mathrm{dom}(\overline{\mathcal{H}^{1/2}}))$, i.e., $p_t F$ is an $\overline{\mathcal{E}}$-a.c., $\mathbb{P}_\Lambda$-version of $\exp(-t\mathcal{H}^{\mathrm{gen}}_\Lambda)F$, for all square integrable $F : \mathbf{N}(\Lambda) \longrightarrow \mathbb{R}$ and $t > 0$ (where $\mathcal{H}^{\mathrm{gen}}_\Lambda$ is the generator of $\overline{\mathcal{E}}$).
3. *$\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}$ is unique up to $\mathbb{P}_\Lambda$-equivalence (we refer the reader to Definition 6.3 page 140 in [26] for the meaning of this notion).*
4. *$\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}$ is $\mathbb{P}_\Lambda$-symmetric, i.e.,*

$$\mathbb{E}\left[G(\eta_\Lambda)\, p_t F(\eta_\Lambda)\right] = \mathbb{E}\left[F(\eta_\Lambda)\, p_t G(\eta_\Lambda)\right],$$

   for square integrable functions F and G on $\mathbf{N}(\Lambda)$.

5. $\mathbf{M}_{\Lambda,\mathbb{P}_\Lambda}$ *has* $\mathbb{P}_\Lambda$ *as invariant measure.*

In dimension $d \geq 2$, the diffusion constructed in the previous theorem is non-colliding. Indeed, the following theorem holds.

Theorem 7 *Assume* $d \geq 2$*, and the hypotheses of Theorem 4. Then*

$$\mathbf{P}_\xi(\{\omega \in \boldsymbol{\Omega} : \mathbf{M}_t(\omega) \in \mathbf{N}_s(\Lambda), \text{ for any } t \in [0,\infty)\}) = 1, \qquad \overline{\mathcal{E}}\text{-a.e. } \xi \in \mathbf{N}_s(\Lambda).$$

4.3 An Illustrative Example

Let $\Lambda := B(0,R) \subset \mathbb{R}^2$ be the closed ball centered at the origin with radius $R \in (0,1)$, let $\{\varphi_k^{(R)}\}_{1\leq k\leq 3}$, denote the orthonormal subset of $L^2(B(0,R),\ell)$ defined by

$$\varphi_k^{(R)}(x) := \frac{1}{R}\sqrt{\frac{k+1}{\pi}}\left(\frac{x^{(1)}}{R} + i\frac{x^{(2)}}{R}\right)^k, \quad x = (x^{(1)}, x^{(2)}) \in B(0,R),\ k = 1,2,3,$$

where $\mu = \ell$ is the Lebesgue measure on $\mathbb{R}^2$ and $i := \sqrt{-1}$ denotes the complex unit. We consider the truncated Bergman kernel (see [18]) restricted to Λ

$$K_{\mathrm{Be}}(x,y) := \sum_{k=1}^{3} R^{2(k+1)}\varphi_k^{(R)}(x)\overline{\varphi_k^{(R)}(y)}, \qquad x,y \in B(0,R),$$

and denote by $\mathcal{K}_{\mathrm{Be}}$ the associated integral operator.

We now discuss the conditions of Theorem 6. First, K_{Be} is readily seen to be Hermitian and locally of trace class with nonzero eigenvalues $\kappa_k := R^{2(k+1)}$, $k = 1,2,3$. As a consequence, the spectrum of $\mathcal{K}_{\mathrm{Be}}$ is contained in $[0,1)$ and the triplet $(\mathcal{K}_{\mathrm{Be}}, K_{\mathrm{Be}}, \ell)$ satisfies assumption (**H1**). In addition, Condition (**H2**) is trivially satisfied since $\mu = \ell$ is the Lebesgue measure.

Denoting by η_Λ the determinantal point process with kernel K_{Be}, the Janossy densities of η_Λ are given by

$$j_\Lambda^{(k)}(x_1,\dots,x_k) = \mathrm{Det}\,(\mathrm{Id} - \mathcal{K}_{\mathrm{Be}})\det J[\Lambda](x_1,\dots,x_k),$$

for $k = 1,2,3,\ (x_1,\dots,x_k) \in \Lambda^k$, and where the kernel $J[\Lambda]$ of $\mathcal{J}[\Lambda]$ is given by

$$J[\Lambda](x,y) := \sum_{h=1}^{3} \frac{R^{2(h+1)}}{1 - R^{2(h+1)}}\varphi_h^{(R)}(x)\overline{\varphi_h^{(R)}(y)}.$$

Moreover, η_Λ has at most 3 points according to Proposition 4, which means that $j^k_\Lambda = 0$, for $k \geq 4$. To prove condition **(H3)** it suffices to remark that the function

$$(x_1, \dots, x_k) \to \det(J[\Lambda](x_p, x_q))_{1\leq p,q\leq k}$$

is continuously differentiable on Λ^k, for $k \leq 3$. Condition (24) is trivially satisfied for $k > 3$ since as already observed in this case $j^k_\Lambda = 0$. Next, we check that Condition (24) is verified for $k = 3$. To that end, note that

$$J[\Lambda](x_1, x_2, x_3) = A(x_1, x_2, x_3)A(x_1, x_2, x_3)^*,$$

where the matrix $A := (A_{ph})_{1\leq p,h\leq 3}$ is given by

$$A_{ph} := \frac{R^{h+1}}{\sqrt{1 - R^{2(h+1)}}}\varphi_h^{(R)}(x_p)$$

and $A(x_1, x_2, x_3)^*$ denotes the transpose conjugate of $A(x_1, x_2, x_3)$. Hence,

$$\det J[\Lambda](x_1, x_2, x_3) = |\det A(x_1, x_2, x_3)|^2,$$

and since the previous determinant is a Vandermonde determinant, we have

$$\det A(x_1, x_2, x_3) = \prod_{p=1}^{3} \sqrt{\frac{1+p}{\pi(1 - R^{2(p+1)})}} \left(\prod_{p=1}^{3} (x_p^{(1)} + i x_p^{(2)}) \right) \prod_{1\leq p<q\leq 3} ((x_p^{(1)} - x_q^{(1)}) + i(x_p^{(2)} - x_q^{(2)})).$$

