# Distributional Approximations and Set-Valued Sublinear Expectations

Inaugural dissertation
of the Faculty of Science,
University of Bern

presented by

Riccardo Turin
from Italy

Supervisor of the doctoral thesis:
Prof. Dr. Ilya Molchanov

Institute of Mathematical Statistics and Actuarial Sciences
University of Bern

# Distributional Approximations and Set-Valued Sublinear Expectations

Inaugural dissertation
of the Faculty of Science,
University of Bern

presented by

Riccardo Turin
from Italy

Supervisor of the doctoral thesis:
Prof. Dr. Ilya Molchanov

Institute of Mathematical Statistics and Actuarial Sciences
University of Bern

Accepted by the Faculty of Science.

Bern, 24.03.2022

The dean:
Prof. Dr. Zoltan Balogh

*To my wife Maria*

# Abstract


This dissertation is composed by two blocks.

The first part is concerned with several types of distributional approximations, namely multivariate Poisson, Poisson process and Gaussian approximation.

Employing the solution of the Stein equation for Poisson distribution, we obtain an explicit bound for the multivariate Poisson approximation of random vectors in the Wasserstein distance. The bound is then utilized in the context of point processes, to provide a Poisson process approximation result in terms of a new metric called $d_\pi$, defined as the supremum over all Wasserstein distances between random vectors obtained evaluating the point processes on arbitrary collections of disjoint sets. As applications, the multivariate Poisson approximation of the sum of $m$-dependent Bernoulli random vectors, the Poisson process approximation of point processes of $U$-statistic structure and the Poisson process approximation of point processes with Papangelou intensity are considered.

Next, we consider a variant of the classical Johnson–Mehl birth-growth model with random growth speed and prove Gaussian approximation results. In this model, seeds appear at random times and locations and start growing instantaneously in all directions with random speeds. The location, birth time and growth speed of the seeds are given by a Poisson process. Under suitable conditions on the random growth speed and birth time distribution, we establish quantitative central limit theorems for the sum of given weights at the exposed points, which are those seeds in the model that are not covered at the time of their birth. Such models have previously been considered, albeit with deterministic growth speed.

In the second part of the dissertation, we propose general construction of convex closed sets obtained by applying sublinear expectations to random vectors in Euclidean space. We show that many well-known transforms in convex geometry (in particular, centroid body, convex floating body, and Ulam floating body) are special instances of our construction. Further, we identify the dual representation of such convex bodies and identify one map that serves as a building block for all so defined convex bodies. Several further properties are investigated.

# Acknowledgments

First, thank you Ilya, for your many advises, and your patience.

Thank you Prof. Thäle, for taking time to review this manuscript.

Thank you Johanna and David, for your affectionate guidance.

Thank you Lutz, for the Friday runs.

Thank you Andrea, for your welcoming presence in our department.

Thank you Matthias, for nice and fruitful discussions.

Thank you Fede, for accepting I do not love pubs as much as you do.

Thank you Alex, Maria Luisa and Chinmoy, for our true, wonderful friendship.

Thank you Sara, for your huge and sensitive heart.

Thank you Matteo, for I can count on you each and every good weather weekend.

Thank you Basel guys, for your help in making life a joy.

Thank you Marco and Madda, for the way you care about us.

Thank you Franci, Dolbi, Gio, Gec, Leti, Edo, Marta, Madda, Marco, Lando and Don Cesare, for you are the sign that He is here.

Thank you Maria and Agnese, there are no words to say what a gift you are.

Thank you my family and friends, sorry that I do not cite you personally one by one.

Finally, and most importantly, thank You Father, for gifting me all that is written above and more. You are and give everything.

# Contents

# Notation

| | |
|---|---|
| a.s. | almost surely |
| i.e. | that is ('id est') |
| $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ | set of integers |
| $\mathbb{N} = \{1, 2, \dots\}$ | set of positive integers |
| $\mathbb{N}_0 = \{0, 1, \dots\}$ | set of non-negative integers |
| $\mathbb{R} = (-\infty, \infty)$ | real line |
| $\mathbb{R}_+ = [0, \infty)$ | non-negative real half-line |
| $\|x\|$ | Euclidean norm of $x \in \mathbb{R}^d$ |
| $\|f\| = \sup_{x \in \mathbb{X}} \lvert f(x)\rvert$ | for a real valued function $f$ with domain $\mathbb{X}$ |
| $\mu f = \int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$ | for a measure $\mu$ on $\mathbb{X}$ and an integrable function $f$ |
| $\langle \cdot, \cdot \rangle$ | Euclidean scalar product on $\mathbb{R}^d$ |
| $B_r(x)$ | $d$-dimensional closed Euclidean ball of radius $r \geq 0$ centered at $x \in \mathbb{R}^d$ |
| $\mathbf{1}_A$ or $\mathbf{1}\{A\}$ | indicator function over a set $A$ |
| $\delta_x$ | Dirac measure at a point $x$ |
| $a \vee b$, $a \wedge b$ | maximum (resp. minimum) of $a$ and $b$ |
| $\mathcal{P}(\mathbb{X})$ | the power set of $\mathbb{X}$, that is, the collection of all subsets of $\mathbb{X}$ |
| $\mathcal{C}_b(\mathbb{X})$ | set of bounded and continuous functions on $\mathbb{X}$ |
| $\mathcal{C}_K^+(\mathbb{X})$ | set of non-negative and continuous functions with compact support |
| $\mathbf{N}_{\mathbb{X}}$ | set of $\sigma$-finite counting measures on $\mathbb{X}$ |
| $\widetilde{\mathbf{N}}_{\mathbb{X}}$, $\widehat{\mathbf{N}}_{\mathbb{X}}$ | set of locally finite (resp. finite) counting measures on $\mathbb{X}$ |
| lcscH | locally compact second countable Hausdorff |
| $(\Omega, \mathfrak{F}, \mathbb{P})$ | reference probability space |
| $\mathbb{P}\{A\}$ or $\mathbb{P}(A)$ | probability of an event $A \in \mathfrak{F}$ |
| $\mathbb{E}[X]$ or $\mathbb{E}(X)$ or $\mathbb{E}X$ | expectation of a random variable $X$ |
| $\mathrm{Var}(X)$ or $\mathrm{Var}\,X$ | variance of a random variable $X$ |
| $\mathscr{L}_X$ | distribution of a random element $X : \Omega \to \mathbb{X}$ over $\mathbb{X}$ |
| $\overset{d}{=}$, $\overset{d}{\to}$ | equality (resp. convergence) in distribution |

# Chapter 1

# Introduction

The content of this thesis has been organized in two blocks. The first parts concerns distributional approximations, while the second one presents a multivariate notion of sublinear expectation and its connections to convex geometric constructions.

## Part I: Distributional approximations

When dealing with distributional approximation, one may be concerned with at least two related aspects. The first, more basic point of interest is that of identifying the limit distribution of a sequence of random elements, in the weak convergence sense. When the limit distribution is standard normal we talk about Gaussian convergence, or central limit theorem (CLT), whereas for Poissonian limits, we talk about Poisson convergence. While weak convergence results like the CLT describe the asymptotic behavior of a sequence of random elements, they do not provide any information concerning how fast this convergence happens or, equivalently, how close a variable from the process is to the limit. This is exactly the second, refined point of view in the context of distributional approximation, which studies the approximation error and provides so-called quantitative limit theorems. In the first part of this manuscript we furnish several quantitative limit theorems in the context of multivariate Poisson, Poisson process and Gaussian approximation.

### Multivariate Poisson approximation

We treat Multivariate Poisson approximation in the first part of Chapter 3. Our aim is to compare the distributions of a non-negative integer-valued random vector $\mathbf{X}$ and a Poisson random vector $\mathbf{P}$ , in terms of the

Wasserstein distance

$$d_W(\mathbf{X},\mathbf{P}) = \sup_{h\in\mathrm{Lip}^d(1)} \left|\mathbb{E}\left[h(\mathbf{X})\right] - \mathbb{E}\left[h(\mathbf{P})\right]\right| ,$$

where $\mathrm{Lip}^d(1)$ denotes the set of 1-Lipschitz functions. This problem has been studied by several authors, e.g. [5, 9, 10, 12, 35, 89, 90], yet mostly in terms of the total variation distance

$$d_{TV}(\mathbf{X},\mathbf{P}) = \sup_{A\in\mathcal{B}(\mathbb{R}^d)} \left|\mathbb{P}\left\{\mathbf{X}\in A\right\} - \mathbb{P}\left\{\mathbf{P}\in A\right\}\right| .$$

Note that, for integer-valued random vectors, the Wasserstein distance always dominates the total variation distance. To date, not many results are known for the multivariate Poisson approximation in Wasserstein distance.

We provide a general upper bound on the approximation error in Theorem 3.1.1. It is obtained applying the Chen-Stein method to each component of the random vectors and combining it with a generalization of the size-bias distribution to the multidimensional framework.

The general result is then applied, in Section 3.2, to Poisson approximate the sum of Bernoulli random vectors. By a Bernoulli random vector, we mean a random vector with values in the set composed by the canonical vectors of $\mathbb{R}^d$ and the null vector. This problem has been mainly studied in terms of the total variation distance and under the assumption that the Bernoulli random vectors are independent, see e.g. [88]. We derive an explicit approximation result in the Wasserstein distance for the more general case of $m$-dependent Bernoulli random vectors.

## Poisson process approximation

Poisson process approximation is the content of the last three sections of Chapter 3. We introduce a new distance $d_\pi$ between point processes with finite intensity measure, defined as the supremum over all Wasserstein distances between random vectors obtained evaluating the point processes on arbitrary collections of disjoint sets:

$$d_\pi(\xi,\eta) = \sup_{(A_1,\dots,A_d)\in\mathcal{X}^d_{\shortparallel},\, d\in\mathbb{N}} d_W\big((\xi(A_1),\dots,\xi(A_d)),(\eta(A_1),\dots,\eta(A_d))\big),$$

where $\mathcal{X}^d_{\shortparallel}$ is the family of $d$-tuples of disjoint measurable subsets of the underlying space.

For a point processes $\xi$ and a Poisson process $\eta$ on a measurable space $\mathbb{X}$, our abstract result on multivariate Poisson approximation, Theorem 3.1.1, provides bounds on the Wasserstein distance

$$d_W((\xi(A_1),\dots,\xi(A_d)),(\eta(A_1),\dots,\eta(A_d)) ,$$

where $A_1, \dots, A_d$ are disjoint measurable subsets of $\mathbb{X}$. In this way, a general bound on the distance $d_\pi$ between a point process $\xi$ and a Poisson point process $\eta$, is directly obtained from our multivariate Poisson approximation result.

In Sections 3.4 and 3.5, we apply the Poisson process approximation result, Theorem 3.3.5, to obtain explicit Poisson process approximation results for point processes with Papangelou intensity and point processes of Poisson $U$-statistic structure. The latter are point processes that, once evaluated on a measurable set, become Poisson $U$-statistics. Analogous results have been already proven for the Kantorovich-Rubinstein distance in [31, Theorem 3.7] and [30, Theorem 3.1], under the additional condition that the configuration space $\mathbb{X}$ is lcscH.

## Gaussian approximation

In Chapter 4, we establish Gaussian approximation results for functionals defined on a generalized version of the Johnson–Mehl growth model.

In the classical Johnson–Mehl growth model, seeds appear at random times $t_i$, $i \in \mathbb{N}$, at random locations $x_i$, $i \in \mathbb{N}$, in $\mathbb{R}^d$, according to a Poisson process $(x_i, t_i)_{i \in \mathbb{N}}$ on $\mathbb{R}^d \times \mathbb{R}_+$. Once a seed is born at time $t$, it begins to form a cell by growing radially in all directions at a constant speed $v \geq 0$, so that by time $t'$ it occupies the ball of radius $v(t' - t)$. The parts of the space claimed by the seeds form the so-called Johnson–Mehl tessellation, see [29] and [76].

The study of such birth-growth processes started with the work of Kolmogorov [53] in two dimensions to model crystal growth. Since then, this model has seen applications in many contexts. For various subsequent developments and applications, see [27, 29, 76] and references therein.

Variants of the classical spatial birth-growth model can be found, sometimes as a particular case of other models, in many subsequent papers. Among them, we mention [81] and [14], where the birth-growth model appears as a particular case of a random sequential packing model, and [94], which studies a variant of the model with non-uniform deterministic growth patterns. The main tools rely on the concept of stabilization by considering regions where the appearance of new seeds influences the functional of interest.

In our work, we consider a generalization of the Johnson–Mehl model by introducing random growth speed for the seeds. This gives rise to many interesting features in the model, most importantly, long-range interactions if the speed can take arbitrarily large values with positive probability. Therefore, the model with random speed is no longer stabilizing in the classical sense of [58] and [82], since distant points may influence the growth pattern if their speeds are sufficiently high. It should be noted that, even in the constant speed setting, we substantially improve and extend limit theorems obtained in [27].

We consider a birth-growth model, determined by a Poisson process $\eta$ in $\mathbb{X} := \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure $\mu := \lambda \otimes \theta \otimes \nu$, where $\lambda$ is the Lebesgue measure on $\mathbb{R}^d$, $\theta$ is a non-null locally finite measure on $\mathbb{R}_+$, and $\nu$ is a probability distribution on $\mathbb{R}_+$ with $\nu(\{0\}) < 1$. Each point $\boldsymbol{x}$ of this point process $\eta$ has three components $(x, t_x, v_x)$, where $v_x \in \mathbb{R}_+$ denotes the random speed of a seed born at location $x \in \mathbb{R}^d$ and whose growth commences at time $t_x \in \mathbb{R}_+$. In a given point configuration, a point $\boldsymbol{x} := (x, t_x, v_x)$ is said to be exposed if there is no other point $(y, t_y, v_y)$ in the configuration with $t_y < t_x$ and $\|x - y\| \leq v_y(t_x - t_y)$, where $\|\cdot\|$ denotes the Euclidean norm. It should be noted that, because of random speeds, it may happen that the cell grown from a non-exposed seed shades a subsequent seed which would be exposed otherwise. Also notice that the event that a point $(x, t_x, v_x) \in \eta$ is exposed depends only on the point configuration in the region

$$L_{x,t_x} := \big\{(y, t_y, v_y) \in \mathbb{X} : \|x - y\| \leq v_y(t_x - t_y)\big\}.$$

Namely, $\boldsymbol{x}$ is exposed if and only if $\eta$ has no points (apart from $\boldsymbol{x}$) in $L_{x,t_x}$.

Given a measurable weight function $h : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ the main object of interest for us is the sum of $h$ over the exposed points in $\eta$:

$$F(\eta) := \sum_{\boldsymbol{x} \in \eta} h(x, t_x) \mathbf{1}\{\boldsymbol{x} \text{ is exposed}\}$$

Our aim is to provide sufficient conditions for Gaussian convergence of such sums. A standard approach for proving Gaussian convergence for such statistics relies on stabilization theory [14, 34, 81, 94]. While in the stabilization literature, one commonly assumes that the so-called stabilization region is a ball around a given reference point, the region $L_{x,t_x}$ is unbounded and it seems that it is not expressible as a ball around $\boldsymbol{x}$ in some different metric. Moreover, our stabilization region is set to be empty if $\boldsymbol{x}$ is not exposed.