So, Condition (24) with $k = 3$ reduces to

$$\int_{B(0,R)^3} \left| \frac{\partial_{x_i^{(h)}} |\det A(x_1, x_2, x_3)|^2 \partial_{x_j^{(k)}} |\det A(x_1, x_2, x_3)|^2}{|\det A(x_1, x_2, x_3)|^2} \right| \ell(\mathrm{d}x_1)\ell(\mathrm{d}x_2)\ell(\mathrm{d}x_3) < \infty,$$

for all $1 \leq i, j \leq 3$ and $1 \leq h, k \leq 2$, and for this it suffices to check

$$\int_{B(0,R)^3} \left| \frac{\partial_{x_1^{(1)}} |\det A(x_1, x_2, x_3)|^2}{|\det A(x_1, x_2, x_3)|^2} \right| \ell(\mathrm{d}x_1)\ell(\mathrm{d}x_2)\ell(\mathrm{d}x_3) < \infty.$$

This latter integral can be written as

$$\int_{B(0,R)^3} \left| \frac{2x_1^{(1)}}{(x_1^{(1)})^2 + (x_1^{(2)})^2} + 2\sum_{j=2}^{3} \frac{x_1^{(1)} - x_j^{(1)}}{(x_1^{(1)} - x_j^{(1)})^2 + (x_1^{(2)} - x_j^{(2)})^2} \right| \ell(\mathrm{d}x_1)\ell(\mathrm{d}x_2)\ell(\mathrm{d}x_3),$$

which is indeed finite. Condition (24) may be verified also for $k < 3$ by taking into account some properties of generalized Vandermonde determinants, we refer the reader to [11] for the details. Consequently, by Theorem 6 we have the existence of a diffusion process properly associated with the determinantal point process with the Bergman-type kernel K_{Be}.

5 Simulation

5.1 Standard Simulation of Determinantal Point Processes

In this section, we describe the standard algorithm to sample from the law of a determinantal point process. The main results of this section can be found in the seminal work of Hough et al. [17], along with the improvements found in [12, 18, 22]. We recall the algorithm introduced there in order to insist on its advantages and disadvantages compared to directly simulating according to the densities. The standard algorithm first yields a way to simulate the number of points $n \in \mathbb{N}$ of a determinantal point process on a given compact $\Lambda \subseteq \mathbb{X}$. Second, it provides a sample from the Janossy density j_Λ^n. Let us now discuss in detail these two steps.

Theorem 8 *Let $\mathcal{K}$ be a trace class integral operator satisfying* **(H1)** *(we often take $\mathcal{K}_\Lambda$, which is indeed of trace class), $\{\varphi_n\}_{n\geq 1}$ an orthonormal basis of $L^2(\mathbb{X}, \mu)$ formed by eigenfunctions of $\mathcal{K}$ and $\{\mu_n\}_{n\geq 1}$ the corresponding sequence of eigenvalues. We write*

$$K(x, y) = \sum_{n\geq 1} \mu_n \varphi_n(x)\overline{\varphi_n(y)}, \quad x, y \in \mathbb{X}. \tag{28}$$

Let $\{B_n\}_{n\geq 1}$ be a sequence of independent Bernoulli random variables of mean $\mathbb{E}[B_n] = \mu_n$. The Bernoulli random variables are defined on a distinct probability space, say $(\Omega, \mathcal{F})$. Then, define the (random) kernel

$$K_B(x, y) = \sum_{n\geq 1} B_n \varphi_n(x)\overline{\varphi_n(y)}, \quad x, y \in \mathbb{X}.$$

Finally, define the point process η on $(\mathbf{N}_s \times \Omega, \mathcal{N}_s \otimes \mathcal{F})$ as the point process obtained by first drawing the Bernoulli random variables, and then the point process with kernel K_B. We have that η is a determinantal point process on $\mathbb{X}$ with kernel K.

For the remainder of this paragraph, we consider a general kernel K of the form (28) and wish to generate a sample of the determinantal point process with kernel $\mathcal{K}$.

According to Theorem 8, the number of points on $\mathbb{X}$ is distributed as the sum of independent Bernoulli random variables. More precisely,

$$|\xi(\mathbb{X})| \sim \sum_{n\geq 1} B_n,$$

where $B_n \sim \text{Be}(\mu_n)$, $n \in \mathbb{N}$. Define $T := \sup\{n \geq 1 \ / \ B_n = 1\}$. Since $\sum_{n\geq 1} \mu_n = \sum_{n\geq 1} \mathbb{P}(B_n = 1) < \infty$, by a direct application of the Borel–Cantelli lemma, we have that $T < \infty$ almost surely. Hence the method is to simulate first a realization of T, say t, and then $t-1$ independent Bernoulli random variables $B_1, \ldots, B_{t-1}$, each B_n with mean μ_n, $n = 1, \ldots, t-1$ Finally, set $B_t = 1$.
The simulation of the random variable T can be obtained by the inversion method, as we know its cumulative distribution function explicitly. Indeed, for $n \in \mathbb{N}$,

$$\mathbb{P}(T = n) = \mu_n \prod_{i=n+1}^{\infty} (1 - \mu_i),$$

hence

$$F(r) = \mathbb{P}(T \leq r) = \sum_{n\leq r} \mu_n \prod_{i=n+1}^{\infty} (1 - \mu_i), \quad \forall r \in \mathbb{N}. \tag{29}$$

To generate a random variable with law F requires the numerical computation of the generalized inverse $F^{-1}(u) := \inf\{t \in \mathbb{N} \ / \ F(t) \geq u\}$. In many practical cases, as in the case of the Ginibre point process, the numerical calculations may augment the complexity of the algorithm and the CPU. This is the main reason for which we shall propose an approximate simulation of the Ginibre point process.