The recent work [17] introduced a new notion of *region-stabilization* which allows for more general regions than balls. We will utilize the main result from [17], which is reported in Section 2.4, to derive bounds on the Wasserstein and Kolmogorov distances, between a suitably normalized sum of weights and the standard Gaussian distribution.

## Part II: Set-valued sublinear expectations

A sublinear expectation $\mathsf{e}$ is a sublinear (positively homogeneous and convex) map from the space $L^p(\mathbb{R})$ (or another linear space of random variables) to $(-\infty, \infty]$, and so may be regarded as a convex function on an infinite-dimensional space, see [110] for a thorough account of convex analysis tools in the infinite-dimensional setting.

The concept of sublinear expectation is essential in mathematical finance, where it is used to quantify the operational risk, see [33, 39]. The sublinearity property reflects the financial paradigm, saying that the diversification decreases the risk, and so the risk of a diversified portfolio is dominated by the sum of the risks of its components. Sublinear expectations are closely related to solutions of backward stochastic differential equations, see [79].

In this work we use a sublinear expectation $\mathsf{e}$ to associate with each $p$-integrable random vector $\xi$ in $\mathbb{R}^d$ a convex closed set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ in $\mathbb{R}^d$. This is done by letting the support function of $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ be the sublinear expectation $\mathsf{e}$ applied to the scalar product $\langle \xi, u \rangle$. For instance, if $\mathsf{e}$ is, up to a sign change, the average value at risk, one obtains convex closed sets called metronoids and studied by Huang and Slomka [48]. Further examples are given by expected random polytopes, which also form a special case of our construction.

We commence with Section 5.1, giving the definition of sublinear expectation of random variables, explaining their dual representation and presenting several examples. We mention the particularly important Kusuoka representation which expresses any law-determined sublinear expectation in terms of integrated quantiles and describe a novel construction (called the maximum extension) suitable to produce parametric families of sublinear expectations from each given one.

Section 5.2.1 presents our construction of convex closed sets $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ generated by a random vector $\xi$ and a given sublinear expectation $\mathsf{e}$. Section 5.2.2 describes a generalization based on relaxing some properties of the underlying numerical sublinear expectations, namely, replacing them with gauge functions. This construction yields centroid bodies [67] and half-space depth-trimmed regions [73], the latter are closely related to convex floating bodies introduced in [97] and their weighted variant from [16].

One of the most important sublinear expectations is based on using weighted integrals of the quantile function. The corresponding convex bodies are studied in Section 5.3, where we show their close connection to metronoids [48] and zonoid-trimmed regions [56]. The Kusuoka representation of numerical sublinear expectations yields Theorem 5.3.4, which provides a representation of a general convex set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ (derived from $\xi$ using a sublinear expectation $\mathsf{e}$) in terms of Aumann integrals of metronoids. We further provide a uniqueness result for the distribution of $\xi$ on the basis of a family of convex bodies generated by it, and also a concentration result for random convex sets constructed from the empirical distribution of $\xi$.

Section 5.4 specialises our general construction to the case when $\xi$ is uniformly distributed on a convex body $K$ (that is, a compact convex set in $\mathbb{R}^d$ with nonempty interior), and so $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ yields a transform $K \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K) = \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$. We derive several properties of this transformation for general $\mathsf{e}$, in particular, establish the continuity of such maps in the Hausdorff metric.

In special cases, our construction yields $L^p$-centroid bodies (see [67] and [92, Sec. 10.8]) and Ulam floating bodies recently introduced in [49]. The latter form a particularly important special setting, which is confirmed by showing that all transformations $K \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ can be expressed in terms of Ulam floating bodies. For instance, Corollary 5.4.8 provides a representation of the centroid body of an origin symmetric $K$ as the convex hull of dilated Ulam floating bodies of $K$. In this course, results for sublinear expectations yield a new insight into the well-known aforementioned constructions of convex bodies, deliver some new relations between them, and provide a general source of nonlinear transformations of convex bodies. Finally, we formulate several conjectures.

# Part I

# Distributional Approximations

# Chapter 2

# Preliminaries

This chapter introduces some basic notions and results from the areas that this part of the thesis touches upon. For each of them, we do not provide a thorough account of the topic, nor a comprehensive literature review, but simply mention definitions and results that prove necessary in an attempt to make this manuscript as self-contained as possible.

## 2.1 Point processes

For the proofs of the results mentioned in this section, the interested reader is referred to the books [52] and [60], where most of the material is taken from.

### 2.1.1 Point processes on measurable spaces

Intuitively, a point process is a random collection of points in a space. This concept can be formalized by describing a (random) collection of points by the (random) measure that provides for each set the number of points in it. Let $(\mathbb{X}, \mathcal{X})$ be a measurable space.

**Definition 2.1.1.** A measure $\mathsf{M}$ on $(\mathbb{X}, \mathcal{X})$ is called

i) *σ-finite* if there exist $A_n \in \mathcal{X}$, $n \in \mathbb{N}$ such that $\mathsf{M}(A_n) < \infty$ and $\mathbb{X} = \bigcup_{n\in\mathbb{N}} A_n$ ;

ii) *counting measure* if $\mathsf{M}(A) \in \overline{\mathbb{N}}_0$ for all $A \in \mathcal{X}$ .

Let $\mathbf{N}_{\mathbb{X}}$ be the collection of all $\sigma$-finite counting measures $\mathsf{M}$ on $(\mathbb{X}, \mathcal{X})$, and let $\mathcal{N}_{\mathbb{X}}$ be the smallest $\sigma$-alegbra on $\mathbf{N}_{\mathbb{X}}$ such that the maps $\mathsf{M} \mapsto \mathsf{M}(A)$ are measurable for all $A \in \mathcal{X}$.

**Definition 2.1.2.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the reference probability space. A *point process* $\xi$ on $\mathbb{X}$ is a random element in $\mathbf{N}_{\mathbb{X}}$, that is, a measurable map $\xi : \Omega \to \mathbf{N}_{\mathbb{X}}$.

A fist example of point process is the *binomial process*, which has the form

$$\xi = \sum_{n=1}^{m} \delta_{X_n} \tag{2.1.1}$$

for some $m \in \mathbb{N}$ and i.i.d. random elements $X_1, \dots, X_m \in \mathbb{X}$. Note that, for a measurable set $A \in \mathcal{X}$, the random variable $\xi(A)$ has the binomial distribution with sample size $m$ and success rate $\mathscr{L}_{X_1}(A)$, where $\mathscr{L}_{X_1}$ is the distribution of $X_1$ in $\mathbb{X}$.

**Definition 2.1.3.** The *intensity measure* of a point process $\xi$ on $\mathbb{X}$ is the measure $\mu$ defined by

$$\mu(A) \coloneqq \mathbb{E}\left[\xi(A)\right], \quad A \in \mathcal{X}\,.$$

In the case of the binomial point process defined at (2.1.1), the intensity measure is $\mu = m\mathscr{L}_{X_1}$.

**Proposition 2.1.4.** *(Campbell's formula) Let $\xi$ be a point process on $\mathbb{X}$ with intensity measure $\mu$, and let $h\colon \mathbb{X} \to \overline{\mathbb{R}}$ be a measurable function. Then $\int_{\mathbb{X}} h(x)\,\xi(\mathrm{d}x)$ is a random variable and*

$$\mathbb{E}\left[\int_{\mathbb{X}} h(x)\,\xi(\mathrm{d}x)\right] = \int_{\mathbb{X}} h(x)\,\mu(\mathrm{d}x)\,.$$

The *distribution* of a point process $\xi$ on $\mathbb{X}$ is the probability measure $\mathscr{L}_\xi$ on $(\mathbf{N}_{\mathbb{X}}, \mathcal{N}_{\mathbb{X}})$ given by

$$\mathscr{L}_\xi(B) \coloneqq \mathbb{P}\left\{\xi \in B\right\}, \quad B \in \mathcal{N}_{\mathbb{X}}\,.$$

We write $\xi \overset{d}{=} \zeta$ when $\mathscr{L}_\xi = \mathscr{L}_\zeta$. The next proposition gives two equivalent characterizations of equality in distributions between point processes.

**Proposition 2.1.5.** *Let $\xi$ and $\zeta$ be two point processes on $\mathbb{X}$. Then $\xi \overset{d}{=} \zeta$ if and only if one of the following equivalent conditions holds:*

*i)* $(\xi(A_1), \dots, \xi(A_d)) \overset{d}{=} (\zeta(A_1), \dots, \zeta(A_d))$*, for all $d \in \mathbb{N}$ and pairwise disjoint $A_1, \dots, A_d \in \mathcal{X}$;*

*ii) for every measurable function $h : \mathbb{X} \to \mathbb{R}_+$, the $\overline{\mathbb{R}}_+$-valued random variables $\int_{\mathbb{X}} h(x)\,\xi(\mathrm{d}x)$ and $\int_{\mathbb{X}} h(x)\,\zeta(\mathrm{d}x)$ have the same distribution.*

As we said at the beginning of this section, a point process can be thought as an at most countable random collection of points in the space $\mathbb{X}$. This intuition is particularly appropriate for the case of proper point processes.

**Definition 2.1.6.** A point process $\xi$ on $\mathbb{X}$ is called *proper* if there exist random elements $X_1, X_2, \dots \in \mathbb{X}$ and a random variable $\kappa$ with values in $\overline{\mathbb{N}}_0$ such that

$$\xi = \sum_{n=1}^{\kappa} \delta_{X_n} .$$

Clearly the binomial process from (2.1.1) is an example of proper point process. Although there are examples of non-proper point processes for pathological choices of the space $\mathbb{X}$, the class of proper point processes is very large, as we shall see in Section 2.1.3.

### 2.1.2 Poisson process

Let us now introduce what is arguably the most prominent example of point process. Recall that a random variable $X$ is Poisson distributed with mean $\gamma \geq 0$, if

$$\mathbb{P}\{X = k\} = \frac{\gamma^k}{k!} e^{-\gamma}, \quad k \in \mathbb{N}_0,$$

where, for $\gamma = 0$, we take $\mathbb{P}(X = 0) = 0^0 := 1$.

**Definition 2.1.7.** Let $\mu$ be a $\sigma$-finte measure on $(\mathbb{X}, \mathcal{X})$. A point process $\eta$ on $\mathbb{X}$ is called a *Poisson process* with intensity measure $\mu$ if

i) for all $A \in \mathcal{X}$, the random variable $\eta(A)$ is Poisson distributed with mean $\mu(A)$,

ii) for every $m \in \mathbb{N}$ and pairwise disjoint sets $A_1, \dots, A_m \in \mathcal{X}$, the random variables $\eta(A_1), \dots, \eta(A_m)$ are independent.

In view of Proposition 2.1.5, two Poisson processes with the same intensity measure have the same distribution. The following result provides existence of the Poisson process and, at the same time, ensures that for any Poisson process one can find a proper Poisson process with the same distribution.

**Proposition 2.1.8.** *Let $\mu$ be a $\sigma$-finte measure on $(\mathbb{X}, \mathcal{X})$. Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with random elements $X_1, X_2, \dots$ in $\mathbb{X}$ and $\kappa \in \overline{\mathbb{N}}_0$ such that*

$$\eta = \sum_{n=1}^{\kappa} \delta_{X_n}$$

*is a Poisson process with intensity measure $\mu$. In particular, when $0 < \mu(\mathbb{X}) < \infty$, $\kappa$ has the Poisson distribution with mean $\mu(\mathbb{X})$ and $X_1, X_2, \dots$ are independent of $\kappa$ and i.i.d. with distribution $\mu(\cdot)/\mu(\mathbb{X})$.*

The next proposition clarifies the relation of the Poisson process to the binomial process.

**Proposition 2.1.9.** *Let $\eta$ be a Poisson process on $\mathbb{X}$ with intensity measure $\mu$ such that $0 < \mu(\mathbb{X}) < \infty$. Then, conditioned on the event $\eta(\mathbb{X}) = m$, for some $m \in \mathbb{N}$, $\eta$ has the same distribution as a binomial point process of $m$ independent points with distribution $\mu(\cdot)/\mu(\mathbb{X})$.*

We recall here a characterization of the Poisson distribution.

**Proposition 2.1.10.** *An $\mathbb{N}_0$-valued random variable $X$ is Poisson distributed with mean $\gamma$ if and only if, for every function $f : \mathbb{N}_0 \to \mathbb{R}_+$*

$$\mathbb{E}\left[X f(X)\right] = \gamma \mathbb{E}\left[f(X+1)\right] .$$

In the context of point processes, the corresponding characterization of the Poisson process is given by the so-called Mecke equation.

**Proposition 2.1.11.** *(Mecke equation) Let $\mu$ be a $\sigma$-finte measure and let $\eta$ be a point process on $\mathbb{X}$. Then $\eta$ is a Poisson process with intensity measure $\mu$ if and only if*

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

*for all measurable functions $f : \mathbb{X} \times \mathbf{N}_{\mathbb{X}} \to \overline{\mathbb{R}}_+$.*

Proposition 2.1.11 admits a useful generalization involving multiple integration. To formulate it, we first need to introduce the factorial measure of a counting measure. Suppose $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ is given by

$$\mathsf{M} = \sum_{n=0}^{k} \delta_{x_n}$$

for some $k \in \overline{\mathbb{N}}_0$ and $x_n \in \mathbb{X}$. Then, for some $m \in \mathbb{N}$, on the power space $(\mathbb{X}^m, \mathcal{X}^m)$ we define the *$m$-th factorial measure*

$$\mathsf{M}^{(m)} := \sum_{i_1,\dots,i_m}^{\neq} \delta_{(x_{i_1},\dots,x_{i_m})} \in \mathbf{N}_{\mathbb{X}^m},$$

where the superscript $\neq$ indicates that the indexes $i_1, \dots, i_m \in \{1, \dots, k\}$ are pairwise different.

**Proposition 2.1.12.** *(Multivariate Mecke equation) Let $\eta$ be a Poisson process on $\mathbb{X}$ with $\sigma$-finte intensity measure $\mu$. Then, for any measurable function $f : \mathbb{X}^m \times \mathbf{N}_{\mathbb{X}} \to \overline{\mathbb{R}}_+$ with $m \in \mathbb{N}$,*

$$\mathbb{E}\left[\int_{\mathbb{X}^m} f(x_1,\dots,x_m,\eta)\,\eta^{(m)}(\mathrm{d}(x_1,\dots,x_m))\right] = \int_{\mathbb{X}^m} \mathbb{E}\left[f(x_1,\dots,x_m,\eta+\delta_{x_1}+\dots+\delta_{x_m})\right]\mu^m(\mathrm{d}(x_1,\dots,x_m)) .$$

### 2.1.3 Point processes on lcscH spaces

Let us now assume that $\mathbb{X}$ is a locally compact second countable Hausdorff (later: lcscH) space, that is, $\mathbb{X}$ is a topological space with countable base such that every point in $\mathbb{X}$ has an open neighborhood whose topological closure is compact and such that any two points of $\mathbb{X}$ can be separated by two disjoint open neighborhoods. Such a space is always separable and completely metrizable. In this case, $\mathcal{X}$ is the Borel $\sigma$-field of $\mathbb{X}$. It is notable that, when $\mathbb{X}$ is a lcscH space, every point process is proper, as follows e.g. by [93, Lemma 3.1.3].

We denote by $(\widetilde{\mathbf{N}}_{\mathbb{X}}, \widetilde{\mathcal{N}}_{\mathbb{X}})$ the measurable space of locally finite counting measures on $(\mathbb{X}, \mathcal{X})$, where $\widetilde{\mathcal{N}}_{\mathbb{X}}$ is the trace $\sigma$-field of $\mathcal{N}_{\mathbb{X}}$.

**Definition 2.1.13.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the reference probability space. A *locally finite point process* $\xi$ on $\mathbb{X}$ is a random element in $\widetilde{\mathbf{N}}_{\mathbb{X}}$, that is, a measurable map $\xi : \Omega \to \widetilde{\mathbf{N}}_{\mathbb{X}}$.

Note that any point process $\xi$ on $\mathbb{X}$ with locally finite intensity measure is almost surely locally finite, meaning that there exists another process $\xi'$ almost surely identical to $\xi$ and with realizations in $\widetilde{\mathbf{N}}_{\mathbb{X}}$.

We endow the space $\widetilde{\mathbf{N}}_{\mathbb{X}}$ with an appropriate topology in such a way that the $\sigma$-field $\widetilde{\mathcal{N}}_{\mathbb{X}}$ is the Borel $\sigma$-field of $\widetilde{\mathbf{N}}_{\mathbb{X}}$. Namely, let us equip the space $\widetilde{\mathbf{N}}_{\mathbb{X}}$ with the *vague topology*, i.e. the topology induced by the mappings

$$\widetilde{\mathbf{N}}_{\mathbb{X}} \ni \mathsf{M} \mapsto \int_{\mathbb{X}} f(x)\, \mathsf{M}(\mathrm{d}x), \quad f \in \mathcal{C}_K^+(\mathbb{X}),$$

where $\mathcal{C}_K^+(\mathbb{X})$ is the set of non-negative and continuous functions on $\mathbb{X}$ with compact support. Then, $\widetilde{\mathbf{N}}_{\mathbb{X}}$ equipped with the vague topology is a Polish space, and $\widetilde{\mathcal{N}}_{\mathbb{X}}$ coincides with the Borel $\sigma$-field generated by the vague topology. Now that the space of locally finite counting measures has a topological structure, we can talk about convergence in distribution of locally finite point process on $\mathbb{X}$.

**Definition 2.1.14.** Let $\xi, \xi_1, \xi_2, \ldots$ be locally finite point processes on a lcscH space $\mathbb{X}$. The sequence $(\xi_n)_{n\in\mathbb{N}}$ *converges in distribution to* $\xi$, if

$$\mathbb{E}\left[F(\xi_n)\right] \to \mathbb{E}\left[F(\xi)\right], \quad \text{for all } F \in \mathcal{C}_b(\widetilde{\mathbf{N}}_{\mathbb{X}}),$$

where $\mathcal{C}_b(\widetilde{\mathbf{N}}_{\mathbb{X}})$is the set of bounded and continuous functions on $\widetilde{\mathbf{N}}_{\mathbb{X}}$.

For a point process $\xi$ on a lcscH space $\mathbb{X}$, denote by $\mathcal{X}_c$ the class of relatively compact sets $A \in \mathcal{X}$, and by $\mathcal{X}_{c,\xi}$ the family of sets in $\mathcal{X}_c$ such that $\xi(\partial A) = 0$ almost surely. The following result characterizes the convergence in distribution of locally finite point processes on a lcscH space.

**Proposition 2.1.15.** *Let $\xi, \xi_1, \xi_2, \ldots$ be locally finite point processes on a lcscH space $\mathbb{X}$. Then, the following assertions are equivalent:*

*i)* $\xi_n \xrightarrow{d} \xi$*;*

*ii)* $\int_{\mathbb{X}} f(x)\, \xi_n(\mathrm{d}x) \xrightarrow{d} \int_{\mathbb{X}} f(x)\, \xi(\mathrm{d}x)$*, for all* $f \in \mathcal{C}_K^+(\mathbb{X})$*;*

*iii)* $(\xi_n(A_1), \ldots, \xi_n(A_k)) \xrightarrow{d} (\xi(A_1), \ldots, \xi(A_k))$ *for all* $A_1, \ldots, A_k \in \mathcal{X}_{c,\xi}$*,* $k \in \mathbb{N}$.

## 2.2 Distances between distributions

At the very base of the approximation concept lies the notion of a metric: Approximating something by something else means being able to identify an upper bound to the distance between the two objects, and this obviously depends on the chosen metric. In the following chapters, we approximate the distribution of different types of random objects, namely, random variables, random vectors and point processes. For each of these elements, a variety of metrics is available in the literature in order to compare their distributions. We hereafter recall, for each object category, the definitions of distances between probability distributions that are used or mentioned in the sequel, and some relationships between them.

### 2.2.1 Probability metrics

For two real-valued random variables $X$ and $Y$, we consider probability metrics of the form

$$d_{\mathcal{H}}(X,Y) = \sup_{h \in \mathcal{H}} \left| \mathbb{E}\left[h(X)\right] - \mathbb{E}\left[h(Y)\right] \right|, \tag{2.2.1}$$

where $\mathcal{H}$ is some family of test functions. For $\mathcal{H} = \{\mathbf{1}\{\cdot \le a\} \colon a \in \mathbb{R}\}$, the expression above corresponds to the *Kolmogorv* distance

$$d_K(X,Y) := \sup_{a \in \mathbb{R}} \left| \mathbb{P}\left\{X \le a\right\} - \mathbb{P}\left\{Y \le a\right\} \right|, \tag{2.2.2}$$

by taking $\mathcal{H} = \{\mathbf{1}\{\cdot \in A\} \colon A \in \mathcal{B}\}$ one obtains the *total variation* distance

$$d_{TV}(X,Y) := \sup_{A \in \mathcal{B}} \left| \mathbb{P}\left\{X \in A\right\} - \mathbb{P}\left\{Y \in A\right\} \right|,$$

while the choice $\mathcal{H} = \mathrm{Lip}(1) := \{h \colon \mathbb{R} \to \mathbb{R} \colon |h(x) - h(y)| \le |x - y|\}$ gives the *Wasserstein* distance

$$d_W(X,Y) := \sup_{h \in \mathrm{Lip}(1)} \left| \mathbb{E}\left[h(X)\right] - \mathbb{E}\left[h(Y)\right] \right|. \tag{2.2.3}$$

From their definitions it is immediately clear that $d_K \le d_{TV}$. When either X or Y has Lebesgue density bounded by a constant $C > 0$, then it is not hard to see that

$$d_K(X,Y) \le \sqrt{2C d_W(X,Y)}\,.$$

The metrics $d_{TV}$ and $d_W$ are in general not comparable, apart from when the random variables are integer-valued. In this case

$$d_{TV}(X,Y) = \frac{1}{2}\sum_{k\in\mathbb{Z}} \left|\mathbb{P}\{X=k\} - \mathbb{P}\{Y=k\}\right| \le d_W(X,Y),$$

and it is not hard to find sequences of integer-valued random variables that converge in total variation distance but not in Wasserstein distance.

Notice that a sequence of random variables that converges with respect to any of the metrics defined in this section is also weakly convergent. The same applies for the distances between multivariate distributions considered in the next section.

### 2.2.2 Distances between probability distributions of random vectors

For two random vectors $\mathbf{X}$ and $\mathbf{Y}$, i.e. random elements with values in $\mathbb{R}^d$ for some $2 \le d \in \mathbb{N}$, the *total variation* distance is defined by

$$d_{TV}(\mathbf{X},\mathbf{Y}) \coloneqq \sup_{A\in\mathcal{B}(\mathbb{R}^d)} \left|\mathbb{P}\{\mathbf{X}\in A\} - \mathbb{P}\{\mathbf{Y}\in A\}\right|,$$

and the *Wasserstein* distance is defined by

$$d_W(\mathbf{X},\mathbf{Y}) \coloneqq \sup_{h\in\mathrm{Lip}^d(1)} \left|\mathbb{E}[h(\mathbf{X})] - \mathbb{E}[h(\mathbf{Y})]\right|.$$

Note that, for convenience, we unusually define $\mathrm{Lip}^d(1)$ to be the collection of 1-Lipschitz functions $h:\mathbb{R}^d\to\mathbb{R}$ with respect to the metric induced by the 1-norm, that is,

$$|h(\boldsymbol{x}) - h(\boldsymbol{y})| \le |\boldsymbol{x}-\boldsymbol{y}|_1 = \sum_{i=1}^{d} |x_i - y_i|, \quad \boldsymbol{x},\boldsymbol{y}\in\mathbb{N}_0^d.$$

Clearly, this family of functions contains the 1-Lipschitz functions with respect to the Euclidean norm.

As for the case of random variables, for integer-valued random vectors $d_{TV} \le d_W$.

### 2.2.3 Distances between probability distributions of point processes

Let $\xi$ and $\zeta$ be two point processes on $\mathbb{X}$ with laws $\mathscr{L}_\xi$ and $\mathscr{L}_\zeta$, respectively. The *total variation* distance between their distributions is defined by

$$d_{TV}(\xi,\zeta) \coloneqq \sup_{B\in\mathcal{N}_{\mathbb{X}}} \left|\mathbb{P}\{\xi\in B\} - \mathbb{P}\{\zeta\in B\}\right| = \sup_{B\in\mathcal{N}_{\mathbb{X}}} \left|\mathscr{L}_\xi(B) - \mathscr{L}_\zeta(B)\right|. \tag{2.2.4}$$

Another way to compare the distributions of two point processes is to look at the minimum transportation cost. When the transportation cost is expressed by the total variation distance between measures,

$$d_{TV}(\mathsf{M},\mathsf{N}) := \sup_{\substack{A\in\mathcal{X}\\ \mathsf{M}(A),\,\mathsf{N}(A)<\infty}} |\mathsf{M}(A)-\mathsf{N}(A)|\,, \tag{2.2.5}$$

one obtains the *Kantorovich-Rubinstein* distance

$$d_{KR}(\xi,\zeta) := \inf_{\gamma\in\Gamma(\mathscr{L}_\xi,\mathscr{L}_\zeta)} \int_{\mathbf{N}_{\mathbb{X}}\times\mathbf{N}_{\mathbb{X}}} d_{TV}(\mathsf{M},\mathsf{N})\ \gamma\left(\mathrm{d}(\mathsf{M},\mathsf{N})\right), \tag{2.2.6}$$

where $\Gamma(\mathscr{L}_\xi,\mathscr{L}_\zeta)$ is the set of couplings of $\mathscr{L}_\xi$ and $\mathscr{L}_\zeta$, i.e. the set of probability measures on $\mathbf{N}_{\mathbb{X}}\times\mathbf{N}_{\mathbb{X}}$ with marginals $\mathscr{L}_\xi$ and $\mathscr{L}_\zeta$. The Kantorovich-Rubinstein distance is also called Wasserstein distance, Rubinstein distance or Monge-Kantorovich distance in the literature. The Kantorovich-Rubinstein distance always dominates the total variation distance because

$$\begin{aligned} d_{TV}(\xi,\zeta) &= \sup_{B\in\mathcal{N}_{\mathbb{X}}} \left| \inf_{\gamma\in\Gamma(\mathscr{L}_\xi,\mathscr{L}_\zeta)} \int_{\mathbf{N}_{\mathbb{X}}\times\mathbf{N}_{\mathbb{X}}} \mathbf{1}\{\mathsf{M}\in B\} - \mathbf{1}\{\mathsf{N}\in B\}\ \gamma\left(\mathrm{d}(\mathsf{M},\mathsf{N})\right)\right| \\ &\le \inf_{\gamma\in\Gamma(\mathscr{L}_\xi,\mathscr{L}_\zeta)} \int_{\mathbf{N}_{\mathbb{X}}\times\mathbf{N}_{\mathbb{X}}} d_{TV}(\mathsf{M},\mathsf{N})\ \gamma\left(\mathrm{d}(\mathsf{M},\mathsf{N})\right) = d_{KR}(\xi,\zeta)\,. \end{aligned}$$

On the other hand, there are sequences of point processes that converge to a limit point point process in total variation distance but diverge from the same reference point in Kantorovich-Rubinstein distance: see e.g. [30, Example 2.2]. The next proposition, which follows from Theorems 4.1 and 5.10 in [105], provides a dual characterization of the Kantorovich-Rubinstein distance between locally finite point processes on a lcscH space.

**Proposition 2.2.1.** *Let $\xi$ and $\zeta$ be locally finite point processes on a lcscH space $\mathbb{X}$. Then the infimum in* (2.2.6) *is attained and*

$$d_{KR}(\xi,\zeta) = \sup_{\mathcal{L}(\xi,\eta)} |\mathbb{E}[h(\xi)]-\mathbb{E}[h(\zeta)]|\,, \tag{2.2.7}$$

*where $\mathcal{L}(\xi,\eta)$ denotes the set of measurable functions $h:\widetilde{\mathbf{N}}_{\mathbb{X}}\to\mathbb{R}$ that are* 1*-Lipschitz with respect to the total variation distance between measures defined at* (2.2.5)*, and make $h(\xi)$ and $h(\zeta)$ integrable.*

A novel notion of distance between point process, denoted as $d_\pi$, has been recently introduced in [85] and will be described in Chapter 3, alongside with a short discussion concerning the connections between $d_{TV}$, $d_{KR}$ and $d_\pi$.

## 2.3 Stein's method

Stein's method is a technique from probability theory that allows to bound the distance of a probability measure to a target distribution. The method made its first appearance nearly fifty years ago, in the context of Normal approximation, in the seminal paper [101] by Charles Stein. Soon thereafter, in [23], Stein's PhD student Louis H.Y. Chen adapted the method so as to obtain approximation results for the Poisson distribution. Since then, the main idea and technique has been extended to an great variety of mathematical problems. Below we give a sketched idea of the method and provide a few notions that are preparatory to the following chapters. For a general introduction to the topic, the interested reader is referred to [25] and [91] and references therein.

### 2.3.1 The idea

The starting point of Stein's method is a characterizing equation for the target distribution. For the Gaussian case, we have the following: A random variable $Z$ has standard normal distribution if and only if

$$\mathbb{E}\left[f'(Z) - Zf(Z)\right] = 0 \tag{2.3.1}$$

for all absolutely continuous functions $f$ for which the expectation above exists. Roughly speaking, the idea behind Stein's method is to measure how close a distribution is to the Gaussian by estimating how close to zero is the expectation appearing in (2.3.1). This rather vague concept can be made precise by recalling that all probability metrics considered in Section 2.2.1 are of the form (2.2.1), for some suitable set $\mathcal{H}$ of test functions. This suggest to look, for each $h \in \mathcal{H}$, at the differential equation

$$f_h'(x) - xf_h(x) = h(x) - \mathbb{E}\left[h(Z)\right], \tag{2.3.2}$$

where $Z$ is now assumed to have the standard normal distribution. The above equation is called *Stein equation for the Gaussian distribution* and, for every $h \in \mathcal{H}$, there exists a unique solution which is given by

$$\begin{aligned} f_h(x) &= e^{x^2/2} \int_x^\infty e^{-t^2/2} \left(\mathbb{E}\left[h(Z)\right] - h(t)\right) \mathrm{d}t \\ &= -e^{x^2/2} \int_{-\infty}^x e^{-t^2/2} \left(\mathbb{E}\left[h(Z)\right] - h(t)\right) \mathrm{d}t\,. \end{aligned}$$

Hence, the Gaussian approximation problem takes the form

$$d_{\mathcal{H}}(X, Z) = \sup_{h \in \mathcal{H}} \mathbb{E}\left|f_h'(X) - Xf_h(X)\right|. \tag{2.3.3}$$

It is remarkable that, now that the characterization of the Gaussian distribution has been encoded into the solution of Equation (2.3.2), the right-hand side of (2.3.3) does not involve the variable $Z$ any more. Once the problem is in the latter form, one needs to find suitable techniques to compare $\mathbb{E}\left[Xf_h(X)\right]$ to $\mathbb{E}\left[f_h'(X)\right]$. This can be done in many cases by exploiting the structure of $X$ and finding good properties of the solution to the Stein equation. The following proposition, taken from [91], handles the latter issue. Recall that, for a real-valued function $f$ with domain $D$, we write $\|f\| = \sup_{x \in D} |f(x)|$.