Assume we have simulated the number of points of the determinantal point process on a compact Λ. For the clarity, we suppose $T = n$ and $B_1 = 1, B_2 = 1, \ldots, B_n = 1$. This assumption is equivalent to a simple reordering of the eigenvectors $(\varphi_n)_{n\in\mathbb{N}}$. Then we have reduced the problem to that of simulating the vector $(X_1, \ldots, X_n)$ of joint density

$$p(x_1, \ldots, x_n) = \frac{1}{n!} \det\left(\tilde{K}(x_i, x_j)\right)_{1\leq i,j\leq n},$$

where $\tilde{K}(x, y) = \sum_{j=1}^{n} \psi_j(x)\overline{\psi_j(y)}$, for $x, y \in \Lambda$, where here $(\psi_j)_{j\in\mathbb{N}}$ is the reordering of $(\varphi_j)_{j\in\mathbb{N}}$. The determinantal point process of kernel $\tilde{K}$ has n points almost surely by Proposition 4, which means that it remains to simulate the unordered vector $(X_1, \ldots, X_n)$ of points of the point process. The idea of the algorithm is to start by simulating X_1, then $X_2|X_1$, until $X_n|X_1, \ldots, X_{n-1}$. The key here is that in the determinantal case, the density of these conditional probabilities takes a computable form. Let us start by observing that

$$\det\left(\tilde{K}(x_i,x_j)\right)_{1\le i,j\le n_i} = \det\left(\psi_k(x_l)\right)_{1\le k,l\le n_i} \det\left(\overline{\psi_l(x_k)}\right)_{1\le k,l\le n},$$

so the density of X_1 on Λ is

$$\begin{aligned}
p_1(x_1) &= \int \dots \int p(x_1,\dots,x_n)\, \mu(\mathrm{d}x_2)\cdots\mu(\mathrm{d}x_n) \\
&= \frac{1}{n!}\sum_{\tau,\sigma\in S_n} \mathrm{sgn}(\tau)\mathrm{sgn}(\sigma)\psi_{\tau(1)}(x_1)\overline{\psi_{\sigma(1)}(x_1)}\prod_{k=2}^{n}\int \psi_{\tau(k)}(x_k)\overline{\psi_{\sigma(k)}(x_k)}\,\mu(\mathrm{d}x_k) \\
&= \frac{1}{n!}\sum_{\sigma\in S_n}|\psi_{\sigma(1)}(x_1)|^2 \\
&= \frac{1}{n}\sum_{k=1}^{n}|\psi_k(x_1)|^2,
\end{aligned}$$

where S_n is the n-th symmetric group and $\mathrm{sgn}(\sigma)$ is the sign of the permutation $\sigma \in S_n$. By a similar computation, we may compute the distribution of $X_2|X_1$, whose density with respect to μ is given by

$$\begin{aligned}
p_{X_2|X_1}(x_2) = \frac{p_2(X_1,x_2)}{p_1(X_1)} &= \frac{1}{(n-1)!\sum|\psi_j(X_1)|^2} \\
&\sum_{\sigma\in S_n}\left(|\psi_{\sigma(1)}(X_1)|^2|\psi_{\sigma(2)}(x_2)|^2 - \psi_{\sigma(1)}(X_1)\overline{\psi_{\sigma(2)}}(X_1)\psi_{\sigma(2)}(x_2)\overline{\psi_{\sigma(1)}}(x_2)\right) \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}|\psi_j(x_2)|^2 - |\sum_{j=1}^{n}\frac{\psi_j(X_1)}{\sqrt{\sum|\psi_j(X_1)|^2}}\overline{\psi_j}(x_2)|^2\right).
\end{aligned}$$

The previous formula can be generalized recursively and has the advantage of giving a natural interpretation of the conditional densities. Indeed, we may write the conditional densities at each step in a way that makes the orthogonalization procedure appear. This is presented in the final algorithm, which was explicitly obtained in [22] (see also [17] for the proof). We define the vector $\mathbf{v}(x) := (\psi_1(x),\dots,\psi_n(x))^t$, where t stands for the transpose operator, denote by $\|\mathbf{v}(x)\|$ its Euclidean norm, and given $\mathbf{x}\in\mathbb{C}^n$, we set $\mathbf{x}^* := \overline{\mathbf{x}}^t$.
It is then known that Algorithm 1 yields a sample $\{X_1,\dots,X_n\}$ of a determinantal point process with kernel $\tilde{K}(x,y) = \sum_{j=1}^{n}\psi_j(x)\overline{\psi_j(y)}$, $x,y\in\Lambda$.

Algorithm 1 Simulation of the determinantal projection point process

sample X_n from the distribution with density $p_n(x) = \|\mathbf{v}(x)\|^2/n,\ x \in \Lambda$
$\mathbf{e}_1 \leftarrow \mathbf{v}(X_n)/\|\mathbf{v}(X_n)\|$
for $j = n-1 \to 1$ **do**
  sample X_j from the distribution with density

$$p_j(x) = \frac{1}{j}\Big[\|\mathbf{v}(x)\|^2 - \sum_{k=1}^{n-j} |\mathbf{e}_k^* \mathbf{v}(x)|^2\Big]$$

  $\mathbf{w}_j \leftarrow \mathbf{v}(X_j) - \sum_{k=1}^{n-j} \big(\mathbf{e}_k^* \mathbf{v}(X_j)\big)\, \mathbf{e}_k, \quad \mathbf{e}_{n-j+1} \leftarrow \mathbf{w}_j/\|\mathbf{w}_j\|$
end for
return $(X_1, \ldots, X_n)$

5.2 Simulation Using Markov Chains

Exploiting the bound (15), an alternative algorithm to sample from the law of a determinantal point process on a finite window is readily obtained by specializing the general theory developed in [19, 20, 23], which allow to sample from the law of finite point processes with bounded Papangelou intensity. Let us give a brief description.

In the remainder of this paragraph, we fix a compact set $\Lambda \subseteq \mathbb{X}$, and turn our attention to the simulation of a determinantal point process with kernel K_Λ. The following bound holds for the Papangelou conditional intensity c_Λ:

$$\forall x \in \Lambda,\ \forall \xi \in \mathbf{N}_s,\ c_\Lambda(x, \xi) \le J[\Lambda](x, x) = J(x, x), \tag{30}$$

where we have specialized the bound (15) and have noticed that $J[\Lambda](x, x) = J(x, x)$ for $x \in \Lambda$. We first simulate a Glauber process associated with the measure $J(x, x)\mathrm{d}\mu(x)$:

- Draw an initial configuration D_0 according to the distribution of a Poisson point process over Λ with mean measure $J(x, x)\mathrm{d}\mu(x)$.
- Define a Poisson process on $\mathbb{R}_+$ of intensity $M = \int_\Lambda J(x, x)\mathrm{d}\mu(x)$ and denote by $(T_n, n \ge 1)$ its arrival times.
- At each time T_n, a particle appears at a position randomly located according to the probability distribution $M^{-1} J(x, x)\mathrm{d}\mu(x)$ independently from any other event.
- To each particle, we assign an exponentially distributed lifetime of mean 1, independently from any other event, i.e., each particle dies after an exponential distributed time.
- The Glauber process D is formed by the random variables D_t denoting the number of particles alive at time t.

Once this process is constructed, we can use the coupling from the past to simulate the determinantal point process:

- Simulate a (dominating) Glauber process D corresponding to the mean measure $J(x,x)\mathrm{d}\mu(x)$ over Λ on a time horizon T, with initial configuration D_0 . Record all birth dates and locations along the sample-path.
- Define two configuration-valued Markov chains, L and U. L stands for *lower* and U for *upper* since we will guarantee $L_t \subset U_t \subset D_t$ at any time $t \geq 0$, $L_0 = \emptyset$ and $U_0 = D_0$.
- Read the time-line of the process D.

1. When there is a death in the sample-path of D, then the corresponding particle dies (in both U and L) provided it exists.
2. When there is a birth at x in D at time t, draw a uniform sample S on $[0, 1]$, independently from everything else. If $S \leq c(x, U_{t}-)$ then x is added to $L_t = L_{t}- \cup x$. If $S \leq c(x, L_{t}-)$, then $U_t = U_{t}- \cup \{x\}$.
3. If at time T, $U_T = L_T$ then U_T is a sample of the determinantal point process of Papangelou intensity c. If not, expand the sample-path of D to $[T, 2T]$ and replay the same algorithm.

A crucial question is then how to choose T to avoid both a too long simulation if T is large and the need to extend several times the sample-path of D if T is too small. A very crude bound on the coalescence time, i.e., the time at which U and L coincide, is the hitting time of the null configuration by D. Indeed, since for any time t, $L_t \subset U_t \subset D_t$, if $D_T = \emptyset$ then $U_T = \emptyset$ and $L_T = \emptyset$. It turns out that the number of points of D follow the dynamics of an M/M/∞ queue. If the initial population of D_0 is large then Proposition 6.8 of [30] entails that T_0 is of the order of $\log(|D_0|)$. This means that the coalescence time of our algorithm is an $O\left(\log \int_{\Lambda} J(x,x)\,\mathrm{d}\mu(x)\right)$, but in practice, we are well below this upper bound (Fig. 1).

Finally, we present some samples of the coalescence time in a practical example known as the Gaussian model (see [22]). More precisely, Fig. 2 shows the distributions of the coalescence time of L_t and U_t and the stopping time of the algorithm for 500 samples of the Gaussian model DPP with $\rho = 50$.

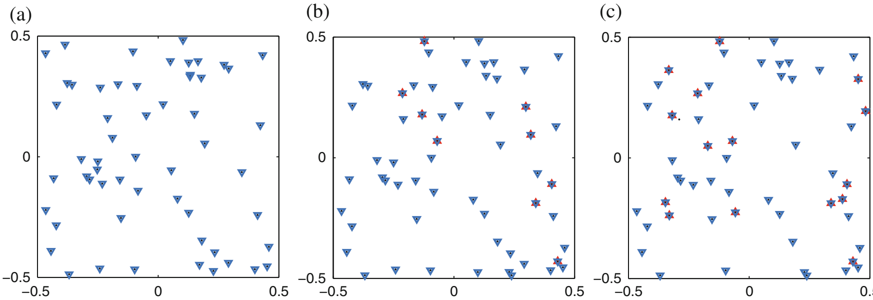

Fig. 1 CFTP simulations for Gaussian model DPP with $\rho = 50$ and $\alpha = 0.04$, respectively at time T_i, the i-th jump time from time $t = -n$. Notations: "·" := D_t, "∇" := U_t and "Δ"(*red*) := L_t. (**a**) at time T_0; (**b**) at time T_{25}; (**c**) at time T_{50}