**Proposition 2.3.1.** *Let $f_h$ be the solution of* (2.3.2).

*i) If $h$ is bounded, then*

$$\|f_h\| \leq \sqrt{\frac{\pi}{2}}\,\|h(\cdot) - \mathbb{E}[h(Z)]\| \quad \textit{and} \quad \|f_h'\| \leq 2\,\|h(\cdot) - \mathbb{E}[h(Z)]\| \,.$$

*ii) If $h$ is absolutely continuous, then*

$$\|f_h\| \leq 2\|h'\|\,, \quad \|f_h'\| \leq \sqrt{\frac{\pi}{2}}\|h'\| \quad \textit{and} \quad \|f_h''\| \leq 2\|h'\|\,.$$

Concerning the techniques used to transform the term $\mathbb{E}[Xf_h(X)]$ into something comparable to the other term, in the next subsection we will see the *size-bias coupling* applied to the context of Poisson approximation. Let us first adapt the steps seen so far to the case of Poisson distribution. As this was first done by Louis H.Y. Chen, this version of the Stein's method is often referred to as Chen-Stein method, see e.g. [5]. Standard references for the application of Stein's method to the context of Poisson approximation are [12] and [24]. We have already expressed the characterizing equation for the Poisson distribution in Proposition 2.1.10, and we recall it here: An $\mathbb{N}_0$-valued random variable $P$ has the Poisson distribution with mean $\mu$ if and only if

$$\mathbb{E}[\mu f(P+1) - Pf(P)] = 0 \tag{2.3.4}$$

for every non-negative function $f$. This characterization is reflected in the corresponding *Stein equation for Poisson distribution*

$$\mu f_h(k+1) - kf_h(k) = h(k) - \mathbb{E}[h(P)]\,, \quad k \in \mathbb{N}_0\,, \tag{2.3.5}$$

where $P$ is assumed to be Poisson distributed with mean $\mu$. Note that the value of the solution $f_h$ at zero can be chosen arbitrarily, and we adopt the convention that $f_h(0) = f_h(1)$. All other values can be computed recursively and are explicitly given by

$$\begin{aligned} f_h(k) &= \frac{(k-1)!}{\mu^k} \sum_{i=0}^{k-1} \frac{\mu^i}{i!} \left(h(i) - \mathbb{E}[h(P)]\right) \\ &= -\frac{(k-1)!}{\mu^k} \sum_{i=k}^{\infty} \frac{\mu^i}{i!} \left(h(i) - \mathbb{E}[h(P)]\right)\,. \end{aligned}$$

Therefore, the Poisson approximation problem can be translated into the task of upper bounding the following expression:

$$d_{\mathcal{H}}(X,P) = \sup_{h \in \mathcal{H}} |\mu\mathbb{E}[f_h(X+1)] - \mathbb{E}[Xf_h(X)]|\,. \tag{2.3.6}$$

The next proposition, taken from [13], provides bounds on the solution to the Stein equation for Poisson distribution.

**Proposition 2.3.2.** *Let $h : \mathbb{N}_0 \to \mathbb{R}$ be a 1-Lipschitz function. Then, the solution $f_h$ to Equation* (2.3.5) *satisfies*

$$\|f_h\| \leq 1 \quad \textit{and} \quad \|f_h(\cdot+1) - f_h(\cdot)\| \leq \min\left\{1, \frac{8}{3\sqrt{2e\mu}}\right\}.$$

### 2.3.2 Size-bias coupling and the law of small numbers

In this subsection we describe the size-bias coupling technique, that is an efficient method of rewriting $\mathbb{E}\left[Xf_h(X)\right]$ so that it can be compared to $\mu\mathbb{E}\left[f_h(X+1)\right]$.

**Definition 2.3.3.** Let $X \geq 0$ be a random variable with $\mathbb{E}\left[X\right] = \mu < \infty$. We say that a random variable $X^s$ has the *size-bias distribution* with respect to $X$ if

$$\mathbb{E}\left[Xf(X)\right] = \mu\mathbb{E}\left[f(X^s)\right] \tag{2.3.7}$$

for all $f$ such that $\mathbb{E}\left[Xf(X)\right] < \infty$.

The size-bias distribution is always well defined and it is absolutely continuous with respect to the law of $X$. From the characterization of the Poisson distribution (2.3.4) it immediately follows that for any Poisson random variable $P$, the variable $P+1$ has the size-bias distribution with respect to $P$. It is also straightforward to see that, for a generic non-negative integer-valued random variable $X$ with finite mean $\mu$, the random variable $X^s$ with the size-bias distribution of $X$ satisfies

$$\mathbb{P}\left\{X^s = k\right\} = \mu^{-1}k\mathbb{P}\left\{X = k\right\}, \quad k \in \mathbb{N}_0.$$

The notion of size-bias distribution provides yet another formulation for the Poisson approximation problem expressed by (2.3.6). Indeed, for a $\mathbb{N}_0$-valued random variable $X$ with finite mean $\mu$ and a Poisson random variable $P$ with same mean, we have

$$d_{\mathcal{H}}(X,P) = \mu \sup_{h\in\mathcal{H}} \left|\mathbb{E}f_h(X+1) - \mathbb{E}f_h(X^s)\right|.$$

Combined with Proposition 2.3.2, the latter equation yields the following upper bound for the Wasserstein distance:

$$d_W(X,P) \leq \min\left\{\mu, \frac{8\sqrt{\mu}}{3\sqrt{2e}}\right\}\mathbb{E}\left|X+1-X^s\right|. \tag{2.3.8}$$

We see through the proof of the next theorem how Equation (2.3.8) proves to be an effective formulation when dealing with the sum of independent random variables.

**Theorem 2.3.4.** *(Law of small numbers) Let $X_1, \dots X_n$ be independent Bernoulli random variables with $\mathbb{P}\{X_i = 1\} = p_i$ for $i = 1, \dots, n$, let $X = X_1 + \cdots + X_n$ and let $P$ be a Poisson random variable with mean $\mu = \mathbb{E}[X] = p_1 + \cdots + p_n$. Then*

$$d_W(X, P) \le \min\left\{1, \frac{8}{3\sqrt{2e\mu}}\right\} \sum_{i=1}^{n} p_i^2 .$$

*Proof.* Let $\tau$ be a random variable independent of $X_1, \dots X_n$ and such that $\mathbb{P}\{\tau = i\} = p_i/\mu$ for $i = 1, \dots, n$. By independence we have

$$\begin{aligned}
\mu\mathbb{E}[f(X+1-X_\tau)] &= \mu \sum_{k\in\mathbb{N}} f(k)\mathbb{P}\{X - X_\tau = k-1\} \\
&= \sum_{k\in\mathbb{N}} f(k) \sum_{i=1}^{n} \mu\mathbb{P}\{\tau = i, X - X_i = k-1\} \\
&= \sum_{k\in\mathbb{N}} f(k) \sum_{i=1}^{n} p_i\mathbb{P}\{X - X_i = k-1\} \\
&= \sum_{k\in\mathbb{N}} f(k) \sum_{i=1}^{n} \mathbb{P}\{X_i = 1, X = k\} \\
&= \sum_{k\in\mathbb{N}} f(k)\mathbb{E}\left[\sum_{i=1}^{n} \mathbf{1}\{X_i = 1\}\mathbf{1}\{X = k\}\right] \\
&= \sum_{k\in\mathbb{N}} kf(k)\mathbb{P}\{X = k\} \\
&= \mathbb{E}[Xf(X)] .
\end{aligned}$$

The above sequence of identities proves that $X + 1 - X_\tau$ has the size-bias distribution with respect to $X$. Then, Equation (2.3.8) yields

$$d_W(X, P) \le \min\left\{\mu, \frac{8\sqrt{\mu}}{3\sqrt{2e}}\right\} \mathbb{E}|X_\tau| = \min\left\{\mu, \frac{8\sqrt{\mu}}{3\sqrt{2e}}\right\} \frac{1}{\mu} \sum_{i=1}^{n} p_i^2 .$$

□

## 2.4 Gaussian approximation for sums of region stabilizing scores

In this section we report, in our notation and in a form slightly adapted to our purpose, the main result from [17], Theorem 2.1, which generalizes Theorem 2.1(a) in [58]. The latter result is obtained by incorporating stabilization methods into the so-called Malliavin–Stein theory. The concept of stabilization and the study of stabilizing functionals originated from the papers [80, 82], while the so called Malliavin-Stein theory, which combines Stein's method and Malliavin calculus, was initiated by Nourdin and Peccati in [74]. We will need the theorem stated below in Chapter 4, in order to prove our Gaussian approximation bounds.

Let $(\mathbb{X}, \mathcal{X})$ be a Borel space, $\mathbf{N}_{\mathbb{X}}$ the space of $\sigma$-finite counting measures $\mathsf{M}$ on $(\mathbb{X}, \mathcal{X})$, equipped with the smallest $\sigma$-alegbra $\mathcal{N}_{\mathbb{X}}$ such that the maps $\mathsf{M} \mapsto \mathsf{M}(A)$ are measurable for all $A \in \mathcal{X}$. We write $x \in \mathsf{M}$ if $\mathsf{M}(\{x\}) \geq 1$, and denote by $\mathsf{M}_A$ the restriction of $\mathsf{M}$ onto the set $A$, for $A \in \mathcal{X}$. Further, for $\mathsf{M}_1, \mathsf{M}_2 \in \mathbf{N}_{\mathbb{X}}$, $\mathsf{M}_1 \leq \mathsf{M}_2$ means $\mathsf{M}_2 - \mathsf{M}_1$ is non-negative.

**Definition 2.4.1.** A *score function* is a Borel measurable map $S \colon \mathbb{X} \times \mathbf{N}_{\mathbb{X}} \to \mathbb{R}$.

For a score function, we assume the following condition. If $S(x, \mathsf{M}_1) = S(x, \mathsf{M}_2)$ for some $\mathsf{M}_1, \mathsf{M}_2 \in \mathbf{N}_{\mathbb{X}}$ with $0 \neq \mathsf{M}_1 \leq \mathsf{M}_2$, then

$$S(x, \mathsf{M}_1) = S(x, \mathsf{M}'), \quad \text{for all } \mathsf{M}' \in \mathbf{N}_{\mathbb{X}},\ \mathsf{M}_1 \leq \mathsf{M}' \leq \mathsf{M}_2 \,. \tag{2.4.1}$$

Let $\eta$ denote a Poisson process on $\mathbb{X}$ with intensity measure $\mu$, and recall that, for $x \in \mathbb{X}$, $\delta_x$ denotes the Dirac measure at $x$. Theorem 2.4.2 below concerns the Gaussian approximation of the (suitably normalized) sum of score functions

$$F = F(\eta) := \sum_{x \in \eta} S(x, \eta) \,. \tag{2.4.2}$$

We need some assumptions.

(A1) There exists a map $R \colon \mathbb{X} \times \mathbf{N}_{\mathbb{X}} \to \mathcal{X}$ such that

(1)

$$\{\mathsf{M} \in \mathbf{N}_{\mathbb{X}} \colon y \in R(x, \mathsf{M} + \delta_x)\} \in \mathcal{N}_{\mathbb{X}}, \quad \text{for all } x, y \in \mathbb{X},$$

and

$$\mathbb{P}\{y \in R(x, \eta + \delta_x)\} \quad \text{and} \quad \mathbb{P}\{\{y_1, y_2\} \subseteq R(x, \eta + \delta_x)\}$$

are measurable functions of $(x, y) \in \mathbb{X}^2$ and $(x, y_1, y_2) \in \mathbb{X}^3$ respectively,

(2) the map $R$ is monotonically decreasing in the second argument, meaning

$$R(x, \mathsf{M}_1) \supseteq R(x, \mathsf{M}_2), \quad \mathsf{M}_1 \leq \mathsf{M}_2,\ x \in \mathsf{M}_1 \,,$$

(3) for all $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ and $x \in \mathsf{M}$, $\mathsf{M}_{R(x,\mathsf{M})} \neq \varnothing$ implies $(\mathsf{M} + \delta_y)_{R(x, \mathsf{M} + \delta_y)} \neq \varnothing$ for all $y \notin R(x, \mathsf{M})$,

(4) for all $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ and $x \in \mathsf{M}$,

$$S(x, \mathsf{M}) = S(x, \mathsf{M}_{R(x,\mathsf{M})}) \,;$$

(A2) there exist a $p \in (0,1]$ and a measurable function $M_p\colon \mathbb{X} \to \mathbb{R}$ such that, for all $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ with $\mathsf{M}(\mathbb{X}) \le 7$,

$$\mathbb{E}\left[S(x, \eta + \delta_x + \mathsf{M})^{4+p}\right] \le M_p(x)^{4+p}, \quad x \in \mathbb{X}.$$

Let $r\colon \mathbb{X} \times \mathbb{X} \to [0, \infty]$ be a measurable function such that

$$\mathbb{P}\{y \in R(x, \eta + \delta_x)\} \le e^{-r(x,y)}, \quad x, y \in \mathbb{X}. \tag{2.4.3}$$

For $x, y \in \mathbb{X}$ denote

$$q(x,y) := \int_{\mathbb{X}} \mathbb{P}\Big\{\{x,y\} \subseteq R\big(z, \eta + \delta_z\big)\Big\}\, \mu(\mathrm{d}z).$$

Furthermore, for $p$ as in (A.2) and $\zeta := p/(40 + 10p)$, define

$$\begin{aligned} g(y) &:= \int_{\mathbb{X}} e^{-\zeta r(x,y)}\, \mu(\mathrm{d}x), \\ G(y) &:= \widetilde{M}_p(y)\big(1 + g(y)^5\big), \quad y \in \mathbb{X}, \end{aligned}$$

where $\widetilde{M}_p(y) = \max\{M_p(y)^2, M_p(y)^4\}$, $y \in \mathbb{X}$. For $\alpha > 0$, let

$$f_\alpha(y) := f_\alpha^{(1)}(y) + f_\alpha^{(2)}(y) + f_\alpha^{(3)}(y), \quad y \in \mathbb{X},$$

where

$$\begin{aligned} f_\alpha^{(1)}(y) &:= \int_{\mathbb{X}} G(x) e^{-\alpha r(x,y)}\, \mu(\mathrm{d}x), \\ f_\alpha^{(2)}(y) &:= \int_{\mathbb{X}} G(x) e^{-\alpha r(y,x)}\, \mu(\mathrm{d}x), \\ f_\alpha^{(3)}(y) &:= \int_{\mathbb{X}} G(x) q(x,y)^\alpha\, \mu(\mathrm{d}x). \end{aligned}$$

Finally, let

$$\kappa(x) := \mathbb{P}\{S(x, \eta + \delta_x) \ne 0\}, \quad x \in \mathbb{X}.$$

For an integrable function $f : \mathbb{X} \to \mathbb{R}$, denote $\mu f := \int_{\mathbb{X}} f(x)\mu(\mathrm{d}x)$. Recall the definitions of Wasserstein and Kolmogorov distances given at (2.2.3) and (2.2.2), respectively.

**Theorem 2.4.2.** *Assume that $S$ satisfies conditions (A1), (A2) and let $F$ be as in* (2.4.2). *Then, for $p$ as in (A2), with $\beta := \frac{p}{32+4p}$,*

$$d_W\left(\frac{F - \mathbb{E}F}{\sqrt{\operatorname{Var} F}}, N\right) \le C\left[\frac{\sqrt{\mu f_\beta^2}}{\operatorname{Var} F} + \frac{\mu((\kappa + g)^{2\beta} G)}{(\operatorname{Var} F)^{3/2}}\right],$$

*and*

$$d_K\left(\frac{F-\mathbb{E}F}{\sqrt{\operatorname{Var} F}}, N\right) \leq C\Bigg[\frac{\sqrt{\mu f_\beta^2}+\sqrt{\mu f_{2\beta}}}{\operatorname{Var} F}+\frac{\sqrt{\mu((\kappa+g)^{2\beta}G)}}{\operatorname{Var} F}$$
$$+\frac{\mu((\kappa+g)^{2\beta}G)}{(\operatorname{Var} F)^{3/2}}+\frac{(\mu((\kappa+g)^{2\beta}G))^{5/4}+(\mu((\kappa+g)^{2\beta}G))^{3/2}}{(\operatorname{Var} F)^2}\Bigg],$$

*where $N$ is a standard normal random variable and $C \in (0,\infty)$ is a constant depending only on $p$.*

# Chapter 3

# Multivariate Poisson and Poisson Process Approximations

This chapter is based on the following article:

*F. Pianoforte and R. Turin. Multivariate Poisson and Poisson process approximations with applications to Bernoulli sums and $U$-statistics. arXiv:2105.01599, 2021.*

The layout of the chapter is as follows. In Section 3.1 we adapt the Chen-Stein method described in Section 2.3 to the multivariate setting, hence obtaining a general bound on the Wasserstain distance between an integer-valued random vector and a Poisson random vector with the same mean. The result from Section 3.1 is then employed in Section 3.2, where we Poisson approximate the sum of dependent Bernoulli random vectors. In Section 3.3 we introduce a novel notion of metric between point processes distributions, $d_\pi$, heavily relying on the Wasserstein distance between multivariate distributions. After a small discussion concerning the connections between $d_\pi$ and the other metrics defined in Chapter 2, we derive a limit theorem for the Poisson process approximation in terms of the newly defined metric. The main theorem from Section 3.3 is then applied in Sections 3.4 and 3.5 to provide explicit Poisson process approximation results for point processes with Papangelou intensity and point processes of Poisson $U$-statistic structure, respectively.

## 3.1 Multivariate Poisson approximation

Let $\mathbf{X} = (X_1, \dots, X_d)$ be an integrable random vector taking values in $\mathbb{N}_0^d$, $d \in \mathbb{N}$, and let $\mathbf{P} = (P_1, \dots, P_d)$ be a Poisson random vector, that is, a random vector with independent and Poisson distributed components. In this section, we provide an upper bound on the Wasserstein distance

$$d_W(\mathbf{X},\mathbf{P}) = \sup_{g\in \mathrm{Lip}^d(1)} \left|\mathbb{E}[g(\mathbf{X})] - \mathbb{E}[g(\mathbf{P})]\right|.$$

between $\mathbf{X}$ and $\mathbf{P}$. Recall that, for any $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$ and indexes $1 \le i < j \le d$, we denote by $x_{i:j}$ the sub-vector $(x_i,\dots,x_j)$.

**Theorem 3.1.1.** *Let $\mathbf{X} = (X_1,\dots,X_d)$ be an integrable random vector with values in $\mathbb{N}_0^d$, $d\in\mathbb{N}$, and let $\mathbf{P} = (P_1,\dots,P_d)$ be a Poisson random vector with $\mathbb{E}[\mathbf{P}] = (\mu_1,\dots,\mu_d) \in [0,\infty)^d$. For $1 \le i \le d$, consider any random vector $\mathbf{Z}^{(i)} = (Z_1^{(i)},\dots,Z_i^{(i)})$ in $\mathbb{Z}^i$ defined on the same probability space as $\mathbf{X}$, and define*

$$q_{m_{1:i}} \coloneqq m_i\mathbb{P}\big(X_{1:i} = m_{1:i}\big) - \mu_i\mathbb{P}\big(X_{1:i} + \mathbf{Z}^{(i)} = (m_{1:i-1}, m_i - 1)\big) \tag{3.1.1}$$

*for $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$. Then,*

$$d_W(\mathbf{X},\mathbf{P}) \le \sum_{i=1}^d \left(\mu_i\mathbb{E}\big|Z_i^{(i)}\big| + 2\mu_i\sum_{j=1}^{i-1}\mathbb{E}\big|Z_j^{(i)}\big| + \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}} |q_{m_{1:i}}|\right). \tag{3.1.2}$$

It should be noted that a slightly improved bound than (3.1.2) can be easily obtained, and this is expressed below in Remark 3.1.3.

In order to give an interpretation of Equation (3.1.1), let us consider the random vectors

$$\mathbf{Y}^{(i)} = (X_{1:i-1}, X_i + 1) + \mathbf{Z}^{(i)}, \quad i = 1,\dots,d, \tag{3.1.3}$$

with $\mathbf{X}$ and $\mathbf{Z}^{(i)}$ defined as in Theorem 3.1.1. Under the additional condition $\mathbb{P}(X_{1:i}+\mathbf{Z}^{(i)} \in \mathbb{N}_0^i) = 1$, a sequence of real numbers $q_{m_{1:i}}$, $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$, satisfies Equation (3.1.1) if and only if

$$\mathbb{E}[X_i f(X_{1:i})] = \mu_i\mathbb{E}[f(\mathbf{Y}^{(i)})] + \sum_{m_{1:i}\in\mathbb{N}_0^i,\, m_i\neq 0} q_{m_{1:i}} f(m_{1:i}) \tag{3.1.4}$$

for all functions $f : \mathbb{N}_0^i \to \mathbb{R}$ such that $\mathbb{E}\,|X_i f(X_{1:i})| < \infty$, where to prove that (3.1.4) implies (3.1.1) it is enough to consider $f$ to be the function with value 1 at $m_{1:i}$ and 0 elsewhere. When the $q_{m_{1:i}}$ are all zeros and $\mathbb{E}[X_i] = \mu_i$, the condition $\mathbb{P}(X_{1:i} + \mathbf{Z}^{(i)} \in \mathbb{N}_0^i) = 1$ is satisfied, as can be seen by taking the sum over $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$ in (3.1.1). In this case, (3.1.4) becomes

$$\mathbb{E}[X_i f(X_{1:i})] = \mathbb{E}[X_i]\mathbb{E}[f(\mathbf{Y}^{(i)})]. \tag{3.1.5}$$

The last equation, for $d = 1$, corresponds to the equation appearing in Definition 2.3.3. Therefore, if for some $1 \le i \le d$ the $q_{m_{1:i}}$ are all zeros and $\mathbb{E}[X_i] = \mu_i$, the distribution of the random vector $\mathbf{Y}^{(i)}$ can be seen as the size bias distribution of the vector $X_{1:i}$. Following this interpretation, when $\mathbb{E}[\mathbf{X}] = (\mu_1, \ldots, \mu_d)$ and the random vectors $\mathbf{Z}^{(i)}$ are chosen such that the $q^{(i)}_{m_{1:i}}$ are not zero, we can think of the distribution of $\mathbf{Y}^{(i)}$ defined by (3.1.3) as an approximate size bias distribution of $X_{1:i}$, where instead of assuming that $\mathbf{Y}^{(i)}$ satisfies (3.1.5) exactly, we allow error terms $q_{m_{1:i}}$. This is an important advantage of Theorem 3.1.1, since one does not need to find random vectors with an exact size bias distribution (in the sense of (3.1.5)), it only matters that the error terms $q^{(i)}_{m_{1:i}}$ are sufficiently small and that the random vectors $\mathbf{Z}^{(i)}$ are null with high probability.

The proof of Theorem 3.1.1 is based on the Chen-Stein method described in Section 2.3 applied to each component of the random vectors and combined with the approximate coupling expressed in (3.1.1). Without loss of generality we may assume that $\mathbf{X}$ and $\mathbf{P}$ are independent and defined on the same reference probability space $(\Omega, \mathfrak{F}, \mathbb{P})$.

For any fixed $h \in \mathrm{Lip}(1)$, we now denote by $\widehat{h}^{(\mu)}$ (instead of $f_h$) the solution of Stein's equation (2.3.5). In this way we highlight the dependency on the parameter lambda and also leave the subscript area free for further notation. By Proposition 2.3.2 we have

$$\sup_{i\in\mathbb{N}_0} \left|\widehat{h}^{(\mu)}(i)\right| \leq 1 \quad \text{and} \quad \sup_{i\in\mathbb{N}_0} \left|\widehat{h}^{(\mu)}(i+1) - \widehat{h}^{(\mu)}(i)\right| \leq \min\left\{1, \frac{8}{3\sqrt{2e\mu}}\right\}. \tag{3.1.6}$$

For $h \in \mathrm{Lip}^d(1)$, let $\widehat{h}^{(\mu)}_{x_{1:i-1}|x_{i+1:d}}$ denote the solution to (2.3.5) for the Lipschitz function $h(x_{1:i-1}, \cdot\ , x_{i+1:d})$ with fixed $x_{1:i-1} \in \mathbb{N}_0^{i-1}$ and $x_{i+1:d} \in \mathbb{N}_0^{d-i}$. Since $\widehat{h}^{(\mu)}$ takes vectors from $\mathbb{N}_0^d$ as input, we do not need to worry about measurability issues. The following proposition is the first building block for the proof of Theorem 3.1.1.

**Proposition 3.1.2.** *For any* $h \in \mathrm{Lip}^d(1)$,

$$\mathbb{E}[h(\mathbf{P}) - h(\mathbf{X})] = \sum_{i=1}^{d} \mathbb{E}\left[X_i \widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i) - \mu_i \widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i+1)\right].$$

*Proof of Proposition 3.1.2.* First, observe that

$$\mathbb{E}\left[h(\mathbf{P}) - h(\mathbf{X})\right] = \sum_{i=1}^{d} \mathbb{E}\left[h(X_{1:i-1}, P_{i:d}) - h(X_{1:i}, P_{i+1:d})\right], \tag{3.1.7}$$

with the conventions $(X_{1:0}, P_{1:d}) = \mathbf{P}$ and $(X_{1:d}, P_{d+1:d}) = \mathbf{X}$. The independence of $P_i$ from $P_{i+1:d}$ and $X_{1:i}$ implies

$$\mathbb{E}\big[h(X_{1:i-1}, P_{i:d}) - h(X_{1:i}, P_{i+1:d})\big] = \mathbb{E}\big[\mathbb{E}^{P_i}[h(X_{1:i-1}, P_{i:d})] - h(X_{1:i}, P_{i+1:d})\big],$$

where $\mathbb{E}^{P_i}$ denotes the expectation with respect to the random variable $P_i$. From the definition of $\widehat{h}^{(\mu_i)}_{x_{1:i-1}|x_{i+1:d}}$ with $x_{1:i-1} = X_{i:i-1}$ and $x_{i+1:d} = P_{i+1:d}$, it follows

$$\mathbb{E}^{P_i}[h(X_{1:i-1}, P_{i:d})] - h(X_{1:i}, P_{i+1:d}) = X_i \widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i) - \mu_i \widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i+1)$$

for all $i = 1, \ldots, d$. Together with (3.1.7), this leads to the desired conclusion. ☐

*Proof of Theorem 3.1.1.* In view of Proposition 3.1.2, it suffices to bound

$$\left|\mathbb{E}\left[X_i\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i)-\mu_i\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i+1)\right]\right|,\quad i=1,\dots,d\,.$$

For the remainder of the proof, the index $i$ is fixed and we omit the superscript $(i)$ in $Z^{(i)}_{1:i}$. Define the function $\widetilde{h}\colon\mathbb{N}_0^i\to\mathbb{R}$ such that

$$\widetilde{h}(X_{1:i})=\mathbb{E}\left[\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i)\,\middle|\,X_{1:i}\right],$$

where $\mathbb{E}[\,\cdot\,|\,Y]$ denotes the conditional expectation with respect to a random element $Y$. With the convention $\widehat{h}^{(\mu_i)}_{m_{1:i-1}|m_{i+1:d}}(m_i)=0$ if $m_{1:d}\notin\mathbb{N}_0^d$, it follows from (3.1.1) that

$$\begin{aligned}
\mathbb{E}\left[X_i\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i)\right]&=\mathbb{E}[X_i\widetilde{h}(X_{1:i})]=\sum_{m_{1:i}\in\mathbb{N}_0^i}m_i\widetilde{h}(m_{1:i})\mathbb{P}(X_{1:i}=m_{1:i})\\
&=\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}}\widetilde{h}(m_{1:i})q_{m_{1:i}}+\mu_i\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}}\widetilde{h}(m_{1:i})\mathbb{P}\left(X_{1:i}+Z_{1:i}=(m_{1:i-1},m_i-1)\right)\\
&=\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}}\widetilde{h}(m_{1:i})q_{m_{1:i}}+\mu_i\mathbb{E}\left[\widehat{h}^{(\mu_i)}_{X_{1:i-1}+Z_{1:i-1}|P_{i+1:d}}(X_i+Z_i+1)\right].
\end{aligned}$$

Since $|\widetilde{h}(X_{1:i})|\leq 1$ by (3.1.6), the triangle inequality establishes

$$\left|\mathbb{E}\left[X_i\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i)-\mu_i\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i+1)\right]\right|\leq\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}}|q_{m_{1:i}}|+\mu_i(H_1+H_2),\tag{3.1.8}$$

with

$$H_1=\left|\mathbb{E}\left[\widehat{h}^{(\mu_i)}_{X_{1:i-1}+Z_{1:i-1}|P_{i+1:d}}(X_i+Z_i+1)-\widehat{h}^{(\mu_i)}_{X_{1:i-1}+Z_{1:i-1}|P_{i+1:d}}(X_i+1)\right]\right|$$

and

$$H_2=\left|\mathbb{E}\left[\widehat{h}^{(\mu_i)}_{X_{1:i-1}+Z_{1:i-1}|P_{i+1:d}}(X_i+1)-\widehat{h}^{(\mu_i)}_{X_{1:i-1}|P_{i+1:d}}(X_i+1)\right]\right|.$$

The inequalities in (3.1.6) guarantee

$$H_1\leq\mathbb{E}|Z_i|\quad\text{and}\quad H_2\leq 2\mathbb{P}(Z_{1:i-1}\neq 0)\leq\sum_{j=1}^{i-1}2\mathbb{P}(Z_j\neq 0)\leq 2\sum_{j=1}^{i-1}\mathbb{E}|Z_j|.$$

Combining (3.1.8) with the bounds for $H_1$ and $H_2$, and summing over $i=1,\dots,d$ concludes the proof. □