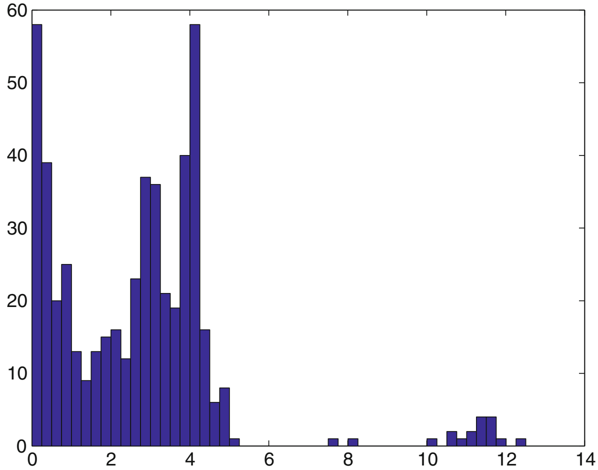

Fig. 2 Histogram of the coalescence time of L_t and U_t and the stopping time on 500 samples of a Gaussian model with $\rho = 50$ and $\alpha = 0.04$

The two simulation methods presented are conceptually quite different and are therefore difficult to compare. To be more precise, in the standard algorithm, there are two time-consuming steps: the simulation of the Bernoulli random variables and the simulation under the density p_i for which we are a priori required to proceed by rejection sampling. This requires an evaluation of the supremum of p_i on a grid which can be unboundedly big. In the algorithm based on Markov chains, we avoid the previous problem by only evaluating elaborate functionals (in our case, the Papangelou conditional intensity c) on a specific configuration, and not on the whole grid. Additionally, the standard algorithm relies on the knowledge of the eigenfunctions and the eigenvalues of the kernel K_Λ whereas the algorithm based on Markov chains works well with any expression of $J[\Lambda]$. However, the time necessary to reach equilibrium can be quite long, which is the main drawback of this algorithm. Thus, quantifying the execution time of the MCMC algorithm is of practical interest. We roughly discussed this question in this section, but a comparison with the standard algorithm is in general quantitatively difficult since the better performing algorithm depends on the kernel K_Λ of the underlying determinantal point process.

5.3 Approximate Simulation of the Ginibre Point Process

In this paragraph, we introduce a specific determinantal point process which is fast to simulate in practice, well suited for applications, and converges weakly to the Ginibre point process.

The Ginibre point process, see [16], is the determinantal point process on $\mathbb{C}$ with kernel

$$K_{\text{Gin}}(z_1, z_2) := \sum_{n=0}^{\infty} \varphi_n(z_1)\overline{\varphi_n(z_2)}, \quad z_1, z_2 \in \mathbb{C}, \tag{31}$$

where $\varphi_n(z) := \frac{1}{\sqrt{\pi n!}} e^{-\frac{1}{2}|z|^2} z^n$ for each $n \geq 0$. Further details concerning the Ginibre point process may be found in [18, 31].
We introduce a new kernel, by setting

$$K^N_{\text{Gin}}(z_1, z_2) := \sum_{n=0}^{N-1} \varphi_n^{\sqrt{N}}(z_1)\overline{\varphi_n^{\sqrt{N}}(z_2)}, \quad z_1, z_2 \in B(0, \sqrt{N}), \tag{32}$$

where we define $\varphi_n^{\sqrt{N}} := \frac{1}{\sqrt{\pi \gamma(n+1, N)}} e^{-\frac{1}{2}|z|^2} z^n \mathbb{1}_{\{z \in B(0,\sqrt{N})\}}$, for $0 \leq n \leq N-1$. Here, $\gamma(z, a) := \int_0^a e^{-t} t^{z-1}\, \mathrm{d}t$, $a \geq 0$, $z \in \mathbb{C}$ is the lower incomplete Gamma function. This kernel defines a determinantal point process named truncated Ginibre point process conditioned on having N points, see [12] for details. Clearly, this determinantal point process can be simulated as described by Algorithm 1. Fixing the number N of points in the ball $B(0, \sqrt{N})$ ensures a fast execution time.
As already noticed, Algorithm 1 yields a sample of the truncated Ginibre point process conditioned on having N points on the ball $B(0, \sqrt{N})$. In order to simulate the process on $B(0, a)$, $a \geq 0$, we need to apply a homothetic transformation to the N points, which translates to a homothety on the eigenfunctions. To summarize, the simulation algorithm for the truncated Ginibre process conditioned on having N points on the ball $B(0, a)$ is done according to Algorithm 2.