*Remark* 3.1.3. It follows directly from the previous proof that the term $\sum_{j=1}^{i-1}\mathbb{E}|Z_j|$ in (3.1.2) could be replaced by $\mathbb{P}(Z_{1:i-1}\neq 0)$. Moreover, applying the more refined bound from (3.1.6) yields

$$H_1\leq\min\left\{1,\frac{8}{3\sqrt{2e\mu_i}}\right\}\mathbb{E}|Z_i|\,.$$

These two observations together lead to the improved bound for Theorem 3.1.1:

$$d_W(\mathbf{X},\mathbf{P}) \le \sum_{i=1}^{d} \left( \min\left\{\mu_i, \frac{8\sqrt{\mu_i}}{3\sqrt{2e}}\right\} \mathbb{E}\big|Z_i^{(i)}\big| + 2\mu_i \mathbb{P}\big(Z_{1:i-1}^{(i)} \neq 0\big) + \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i \\ m_i\neq 0}} |q_{m_{1:i}}| \right).$$

## 3.2 Sum of $m$-dependent Bernoulli random vectors

In this section, we consider a finite family of Bernoulli random vectors $\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(n)}$ and investigate the multivariate Poisson approximation of $\mathbf{X} = \sum_{r=1}^n \mathbf{Y}^{(r)}$ in the Wasserstein distance. The distributions of $\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(n)}$ are given by

$$\begin{aligned} &\mathbb{P}(\mathbf{Y}^{(r)} = \mathbf{e}_j) = p_{r,j} \in [0,1], \quad r=1,\dots,n\,, \quad j=1,\dots,d, \\ &\mathbb{P}(\mathbf{Y}^{(r)} = \mathbf{0}) = 1 - \sum_{j=1}^{d} p_{r,j} \in [0,1], \quad r=1,\dots,n, \end{aligned} \tag{3.2.1}$$

where $\mathbf{e}_j$ denotes the vector with entry 1 at position $j$ and entry 0 otherwise. If the Bernoulli random vectors are i.i.d., $\mathbf{X}$ has the so called multinomial distribution. The multivariate Poisson approximation of the multinomial distribution, and more generally of the sum of independent Bernoulli random vectors, has already been tackled by many authors in terms of the total variation distance. Among others, we refer the reader to [10, 32, 88, 90] and the survey [75]. Unlike the mentioned papers, we assume that $\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(n)}$ are $m$-dependent. Note that the case of sums of 1-dependent random vectors has recently been treated in [35] using metrics that are weaker than the total variation distance. To the best of our knowledge, this is the first time that the Poisson approximation of the sum of $m$-dependent Bernoulli random vectors is investigated in terms of the Wasserstein distance.

More precisely, for $n \in \mathbb{N}$, let $\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(n)}$ be Bernoulli random vectors with distributions given by (3.2.1), and assume that for a given fixed $m \in \mathbb{N}_0$ and any two subsets $S$ and $T$ of $\{1,\dots,n\}$ such that $\min(S) - \max(T) > m$, the collections $(\mathbf{Y}^{(s)})_{s\in S}$ and $(\mathbf{Y}^{(t)})_{t\in T}$ are independent. Define the random vector $\mathbf{X} = (X_1,\dots,X_d)$ as

$$\mathbf{X} = \sum_{r=1}^{n} \mathbf{Y}^{(r)}. \tag{3.2.2}$$

Note that if $\mathbf{Y}^{(r)}, r=1,\dots,n$, are i.i.d., then $m=0$ and $\mathbf{X}$ has the multinomial distribution. The mean vector of $\mathbf{X}$ is $\mathbb{E}[\mathbf{X}] = (\mu_1,\dots,\mu_d)$ with

$$\mu_j = \sum_{r=1}^{n} p_{r,j}, \quad j=1,\dots,d. \tag{3.2.3}$$

For $k=1,\dots,n$ and $m\geq 1$ let $Q(k)$ be the quantity given by

$$Q(k)=\max_{\substack{r\in\{1,\dots,n\}\,:\,1\leq|k-r|\leq m\\ i,j=1,\dots,d}}\mathbb{E}\big[\mathbf{1}\{\mathbf{Y}^{(k)}=\mathbf{e}_i\}\mathbf{1}\{\mathbf{Y}^{(r)}=\mathbf{e}_j\}\big].$$

We now state the main result of this subsection.

**Theorem 3.2.1.** *Let* $\mathbf{X}$ *be as in* (3.2.2), *and let* $\mathbf{P}=(P_1,\dots,P_d)$ *be a Poisson random vector with mean* $\mathbb{E}[\mathbf{P}]=(\mu_1,\dots,\mu_d)$ *given by* (3.2.3). *Then,*

$$d_W(\mathbf{X},\mathbf{P})\leq\sum_{k=1}^{n}\sum_{i=1}^{d}\Bigg[\sum_{\substack{r=1,\dots,n,\\|r-k|\leq m}}p_{r,i}+2\sum_{j=1}^{i-1}\sum_{\substack{r=1,\dots,n,\\|r-k|\leq m}}p_{r,j}\Bigg]p_{k,i}+2d(d+1)m\sum_{k=1}^{n}Q(k).$$

The proof of Theorem 3.2.1 is obtained by applying Theorem 3.1.1. When $d=1$, Equation (3.1.1) corresponds to the condition required in [84, Theorem 1.2], which establishes sharper Poisson approximation results than the one obtained in the univariate case from Theorem 3.1.1. Therefore, for the sum of dependent Bernoulli random variables, a sharper bound for the Wasserstein distance can be derived from [84, Theorem 1.2], while for the total variation distance a bound may be deduced from [5, Theorem 1], [84, Theorem 1.2] or [100, Theorem 1].

As a consequence of Theorem 3.2.1, we obtain the following result for the sum of independent Bernoulli random vectors.

**Corollary 3.2.2.** *For* $n\in\mathbb{N}$, *let* $\mathbf{Y}^{(1)},\dots,\mathbf{Y}^{(n)}$ *be independent Bernoulli random vectors with distribution given by* (3.2.1), *and let* $\mathbf{X}$ *be the random vector defined by* (3.2.2). *Let* $\mathbf{P}=(P_1,\dots,P_d)$ *be a Poisson random vector with mean* $\mathbb{E}[\mathbf{P}]=(\mu_1,\dots,\mu_d)$ *given by* (3.2.3). *Then*

$$d_W(\mathbf{X},\mathbf{P})\leq\sum_{k=1}^{n}\Bigg[\sum_{i=1}^{d}p_{k,i}\Bigg]^2.$$

In [88, Theorem 1], a sharper bound for the total variation distance than the one obtained by Corollary 3.2.2 is proven. When the vectors are identically distributed and $\sum_{j=1}^{d}p_{1,j}\leq\alpha/n$ for some constant $\alpha>0$, our bound for the Wasserstein distance and the one in [88, Theorem 1] for the total variation distance only differ by a constant that does not depend on $n$, $d$ and the probabilities $p_{i,j}$.

*Proof of Theorem 3.2.1.* Without loss of generality we may assume that $\mu_1,\dots,\mu_d>0$. Define the random vectors

$$\mathbf{W}^{(k)}=\big(W_1^{(k)},\dots,W_d^{(k)}\big)=\sum_{\substack{r=1,\dots,n,\\1\leq|r-k|\leq m}}\mathbf{Y}^{(r)},$$

$$\mathbf{X}^{(k)}=\big(X_1^{(k)},\dots,X_d^{(k)}\big)=\mathbf{X}-\mathbf{Y}^{(k)}-\mathbf{W}^{(k)},$$

for $k = 1, \dots, n$. Let us fix $1 \le i \le d$ and $\ell_{1:i} \in \mathbb{N}_0^i$ with $\ell_i \neq 0$. From straightforward calculations it follows that

$$\begin{aligned} \ell_i \mathbb{P}(X_{1:i} = \ell_{1:i}) &= \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \mathbf{1}\{X_{1:i} = \ell_{1:i}\} \\ &= \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \mathbf{1}\big\{X_{1:i}^{(k)} + W_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\}. \end{aligned} \tag{3.2.4}$$

Let $H_{\ell_{1:i}}$ and $q_{\ell_{1:i}}$ be the quantities given by

$$\begin{aligned} H_{\ell_{1:i}} &\coloneqq \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \mathbf{1}\big\{X_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\}, \\ q_{\ell_{1:i}} &\coloneqq \ell_i \mathbb{P}(X_{1:i} = \ell_{1:i}) - H_{\ell_{1:i}}. \end{aligned}$$

For $i = 1, \dots, d$, let $\tau_i$ be a random variable independent of $(\mathbf{Y}^{(r)})_{r=1}^n$ with distribution

$$\mathbb{P}(\tau_i = k) = p_{k,i}/\mu_i, \quad k = 1, \dots, n\,.$$

Since $\mathbf{Y}^{(r)}$, $r = 1, \dots, n$, are $m$-dependent, the random vectors $\mathbf{Y}^{(k)} = (Y_1^{(k)}, \dots, Y_d^{(k)})$ and $\mathbf{X}^{(k)}$ are independent for all $k = 1, \dots, n$. Therefore

$$\begin{aligned} H_{\ell_{1:i}} &= \sum_{k=1}^{n} p_{k,i} \mathbb{P}\big(X_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big) \\ &= \sum_{k=1}^{n} p_{k,i} \mathbb{P}\big(X_{1:i} - W_{1:i}^{(k)} - Y_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big) \\ &= \mu_i \mathbb{P}\big(X_{1:i} - W_{1:i}^{(\tau_i)} - Y_{1:i}^{(\tau_i)} = (\ell_{1:i-1}, \ell_i - 1)\big). \end{aligned}$$

Then, by Theorem 3.1.1 we obtain

$$d_W(\mathbf{X}, \mathbf{P}) \le \sum_{i=1}^{d} \left( \mu_i \mathbb{E}\left[W_i^{(\tau_i)} + Y_i^{(\tau_i)}\right] + 2\mu_i \sum_{j=1}^{i-1} \mathbb{E}\left[W_j^{(\tau_i)} + Y_j^{(\tau_i)}\right] + \sum_{\substack{\ell_{1:i} \in \mathbb{N}_0^d \\ \ell_i \neq 0}} \left|q_{\ell_{1:i}}\right| \right). \tag{3.2.5}$$

From (3.2.4) and the definition of $q_{\ell_{1:i}}$ it follows that

$$\begin{aligned} |q_{\ell_{1:i}}| &\le \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \left| \mathbf{1}\big\{X_{1:i}^{(k)} + W_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\} - \mathbf{1}\big\{X_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\} \right| \\ &\le \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \mathbf{1}\{W_{1:i}^{(k)} \neq 0\} \mathbf{1}\big\{X_{1:i}^{(k)} + W_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\} \\ &\quad + \mathbb{E} \sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\} \mathbf{1}\{W_{1:i}^{(k)} \neq 0\} \mathbf{1}\big\{X_{1:i}^{(k)} = (\ell_{1:i-1}, \ell_i - 1)\big\}. \end{aligned}$$

Thus, by the inequality $\mathbf{1}\{W_{1:i}^{(k)} \neq 0\} \leq \sum_{j=1}^{i} W_j^{(k)}$ we obtain

$$\begin{aligned}\sum_{\substack{\ell_{1:i}\in\mathbb{N}_0^i\\ \ell_i\neq 0}} \left|q_{\ell_{1:i}}\right| &\leq 2\mathbb{E}\sum_{k=1}^{n} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\}\mathbf{1}\{W_{1:i}^{(k)} \neq 0\} \\ &\leq 2\mathbb{E}\sum_{k=1}^{n}\sum_{j=1}^{i} \mathbf{1}\{\mathbf{Y}^{(k)} = \mathbf{e}_i\}W_j^{(k)} \leq 4mi\sum_{k=1}^{n} Q(k).\end{aligned} \tag{3.2.6}$$

Moreover, for any $i, j = 1, \ldots, d$ we have

$$\begin{aligned}\mu_i\mathbb{E}\Big[W_j^{(\tau_i)} + Y_j^{(\tau_i)}\Big] &= \mu_i\mathbb{E}\sum_{\substack{r=1,\ldots,n,\\ |r-\tau_i|\leq m}} \mathbf{1}\{\mathbf{Y}^{(r)} = \mathbf{e}_j\} \\ &= \sum_{k=1}^{n} p_{k,i}\,\mathbb{E}\sum_{\substack{r=1,\ldots,n,\\ |r-k|\leq m}} \mathbf{1}\{\mathbf{Y}^{(r)} = \mathbf{e}_j\} = \sum_{\substack{k,r=1,\ldots,n,\\ |r-k|\leq m}} p_{k,i}p_{r,j}.\end{aligned}$$

Together with (3.2.5) and (3.2.6), this leads to

$$\begin{aligned}d_W(\mathbf{X}, \mathbf{P}) &\leq \sum_{i=1}^{d}\sum_{\substack{k,r=1,\ldots,n,\\ |r-k|\leq m}} p_{k,i}p_{r,i} + 2\sum_{i=1}^{d}\sum_{j=1}^{i-1}\sum_{\substack{k,r=1,\ldots,n,\\ |r-k|\leq m}} p_{k,i}p_{r,j} + 2d(d+1)m\sum_{k=1}^{n} Q(k) \\ &= \sum_{k=1}^{n}\sum_{i=1}^{d}\Bigg[\sum_{\substack{r=1,\ldots,n,\\ |r-k|\leq m}} p_{r,i} + 2\sum_{j=1}^{i-1}\sum_{\substack{r=1,\ldots,n,\\ |r-k|\leq m}} p_{r,j}\Bigg]p_{k,i} + 2d(d+1)m\sum_{k=1}^{n} Q(k),\end{aligned}$$

which completes the proof. □

## 3.3 Poisson process approximation

The Poisson process approximation has mostly been studied in terms of the total variation distance (see Equation (2.2.4)) in the literature; see e.g. [6, 9, 11, 21, 26, 95, 96] and references therein. In contrast, [30, 31] deal with Poisson process approximation using the Kantorovich–Rubinstein distance (see Equation (2.2.6)). In [85], the authors introduce a novel definition of distributional metric for the class of point processes with finite intensity measure. The idea is that of evaluating the Wasserstein distance between the finite dimensional distributions of the point processes indexed by arbitrary collections of disjoint measurable sets.

**Definition 3.3.1.** Let $(\mathbb{X}, \mathcal{X})$ be a measurable space, and let $\xi$ and $\zeta$ be point processes on $\mathbb{X}$ with finite intensity measure. The distance $d_\pi$ between the distributions of $\xi$ and $\zeta$ is defined as

$$d_\pi(\xi, \zeta) := \sup_{(A_1,\ldots,A_d)\in\mathcal{X}_{\shortparallel}^d,\, d\in\mathbb{N}} d_W\big((\xi(A_1), \ldots, \xi(A_d)), (\zeta(A_1), \ldots, \zeta(A_d))\big),$$

where

$$\mathcal{X}_{\shortparallel}^{d} = \{(A_1, \dots, A_d) \in \mathcal{X}^d \,:\, A_i \cap A_j = \emptyset, i \neq j\}.$$

The function $d_\pi$ is a well defined distance between the distributions of point processes with finite intensity measure, as it immediately follows by Proposition 2.1.5. To the best of the author's knowledge, this is the first time the distance $d_\pi$ is defined and employed in Poisson process approximation. We believe that it is possible to extend $d_\pi$ to larger classes of point processes by restricting $\mathcal{X}_{\shortparallel}^{d}$ to suitable families of sets. For example, for locally finite point processes on a locally compact second countable Hausdorff space (lcscH), we may define the distance $d_\pi$ by replacing $\mathcal{X}_{\shortparallel}^{d}$ with the family of $d$-tuples of disjoint and relatively compact Borel sets. An interesting property of $d_\pi$, is that it dominates the total variation distance defined at (2.2.4), and it is actually stronger than the latter.