Algorithm 2 Simulation of the truncated Ginibre point process

define $\varphi_k(z) = \frac{N}{\pi a^2 \gamma(k+1, N)} e^{-\frac{N}{2a^2}|z|^2} (\frac{Nz}{a^2})^k$, for $z \in B(0, \sqrt{N})$ and $0 \leq k \leq N-1$.
define $\mathbf{v}(z) := (\varphi_0(z), \ldots, \varphi_{N-1}(z))$, for $z \in B(0, \sqrt{N})$.
sample X_N from the distribution with density $p_N(z) = \|\mathbf{v}(z)\|^2/N,\ z \in B(0, \sqrt{N})$
set $\mathbf{e}_1 = \mathbf{v}(X_N)/\|\mathbf{v}(X_N)\|$
for $i = N-1 \to 1$ **do**
  sample X_i from the distribution with density

$$p_i(x) = \frac{1}{i}\Big[\|\mathbf{v}(x)\|^2 - \sum_{j=1}^{N-i} |\mathbf{e}_j^* \mathbf{v}(x)|^2\Big]$$

  set $\mathbf{w}_i = \mathbf{v}(X_i) - \sum_{j=1}^{N-i} \left(\mathbf{e}_j^* \mathbf{v}(X_i)\right) \mathbf{e}_j, \quad \mathbf{e}_{N-i+1} = \mathbf{w}_i/\|\mathbf{w}_i\|$
end for
return $(X_1, \ldots, X_N)$

The next theorem from [12] and the subsequent comment guarantee that the above algorithm can be interpreted as an approximate simulation algorithm for the Ginibre point process.

Theorem 9 *The kernel K^N_{Gin} converges to K_{Gin}, as N tends to infinity, uniformly on compacts.*

As a consequence of Theorem 9 and Proposition 3.10 in [32], the truncated Ginibre point process conditioned on having N points converges weakly to the Ginibre point process.

6 Open Questions

We mention here a few open questions.

- Let $\mathbb{P}$ be the law of a determinantal point process η on $\mathbb{X}$, and φ a diffeomorphism of the whole space. Is the image of $\mathbb{P}$ by φ absolutely continuous with respect to $\mathbb{P}$? If yes, is it possible to compute the corresponding Radon–Nikodym derivative?
- Is the diffusion constructed in Theorem 6 ergodic?
- Consider a sequence of diffusions defined by Theorem 6 and indexed by compacts Λ_n increasing to $\mathbb{R}^d$. Does $\mathbf{M}_{\Lambda_n,\mathbb{P}_{\Lambda_n}}$ converge weakly to some limiting diffusion as $n \to \infty$? If yes, may we compute the properly associated Dirichlet form?
- Is it possible to approximate in distribution the diffusion constructed in Theorem 6 by a continuous-time Markov process (such as a Glauber dynamics)?
- What is the error committed by the approximate simulation algorithm to sample from the target law, i.e., the law of the Ginibre point process?
- Let η be a determinantal point process with integral operator $\mathcal{K}$. Can one generalize the results presented in this chapter to include the case where 1 is an eigenvalue of $\mathcal{K}$?

Acknowledgements The authors wish to thank Low Kah Choon who produced the graphs of Sect. 5 and who contributed to some heuristics in the section.

Appendix

First, we recall some results and properties on the closability of linear operators. Given $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ two Banach spaces, and $A : \mathrm{dom}(A) \longrightarrow Y$ a linear operator defined on a subspace $\mathrm{dom}(A)$ of X, the domain of A, the operator A is said to be closed if, for any sequence $(x_n)_{n\geq 1} \subset \mathrm{dom}(A)$, such that x_n converges to x in X and Ax_n converges to y in Y we have $x \in \mathrm{dom}(A)$ and $y = Ax$, i.e., $\mathrm{dom}(A)$ is closed (or equivalently complete) *w.r.t.* the graph norm $\|\cdot\|_G := \|\cdot\|_X + \|A\cdot\|_Y$. A linear operator $A : \mathrm{dom}(A) \longrightarrow Y$ is said closable if, for any sequence $(x_n)_{n\geq 1} \subset \mathrm{dom}(A)$ such that x_n converges to 0 in X and Ax_n converges to y in Y it holds $y = 0$. In other words, A is closable if, for any sequence $(x_n)_{n\geq 1} \subset \mathrm{dom}(A)$ such that x_n converges to 0 in X and $(x_n)_{n\geq 1}$ is Cauchy *w.r.t.* the graph norm $\|\cdot\|_G$ it holds Ax_n converges to 0 in Y. The minimal closed extension of the closable operator A is the closed operator $\overline{A}$ whose domain $\mathrm{dom}(\overline{A})$ is the completion of $\mathrm{dom}(A)$ *w.r.t.* $\|\cdot\|_G$, i.e.,

$$\mathrm{dom}(\overline{A}) := \{x \in X : \exists\, (x_n)_{n\geq 1} \subset \mathrm{dom}(A) : \ x_n \to x \text{ in } X \text{ and } (Ax_n)_{n\geq 1} \text{ converges in } Y\}$$

and we define

$$\overline{A}x := \lim_{n\to\infty} Ax_n, \qquad x \in \mathrm{dom}(\overline{A}),$$

where the limit is in Y and $(x_n)_{n\geq 1}$ is some sequence in $\mathrm{dom}(A)$ such that x_n converges to x in X and $(Ax_n)_{n\geq 1}$ converges in Y.

Next, we recall some notions of Dirichlet forms theory. We begin with some definitions related to bilinear forms (see [25] for details). Let H be a Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and $\mathcal{A} : \mathrm{dom}(\mathcal{A}) \times \mathrm{dom}(\mathcal{A}) \longrightarrow \mathbb{R}$ a bilinear form defined on a dense subspace $\mathrm{dom}(\mathcal{A})$ of H, the domain of $\mathcal{A}$. The form $\mathcal{A}$ is said to be symmetric if $\mathcal{A}(F,G) = \mathcal{A}(G,F)$, for any $F, G \in \mathrm{dom}(\mathcal{A})$, and nonnegative definite if $\mathcal{A}(F,F) \geq 0$, for any $F \in \mathrm{dom}(\mathcal{A})$. Let $\mathcal{A}$ be symmetric and nonnegative definite, $\mathcal{A}$ is said closed if $\mathrm{dom}(\mathcal{A})$ equipped with the norm

$$\|F\|_{\mathcal{A}} := \sqrt{\mathcal{A}(F,F) + \langle F,F\rangle}, \quad F \in \mathrm{dom}(\mathcal{A}),$$

is a Hilbert space. A symmetric and nonnegative definite bilinear form $\mathcal{A}$ is said closable if, for any sequence $(F_n)_{n\geq 1} \subset \mathrm{dom}(\mathcal{A})$ such that F_n goes to 0 in H and $(F_n)_{n\geq 1}$ is Cauchy *w.r.t.* $\|\cdot\|_{\mathcal{A}}$ it holds that $\mathcal{A}(F_n, F_n)$ converges to 0 in $\mathbb{R}$ as n goes to infinity. Let $\mathcal{A}$ be closable and denote by $\mathrm{dom}(\overline{\mathcal{A}})$ the completion of $\mathrm{dom}(\mathcal{A})$ *w.r.t.* the norm $\|\cdot\|_{\mathcal{A}}$. It turns out that $\mathcal{A}$ is uniquely extended to $\mathrm{dom}(\overline{\mathcal{A}})$ by the closed, symmetric, and nonnegative definite bilinear form

$$\overline{\mathcal{A}}(F,G) = \lim_{n\to\infty} \mathcal{A}(F_n, G_n), \quad (F,G) \in \mathrm{dom}(\overline{\mathcal{A}}) \times \mathrm{dom}(\overline{\mathcal{A}}),$$

where $\{(F_n, G_n)\}_{n\geq 1}$ is any sequence in $\mathrm{dom}(\mathcal{A}) \times \mathrm{dom}(\mathcal{A})$ such that (F_n, G_n) converges to $(F,G) \in \mathrm{dom}(\overline{\mathcal{A}}) \times \mathrm{dom}(\overline{\mathcal{A}})$ *w.r.t.* the norm $\|\cdot\|_{\overline{\mathcal{A}}} + \|\cdot\|_{\overline{\mathcal{A}}}$. Suppose $H = L^2(B, \mathcal{B}, \beta)$ where $(B, \mathcal{B}, \beta)$ is a measure space. A symmetric, nonnegative definite, and closed form $\mathcal{A}$ is said to be a symmetric Dirichlet form if

$$\mathcal{A}(F^+ \wedge 1, F^+ \wedge 1) \leq \mathcal{A}(F,F), \qquad F \in \mathrm{dom}(\mathcal{A}),$$

where F^+ denotes the positive part of F. Suppose that B is a Hausdorff topological space and let $\mathcal{A}$ be a symmetric Dirichlet form. An $\mathcal{A}$-nest is an increasing sequence $(C_n)_{n\geq 1}$ of closed subsets of B such that

$$\bigcup_{n\geq 1}\{F \in \mathrm{dom}(\mathcal{A}) : \; F = 0 \; \beta\text{-a.e. on } B \setminus C_n\}$$

is dense in $\mathrm{dom}(\mathcal{A})$ *w.r.t.* the norm $\|\cdot\|_{\mathcal{A}}$. We say that a subset $B' \subset B$ is $\mathcal{A}$-exceptional if there exists an $\mathcal{A}$-nest $(C_n)_{n\geq 1}$ with $B' \subset B \setminus \bigcup_{n\geq 1} C_n$. Throughout this paper we say that a property holds $\mathcal{A}$-almost everywhere ($\mathcal{A}$-a.e.) if it holds up to an $\mathcal{A}$-exceptional set. Moreover, a function $f : B \to \mathbb{R}$ is called $\mathcal{A}$-almost continuous ($\mathcal{A}$-a.c.) if there exists an $\mathcal{A}$-nest $(C_n)_{n\geq 1}$ such that the restriction $f_{|C_n}$ of f to C_n is continuous for each $n \geq 1$.

Let B be again a Hausdorff topological space. A symmetric Dirichlet form $\mathcal{A}$ on the Hilbert space $L^2(B, \mathcal{S}(B), \beta)$ is called quasi-regular if

(1) There exists an $\mathcal{A}$-nest $(C_n)_{n\geq 1}$ consisting of compact sets.
(2) There exists a $\|\cdot\|_{\mathcal{A}}$-dense subset of $\mathrm{dom}(\mathcal{A})$ whose elements have $\mathcal{A}$-a.c. β-versions.
(3) There exist $F_k \in \mathrm{dom}(\mathcal{A})$, $k \geq 1$, having $\mathcal{A}$-a.c. β-versions $\tilde{F}_k$, $k \geq 1$, such that $(\tilde{F}_k)_{k\geq 1}$ is a separating set for $B \setminus N$ (i.e., for any $x, y \in B \setminus N$, $x \neq y$, there exists $\tilde{F}_{k^*}$ such that $\tilde{F}_{k^*}(x) \neq \tilde{F}_{k^*}(y)$), where N is a subset of B which is $\mathcal{A}$-exceptional.

Acknowledgement This work was supported in part by the Singapore MOE Tier 2 grant MOE2012-T2-2-033 ARC 3/13.