**Proposition 3.3.2.** *Let $\xi$ and $\zeta$ be two point processes on $\mathbb{X}$ with finite intensity measure. Then,*

$$d_{TV}(\xi, \zeta) \leq d_\pi(\xi, \zeta).$$

*Moreover, there are sequences of point processes with finite intensity measure that converge in total variation distance, but are not Cauchy sequences with respect to $d_\pi$.*

The result is obtained by a monotone class Theorem, [64, Theorem 1.3], which is stated hereafter as a Lemma. Recall that a *monotone class* $\mathcal{A}$ is a collection of sets closed under monotone limits, that is, for any $A_1, A_2, \dots \in \mathcal{A}$ with $A_n \uparrow A$ or $A_n \downarrow A$, then $A \in \mathcal{A}$.

**Lemma 3.3.3.** *Let $U$ be a set and let $\mathcal{U}$ be an algebra of subsets of $U$. Then, the monotone class generated by $\mathcal{U}$ coincides with the $\sigma$-field generated by $\mathcal{U}$.*

*Proof of Proposition 3.3.2.* Let us first introduce the set of finite counting measures

$$\widehat{\mathbf{N}}_{\mathbb{X}} = \{\mathsf{M} \in \mathbf{N}_{\mathbb{X}} : \mathsf{M}(\mathbb{X}) < \infty\}, \tag{3.3.1}$$

with the trace $\sigma$-field

$$\widehat{\mathcal{N}}_{\mathbb{X}} = \{B \cap \widehat{\mathbf{N}}_{\mathbb{X}} : B \in \mathcal{N}_{\mathbb{X}}\}. \tag{3.3.2}$$

As we are dealing with finite point processes, the total variation distance is equivalently obtained if $\mathcal{N}_X$ is replaced by $\widehat{\mathcal{N}}_{\mathbb{X}}$ in (2.2.4):

$$d_{TV}(\xi, \zeta) = \sup_{B \in \widehat{\mathcal{N}}_{\mathbb{X}}} |\mathbb{P}(\xi \in B) - \mathbb{P}(\zeta \in B)|.$$

Let $\mathcal{P}(\mathbb{N}_0^d)$ denote the power set of $\mathbb{N}_0^d$, that is, the collection of all subsets of $\mathbb{N}_0^d$. For any $d \in \mathbb{N}$ and $M \in \mathcal{P}(\mathbb{N}_0^d)$ note that $\mathbf{1}_M(\cdot) \in \mathrm{Lip}^d(1)$, therefore

$$d_\pi(\xi, \zeta) \geq \sup_{U \in \mathcal{U}} |\mathbb{P}(\xi \in U) - \mathbb{P}(\zeta \in U)| \,, \tag{3.3.3}$$

with

$$\mathcal{U} = \left\{ \left\{ \mathsf{M} \in \widehat{\mathbf{N}}_{\mathbb{X}} : (\mathsf{M}(A_1), \dots, \mathsf{M}(A_d)) \in M \right\} : \, d \in \mathbb{N}, \, (A_1, \dots, A_d) \in \mathcal{X}_{\shortparallel}^d, \; M \in \mathcal{P}(\mathbb{N}_0^d) \right\}.$$

It can be easily verified that $\mathcal{U}$ is an algebra, $\mathcal{U} \subseteq \widehat{\mathcal{N}}_{\mathbb{X}}$ and $\sigma(\mathcal{U}) = \widehat{\mathcal{N}}_{\mathbb{X}}$. Moreover, by (3.3.3), $\mathcal{U}$ is a subset of the monotone class

$$\left\{ U \in \widehat{\mathcal{N}}_{\mathbb{X}} : |\mathbb{P}(\xi \in U) - \mathbb{P}(\zeta \in U)| \leq d_\pi(\xi, \zeta) \right\}.$$

Lemma 3.3.3 concludes the first part of the proof.

The second part of the statement can be seen by the following simple example: Let $Y_1, Y_2, \dots$ be independent Bernulli random variables, with $\mathbb{P}\{Y_n = 1\} = 1/n$, and let $\xi_n := n^2 \mathbf{1}\{Y_n = 1\}\delta_x$, for $n \in \mathbb{N}$ and some fixed $x \in \mathbb{X}$. Denote by $\varnothing$ the null-measure on $\mathbb{X}$. Then $\mathbb{P}\{\xi_n = \varnothing\} = 1/n$ implies $d_{TV}(\xi_n, \varnothing) \leq 1/n \to 0$, while $d_\pi(\xi_n, \varnothing) \geq |\mathbb{E}[\xi_n(\mathbb{X})]| = n \to \infty$. □

A second interesting relation between the above defined metrics is that, when $\mathbb{X}$ is lcscH, $d_\pi$ is dominated by $2d_{KR}$. It remains an open problem whether the two distances are equivalent or not.

**Proposition 3.3.4.** *Let $\xi$ and $\zeta$ be two point processes with finite intensity measure on a lcscH space $\mathbb{X}$. Then*

$$d_\pi(\xi, \zeta) \leq 2 d_{KR}(\xi, \zeta) \,.$$

*Proof.* Take $g \in \mathrm{Lip}^d(1)$ and disjoint sets $A_1, \dots, A_d \in \mathcal{X}, d \in \mathbb{N}$, define $h : \widehat{\mathbf{N}}_{\mathbb{X}} \to \mathbb{R}$ by $h(\mathsf{M}) = g(\mathsf{M}(A_1), \dots, \mathsf{M}(A_d))$, with $\widehat{\mathbf{N}}_{\mathbb{X}}$ defined at (3.3.1). For finite point configurations $\mathsf{M}_1$ and $\mathsf{M}_2$, we obtain

$$\begin{aligned} |h(\mathsf{M}_1) - h(\mathsf{M}_2)| &= |g(\mathsf{M}_1(A_1), \dots, \mathsf{M}_1(A_d)) - g(\mathsf{M}_2(A_1), \dots, \mathsf{M}_2(A_d))| \\ &\leq \sum_{i=1}^d |\mathsf{M}_1(A_i) - \mathsf{M}_2(A_i)| \leq 2 d_{TV}(\mathsf{M}_1, \mathsf{M}_2). \end{aligned}$$

By [68, Theorem 1], there exists a function $\tilde{h}$ on $\widetilde{\mathbf{N}}_{\mathbb{X}}$ that coincides with $h$ on $\widehat{\mathbf{N}}_{\mathbb{X}}$, and such that $\tilde{h}/2 \in \mathcal{L}(\xi, \eta)$. Together with (2.2.7), this implies $|\mathbb{E}[\tilde{h}(\xi)] - \mathbb{E}[\tilde{h}(\zeta)]| \leq 2 d_{KR}(\xi, \zeta)$ and concludes the proof. □

Finally, it is worth noticing that, when considering locally finite point processes on a lcscH space, all the above defined metrics imply convergence in distribution. For $d_\pi$ it directly follows by Proposition 2.1.5 (iii), for

$d_{TV}$ it descends from the structure of the vague topology on $\widetilde{\mathbf{N}}_{\mathbb{X}}$ and Proposition 2.1.5 (ii), whence the property holds for $d_{KR}$ as it dominates the total variation distance.

We now use the main result from Section 3.1 to derive a limit theorem for the Poisson process approximation. For a point processes $\xi$ and a Poisson process $\eta$ on a measurable space $\mathbb{X}$ with finite intensity measure, Theorem 3.1.1 provides bounds on the Wasserstein distance

$$d_W((\xi(A_1), \dots, \xi(A_d)), (\eta(A_1), \dots, \eta(A_d))\,,$$

where $A_1, \dots, A_d$ are measurable subsets of $\mathbb{X}$. This allows for a way to compare the distributions of $\xi$ and $\eta$ in terms of the distance $d_\pi$.

**Theorem 3.3.5.** *Let $\xi$ be a point process on $\mathbb{X}$ with finite intensity measure, and let $\eta$ be a Poisson process on $\mathbb{X}$ with finite intensity measure $\mu$. For any $i$-tuple $(A_1, \dots, A_i) \in \mathcal{X}_{\shortparallel}^i$ with $i \in \mathbb{N}$, consider a random vector $\mathbf{Z}^{A_{1:i}} = (Z_1^{A_{1:i}}, \dots, Z_i^{A_{1:i}})$ defined on the same probability space as $\xi$ with values in $\mathbb{Z}^i$, and define*

$$q_{m_{1:i}}^{A_{1:i}} = m_i \mathbb{P}\big((\xi(A_1), \dots, \xi(A_i)) = m_{1:i}\big) \\ - \mu(A_i)\mathbb{P}\big((\xi(A_1), \dots, \xi(A_i)) + \mathbf{Z}^{A_{1:i}} = (m_{1:i-1}, m_i - 1)\big) \tag{3.3.4}$$

*for $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$. Then,*

$$d_\pi(\xi, \eta) \leq \sup_{(A_1,\dots,A_d)\in\mathcal{X}_{\shortparallel}^d, d\in\mathbb{N}} \sum_{i=1}^d \left( \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i \\ m_i\neq 0}} \big|q_{m_{1:i}}^{A_{1:i}}\big| + 2\mu(A_i) \sum_{j=1}^i \mathbb{E}\big|Z_j^{A_{1:i}}\big| \right). \tag{3.3.5}$$

Note that a slightly sharper bound than (3.3.5) can be derived, as expressed in below Remark 3.3.6. We now prove Theorem 3.3.5 by mimicking the approach used in [5] to prove Theorem 2, as we derive the process bound as a consequence of the $d$-dimensional bound.

*Proof of Theorem 3.3.5.* Let $d \in \mathbb{N}$ and $\mathbf{A} = (A_1, \dots, A_d) \in \mathcal{X}_{\shortparallel}^d$. Define

$$\mathbf{X}^{\mathbf{A}} = (\xi(A_1), \dots, \xi(A_d)) \quad \text{and} \quad \mathbf{P}^{\mathbf{A}} = (\eta(A_1), \dots, \eta(A_d)),$$

where $\mathbf{P}^{\mathbf{A}}$ is a Poisson random vector with mean $\mathbb{E}[\mathbf{P}^{\mathbf{A}}] = (\mu(A_1), \dots, \mu(A_d))$. By Theorem 3.1.1 with $\mathbf{Z}^{(i)} = \mathbf{Z}^{A_{1:i}}$, we obtain

$$d_W(\mathbf{X}^{\mathbf{A}}, \mathbf{P}^{\mathbf{A}}) \leq \sum_{i=1}^d \left( \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i \\ m_i\neq 0}} \big|q_{m_{1:i}}^{A_{1:i}}\big| + 2\mu(A_i) \sum_{j=1}^i \mathbb{E}|Z_j^{A_{1:i}}| \right).$$

Taking the supremum over all $d$-tuples of disjoint measurable sets concludes the proof. ∎

*Remark* 3.3.6. By taking into account Remark 3.1.3, one immediately obtains

$$d_\pi(\xi,\eta) \le \sup_{(A_1,\dots,A_d)\in\mathcal{X}^d_{\shortparallel},d\in\mathbb{N}} \sum_{i=1}^{d} \Bigg( \min\left\{\mu(A_i), \frac{8\sqrt{\mu(A_i)}}{3\sqrt{2e}}\right\} \mathbb{E}\big|Z_i^{A_{1:i}}\big| \\ + 2\mu(A_i)\mathbb{P}\big(Z_{1:i-1}^{A_{1:i}} \ne 0\big) + \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i \\ m_i\ne 0}} \big|q_{m_{1:i}}^{A_{1:i}}\big| \Bigg).$$

In Sections 3.4 and 3.5, we apply Theorem 3.3.5 to obtain explicit Poisson process approximation results for point processes with Papangelou intensity and point processes of Poisson $U$-statistic structure. The latter are point processes that, once evaluated on a measurable set, become Poisson $U$-statistics. Analogous results were already proven for the Kantorovich-Rubinstein distance in [31, Theorem 3.7] and [30, Theorem 3.1], under the additional condition that the configuration space $\mathbb{X}$ is lcscH. It is interesting to note that the proof of our result for point processes with Papangelou intensity employs Theorem 3.3.5 with $\mathbf{Z}^{A_{1:i}}$ set to zero for all $i$, while for point processes of $U$-statistic structure, we find $\mathbf{Z}^{A_{1:i}}$ such that Equation (3.3.4) in Theorem 3.3.5 is satisfied with $q_{m_{1:i}}^{A_{1:i}} \equiv 0$ for all collections of disjoint sets.

## 3.4 Point processes with Papangelou intensity

Let $\xi$ be a proper point process on a measurable space $(\mathbb{X},\mathcal{X})$, that is, a point process that can be written as $\xi = \delta_{X_1} + \cdots + \delta_{X_\tau}$, for some random elements $X_1, X_2, \dots$ in $\mathbb{X}$ and a random variable $\tau \in \mathbb{N}_0 \cup \{\infty\}$. Note that any Poisson process can be seen as a proper point process, and that all locally finite point processes are proper if $(\mathbb{X},\mathcal{X})$ is a Borel space; see e.g. [61, Corollaries 3.7 and 6.5]. The so-called reduced Campbell measure $\mathcal{C}$ of $\xi$ is defined on the product space $(\mathbb{X}\times\mathbf{N}_{\mathbb{X}}, \mathcal{X}\otimes\mathcal{N}_{\mathbb{X}})$ by

$$\mathcal{C}(A) = \mathbb{E}\int_{\mathbb{X}} \mathbf{1}_A(x,\xi\setminus x)\,\xi(dx), \quad A\in\mathcal{X}\otimes\mathcal{N}_{\mathbb{X}},$$

where $\xi\setminus x$ denotes the point process $\xi-\delta_x$ if $x\in\xi$, and $\xi$ otherwise. Let $\nu$ be a $\sigma$-finite measure on $(\mathbb{X},\mathcal{X})$ and let $\mathbb{P}_\xi$ be the distribution of $\xi$ on $(\mathbf{N}_{\mathbb{X}},\mathcal{N}_{\mathbb{X}})$. If $\mathcal{C}$ is absolutely continuous with respect to $\nu\otimes\mathbb{P}_\xi$, any density $c$ of $\mathcal{C}$ with respect to $\nu\otimes\mathbb{P}_\xi$ is called (a version of) the Papangelou intensity of $\xi$. This notion was originally introduced by Papangelou in [78]. In other words, $c$ is a Papangelou intensity of $\xi$ relative to the measure $\nu$ if the Georgii–Nguyen–Zessin equation

$$\mathbb{E}\int_{\mathbb{X}} u(x,\xi\setminus x)\,\xi(dx) = \int_{\mathbb{X}} \mathbb{E}[c(x,\xi)u(x,\xi)]\nu(dx), \tag{3.4.1}$$

is satisfied for all measurable functions $u:\mathbb{X}\times\mathbf{N}_{\mathbb{X}}\to[0,\infty)$. Intuitively $c(x,\xi)$ is a random variable that measures the interaction between $x$ and $\xi$; as a reinforcement of this exposition, it is well-known that if $c$ is deterministic, that is, $c(x, \xi) = f(x)$ for some positive and measurable function $f$, then $\xi$ is a Poisson process with intensity measure $\mu(A) = \int_A f(x)\nu(dx)$, $A \in \mathcal{X}$; see e.g. [61, Theorem 4.1]. For more details on this interpretation we refer the reader to [31, Section 4]; see also [59] and [96] for connections between the Papangelou intensity and Gibbs point processes.