References

1. Albeverio, S., Kondratiev, Y.G., Röckner, M.: Analysis and geometry on configuration spaces. J. Funct. Anal. **154**, 444–500 (1998)
2. Borodin, A., Okounkov, A., Olshanski, G.: Asymptotics of Plancherel measures for symmetric groups. J. Am. Math. Soc. **13**, 481–515 (2000)
3. Borodin, A., Olshanski, G.: Point processes and the infinite symmetric group. Math. Res. Let. **5**, 799–816 (1998)
4. Borodin, A., Olshanski, G.: Distributions on partitions, point processes, and the hypergeometric kernel. Commun. Math. Phys. **211**, 335–358 (2000)
5. Borodin, A., Olshanski, G.: Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes Ann. Math. **161**, 1319–1422 (2005)
6. Borodin, A., Olshanski, G.: Random partitions and the gamma kernel. Adv. Math. **194**, 141–202 (2005)
7. Brezis, H.: Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris (1983)
8. Camilier, I., Decreusefond, L.: Quasi-invariance and integration by parts for determinantal and permanental processes. J. Funct. Anal. **259**, 268–300 (2010)

9. Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. Elementary Theory and Methods. Probability and its Applications (New York), vol. I, 2nd edn. Springer, New York (2003)
10. Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. General Theory and Structure. Probability and its Applications (New York), vol. II, 2nd edn. Springer, New York (2008)
11. Decreusefond, L., Flint, I., Privault, N., Torrisi, G.L.: Stochastic dynamics of determinantal processes by integration by parts. Commun. Stoch. Anal. **9**, 375–399 (2015)
12. Decreusefond, L., Flint, I., Vergne, A.: Efficient simulation of the Ginibre point process. Adv. Appl. Probab. **52**(4), 1003–1012 (2015)
13. Dunford, N., Schwartz, J.T.: Linear Operators. Part I. General Theory. With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication. Wiley Classics Library. Wiley, New York (1988)
14. Fukushima, M., Ōshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. de Gruyter Studies in Mathematics, vol. 19. Walter de Gruyter and Co., Berlin (1994)
15. Georgii, H., Yoo, H.J.: Conditional intensity and Gibbsianness of determinantal point processes. J. Stat. Phys. **118**, 55–84 (2005)
16. Ginibre, J.: Statistical ensembles of complex, quaternion, and real matrices. J. Math. Phys. **6**, 440–449 (1965)
17. Hough, J.B., Krishnapur, M., Peres, Y., Virág, B.: Determinantal processes and independence. Probab. Surv. **3**, 206–229 (electronic) (2006)
18. Hough, J.-B., Krishnapur, M., Peres, Y., Virág, B.: Zeros of Gaussian Analytic Functions and Determinantal Point Processes. University Lecture Series, vol. 51. American Mathematical Society, Providence (2009)
19. Kendall, W.S.: Perfect simulation for the area-interaction point process. In: Accardi, L., Heyde, C.C. (eds.) Probability Towards 2000. Lecture Notes in Statistics, vol. 128, pp. 218–234. Springer, New York (1998)
20. Kendall, W.S., Møller, J.: Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. Adv. Appl. Probab. **32**, 844–865 (2000)
21. Last, G.: Stochastic analysis for Poisson processes. In: Peccati, G., Reitzner, M. (eds.) Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Ito Chaos Expansions and Stochastic Geometry. Bocconi & Springer Series, vol. 7, pp. 1–36. Springer, Cham (2016)
22. Lavancier, F., Møller, J., Rubak, E.: Determinantal point process models and statistical inference. J. R. Stat. Soc. Ser. B **77**, 853–877 (2015)
23. van Lieshout, M.N.M.: Markov Point Processes and Their Applications. Imperial College Press, London (2000)
24. Lytvynov, E.: Determinantal point processes with J-Hermitian correlation kernels. Ann. Probab. **41**, 2513–2543 (2013)
25. Ma, Z.M., Röckner, M.: Introduction to the Theory of (Nonsymmetric) Dirichlet Forms. Universitext. Springer, Berlin (1992)
26. Ma, Z.M., Röckner, M.: Construction of diffusions on configuration spaces. Osaka J. Math. **37**, 273–314 (2000)
27. Macchi, O.: The coincidence approach to stochastic point processes. Adv. Appl. Probab. **7**, 83–122 (1975)
28. Miyoshi, N., Shirai, T.: Spatial stochastic models for analysis of heterogeneous cellular networks with repulsively deployed base stations. Perform. Eval. **77**, 7–17 (2014)
29. Møller, J., Waagepetersen, R.P.: Statistical Inference and Simulation for Spatial Point Processes. Monographs on Statistics and Applied Probability, vol. 100. Chapman and Hall/CRC, Boca Raton (2004)
30. Robert, P.: Stochastic Networks and Queues. Stochastic Modelling and Applied Probability Series. Springer, New York (2003)
31. Shirai, T.: Large deviations for the fermion point process associated with the exponential kernel. J. Stat. Phys. **123**, 615–629 (2006)

32. Shirai, T., Takahashi, Y.: Random point fields associated with certain Fredholm determinants I. fermion, Poisson and boson point processes. J. Funct. Anal. **205**, 414–463 (2003)
33. Simon, B.: Trace Ideals and Their Applications. Mathematical Surveys and Monographs, vol. 120, 2nd edn. American Mathematical Society, Providence (2005)
34. Soshnikov, A.: Determinantal random point fields. Usp. Mat. Nauk **55**, 107–160 (2000)
35. Torrisi, G.L.: Quasi-invariance of fermion processes with J-Hermitian kernel. Commun. Stoch. Anal. **8**, 169–190 (2014)
36. Torrisi, G.L.: Point processes with Papangelou conditional intensity: from the Skorohod integral to the Dirichlet form. Markov Process. Relat. Fields **19**, 195–248 (2013)
37. Torrisi, G.L., Leonardi, E.: Large deviations of the interference in the Ginibre network model. Stoch. Syst. **4**, 173–205 (2014)

Index

G. Peccati, M. Reitzner (eds.), *Stochastic Analysis for Poisson Point Processes*,
Bocconi & Springer Series 7, DOI 10.1007/978-3-319-05233-5