In the next theorem we prove a bound for the $d_\pi$ distance between a point process $\xi$ that admits Papangelou intensity relative to a measure $\nu$, and a Poisson process $\eta$ with intensity measure $\mu$ absolutely continuous with respect to $\nu$. For a locally compact metric space, Theorem 3.4.1 yields the same bound as [31, Theorem 3.7], but for the metric $d_\pi$ instead of the Kantorovich-Rubinstein distance.

**Theorem 3.4.1.** *Let $\xi$ be a proper point process on $\mathbb{X}$ that admits Papangelou intensity $c$ with respect to a $\sigma$-finite measure $\nu$ such that $\int_{\mathbb{X}} \mathbb{E}|c(x, \xi)|\nu(dx) < \infty$. Let $\eta$ be a Poisson process on $\mathbb{X}$ with finite intensity measure $\mu$ having density $f$ with respect to $\nu$. Then*

$$d_\pi(\xi, \eta) \leq \int_{\mathbb{X}} \mathbb{E}\,|c(x, \xi) - f(x)|\, \nu(dx).$$

*Proof of Theorem 3.4.1.* The condition $\int_{\mathbb{X}} \mathbb{E}|c(x, \xi)|\nu(dx) < \infty$ and Equation (3.4.1) ensure that $\xi$ has finite intensity measure. Consider $i \in \mathbb{N}$ and $(A_1, \ldots, A_i) \in \mathcal{X}_{\shortparallel}^i$. Hereafter, $\xi(A_{1:i})$ is shorthand notation for $(\xi(A_1), \ldots, \xi(A_i))$. The idea of the proof is to apply Theorem 3.3.5 with the random vectors $\mathbf{Z}^{A_{1:i}}$ assumed to be $\mathbf{0}$. In this case,

$$\begin{aligned} q_{m_{1:i}}^{A_{1:i}} &= m_i \mathbb{P}\big(\xi(A_{1:i}) = m_{1:i}\big) - \mu(A_i)\mathbb{P}\big(\xi(A_{1:i}) = (m_{1:i-1}, m_i - 1)\big) \\ &= m_i \mathbb{P}\big(\xi(A_{1:i}) = m_{1:i}\big) - \int_{\mathbb{X}} \mathbb{E}\big[f(x)\mathbf{1}_{A_i}(x)\mathbf{1}\{\xi(A_{1:i}) = (m_{1:i-1}, m_i - 1)\}\big]\nu(dx) \end{aligned}$$

for $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$, $i = 1, \ldots, d$. It follows from (3.4.1) that

$$\begin{aligned} m_i \mathbb{P}\big(\xi(A_{1:i}) = m_{1:i}\big) &= \mathbb{E} \int_{\mathbb{X}} \mathbf{1}_{A_i}(x)\mathbf{1}\{\xi \setminus x(A_{1:i}) = (m_{1:i-1}, m_i - 1)\}\, \xi(dx) \\ &= \int_{\mathbb{X}} \mathbb{E}\big[c(x, \xi)\mathbf{1}_{A_i}(x)\mathbf{1}\{\xi(A_{1:i}) = (m_{1:i-1}, m_i - 1)\}\big]\nu(dx), \end{aligned}$$

hence

$$q_{m_{1:i}}^{A_{1:i}} = \int_{\mathbb{X}} \mathbb{E}\big[(c(x, \xi) - f(x))\mathbf{1}_{A_i}(x)\mathbf{1}\{\xi(A_{1:i}) = (m_{1:i-1}, m_i - 1)\}\big]\nu(dx).$$

Theorem 3.3.5 yields

$$d_\pi(\xi, \eta) \leq \sup_{(A_1, \ldots, A_d) \in \mathcal{X}_{\shortparallel}^d, d \in \mathbb{N}} \sum_{i=1}^{d} \sum_{\substack{m_{1:i} \in \mathbb{N}_0^i \\ m_i \neq 0}} \big|q_{m_{1:i}}^{A_{1:i}}\big|\,.$$

The inequalities

$$\begin{aligned}\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}} \left|q_{m_{1:i}}^{A_{1:i}}\right| &\leq \sum_{\substack{m_{1:i}\in\mathbb{N}_0^i,\\ m_i\neq 0}} \int_{\mathbb{X}} \mathbb{E}\big[|c(x,\xi)-f(x)|\mathbf{1}_{A_i}(x)\mathbf{1}\{\xi(A_{1:i})=(m_{1:i-1},m_i-1)\}\big]\nu(dx)\\ &\leq \int_{\mathbb{X}} \mathbb{E}\Big[|c(x,\xi)-f(x)|\mathbf{1}_{A_i}(x)\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}}\mathbf{1}\{\xi(A_{1:i})=(m_{1:i-1},m_i-1)\}\Big]\nu(dx)\\ &\leq \int_{\mathbb{X}} \mathbb{E}\big[|c(x,\xi)-f(x)|\mathbf{1}_{A_i}(x)\big]\nu(dx)\end{aligned}$$

imply that

$$\sum_{i=1}^{d}\sum_{\substack{m_{1:i}\in\mathbb{N}_0^i\\ m_i\neq 0}} \left|q_{m_{1:i}}^{A_{1:i}}\right| \leq \int_{\mathbb{X}} \mathbb{E}\,|c(x,\xi)-f(x)|\,\nu(dx)$$

for any $A_{1:d}\in\mathcal{X}_{\shortparallel}^{d}$ with $d\in\mathbb{N}$. Thus, we obtain the assertion. □

## 3.5 Point processes of Poisson $U$-statistic structure

Let $(\mathbb{X},\mathcal{X})$ and $(\mathbb{Y},\mathcal{Y})$ be measurable spaces. For $k\in\mathbb{N}$ and a symmetric domain $D\in\mathcal{X}^k$, let $g: D\to\mathbb{Y}$ be a symmetric measurable function, i.e., for any $(x_1,\dots,x_k)\in D$ and any index permutation $\sigma$, $(x_{\sigma(1)},\dots,x_{\sigma(k)})\in D$ and $g(x_1,\dots,x_k)=g(x_{\sigma(1)},\dots,x_{\sigma(k)})$. Let $\eta$ be a Poisson process on $\mathbb{X}$ with finite intensity measure $\nu$. We are interested in the point process on $\mathbb{Y}$ given by

$$\xi=\frac{1}{k!}\sum_{(x_1,\dots,x_k)\in\eta^k_{\neq}\cap D}\delta_{g(x_1,\dots,x_k)}, \tag{3.5.1}$$

where $\eta^k_{\neq}$ denotes the collection of all $k$-tuples $(x_1,\dots,x_k)$ of points from $\eta$ with pairwise distinct indexes. The point process $\xi$ has a Poisson $U$-statistic structure in the sense that, for any $B\in\mathcal{Y}$, $\xi(B)$ is a Poisson $U$-statistic of order $k$. We refer to the monographs [55, 62] for more details on $U$-statistics and their applications to statistics. Hereafter we discuss the Poisson process approximation in the metric $d_\pi$ for the point process $\xi$. We prove the exact analogue of [30, Theorem 3.1], with the Kantorovich–Rubinstein distance replaced by $d_\pi$. Several applications of this result are presented in [30], alongside with the case of underlying binomial point processes. It is worth mentioning that [30] relies on a slightly less general setup: $\mathbb{X}$ is assumed to be a locally compact second countable Hausdorff space, while in the present work any measurable space is allowed.

Let $\mu$ denote the intensity measure of $\xi$, and note that, since $\nu$ is a finite measure on $\mathbb{X}$, the multivariate Mecke equation (Proposition 2.1.12) implies $\mu(\mathbb{Y})<\infty$. Define

$$R=\max_{1\leq i\leq k-1}\int_{\mathbb{X}^i}\left(\int_{\mathbb{X}^{k-i}}\mathbf{1}\{(x_1,\dots,x_k)\in D\}\,\nu^{k-i}(d(x_{i+1},\dots,x_k))\right)^2\nu^i(d(x_1,\dots,x_i))$$

for $k \geq 2$, and put $R = 0$ for $k = 1$.

**Theorem 3.5.1.** *Let $\xi$, $\mu$ and $R$ be as above, and let $\gamma$ be a Poisson process on $\mathbb{Y}$ with intensity measure $\mu$. Then,*

$$d_\pi(\xi, \gamma) \leq \frac{2^{k+1}}{k!} R.$$

If the intensity measure $\mu$ of $\xi$ is the zero measure, then the proof of Theorem 3.5.1 is trivial. From now on, we assume $0 < \mu(\mathbb{Y}) < \infty$. The multivariate Mecke equation yields for every $A \in \mathcal{Y}$ that

$$\mu(A) = \mathbb{E}[\xi(A)] = \frac{1}{k!}\mathbb{E} \sum_{\boldsymbol{x} \in \eta^k_{\neq} \cap D} \mathbf{1}\{g(\boldsymbol{x}) \in A\} = \frac{1}{k!} \int_D \mathbf{1}\{g(\boldsymbol{x}) \in A\}\, \nu^k(d\boldsymbol{x}).$$

Define the random element $\mathbf{X}^A = (X_1^A, \dots, X_k^A)$ in $\mathbb{X}^k$ independent of $\eta$ and distributed according to

$$\mathbb{P}\left(\mathbf{X}^A \in B\right) = \frac{1}{k!\mu(A)} \int_D \mathbf{1}\{g(\boldsymbol{x}) \in A\}\mathbf{1}\{\boldsymbol{x} \in B\}\, \nu^k(d\boldsymbol{x})$$

for all $B$ in the product $\sigma$-field of $\mathbb{X}^k$ when $\mu(A) > 0$, and set $\mathbf{X}^A = \boldsymbol{x}_0$ for some $\boldsymbol{x}_0 \in \mathbb{X}^k$ when $\mu(A) = 0$. For any vector $\boldsymbol{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$, denote by $\Delta(\boldsymbol{x})$ the sum of $k$ Dirac measures located at the vector components, that is

$$\Delta(\boldsymbol{x}) = \Delta(x_1, \dots, x_k) = \sum_{i=1}^k \delta_{x_i}\,.$$

In what follows, for any point process $\zeta$ on $\mathbb{X}$, $\xi(\zeta)$ is the point process defined as in (3.5.1) with $\eta$ replaced by $\zeta$. Further, like in Section 3.4, $\xi(A_{1:i})$ denotes the random vector $(\xi(A_1), \dots, \xi(A_i))$, for any $A_1, \dots, A_i \in \mathcal{Y}$, $i \in \mathbb{N}$.

*Proof of Theorem 3.5.1.* For $k = 1$, Theorem 3.5.1 is a direct consequence of [61, Theorem 5.1]. Whence, we assume $k \geq 2$. Let $A_1, \dots, A_i \in \mathcal{Y}$ with $i \in \mathbb{N}$ be disjoint sets and let $m_{1:i} \in \mathbb{N}_0^i$ with $m_i \neq 0$. Suppose $\mu(A_i) > 0$. Again Proposition 2.1.12 implies that

$$\begin{aligned}
m_i\mathbb{P}(\xi(A_{1:i}) = m_{1:i}) &= \frac{1}{k!}\mathbb{E} \sum_{\boldsymbol{x} \in \eta^k_{\neq} \cap D} \mathbf{1}\{g(\boldsymbol{x}) \in A_i\}\mathbf{1}\{\xi(A_{1:i}) = m_{1:i}\} \\
&= \frac{1}{k!} \int_D \mathbf{1}\{g(\boldsymbol{x}) \in A_i\}\mathbb{P}(\xi(\eta + \Delta(\boldsymbol{x}))(A_{1:i}) = m_{1:i})\, \nu^k(d\boldsymbol{x}) \\
&= \frac{1}{k!} \int_D \mathbf{1}\{g(\boldsymbol{x}) \in A_i\}\mathbb{P}\left(\xi(\eta + \Delta(\boldsymbol{x}))(A_{1:i}) - \delta_{g(\boldsymbol{x})}(A_{1:i}) = (m_{1:i-1}, m_i - 1)\right) \nu^k(d\boldsymbol{x}) \\
&= \mu(A_i)\mathbb{P}\left(\xi\left(\eta + \Delta\left(\mathbf{X}^{A_i}\right)\right)(A_{1:i}) - \delta_{g\left(\mathbf{X}^{A_i}\right)}(A_{1:i}) = (m_{1:i-1}, m_i - 1)\right),
\end{aligned} \tag{3.5.2}$$

where the second last equality holds true because $\delta_{g(\boldsymbol{x})}(A_{1:i})$ is the vector $(0, \dots, 0, 1) \in \mathbb{N}_0^i$ when $g(\boldsymbol{x}) \in A_i$. The previous identity is verified also if $\mu(A_i) = 0$. Hence, for

$$\mathbf{Z}^{A_{1:i}} = \xi\left(\eta + \Delta\left(\mathbf{X}^{A_i}\right)\right)(A_{1:i}) - \xi(A_{1:i}) - \delta_{g\left(\mathbf{X}^{A_i}\right)}(A_{1:i})\,,$$

the quantity $q_{m_{1:i}}^{A_{1:i}}$ defined by Equation (3.3.4) in Theorem 3.3.5 is zero. Note that $\mathbf{Z}^{A_{1:i}}$ has non-negative components. Hence, for any $d \in \mathbb{N}$ and $(A_1, \dots, A_d) \in \mathcal{X}_{\shortparallel}^d$,

$$\begin{aligned}
\sum_{i=1}^{d} \mu(A_i) \sum_{j=1}^{i} \mathbb{E}\left|\mathbf{Z}_j^{A_{1:i}}\right| &= \sum_{i=1}^{d} \mu(A_i) \sum_{j=1}^{i} \mathbb{E}\left[\xi\left(\eta + \Delta\left(\mathbf{X}^{A_i}\right)\right)(A_j) - \xi(A_j) - \delta_{g\left(\mathbf{X}^{A_i}\right)}(A_j)\right] \\
&\leq \sum_{i=1}^{d} \mu(A_i) \mathbb{E}\left[\xi\left(\eta + \Delta\left(\mathbf{X}^{A_i}\right)\right)(\mathbb{Y}) - \xi(\mathbb{Y}) - 1\right] \\
&= \frac{1}{k!} \sum_{i=1}^{d} \int_D \mathbf{1}\{g(\boldsymbol{x}) \in A_i\} \mathbb{E}\left[\xi(\eta + \Delta(\boldsymbol{x}))(\mathbb{Y}) - \xi(\mathbb{Y}) - 1\right] \nu^k(d\boldsymbol{x}) \\
&\leq \mu(\mathbb{Y}) \mathbb{E}\left[\xi\left(\eta + \Delta\left(\mathbf{X}^{\mathbb{Y}}\right)\right)(\mathbb{Y}) - \xi(\mathbb{Y}) - 1\right].
\end{aligned}$$

Thus, Theorem 3.3.5 gives

$$d_\pi(\xi, \gamma) \leq 2\mu(\mathbb{Y}) \mathbb{E}\left[\xi\left(\eta + \Delta\left(\mathbf{X}^{\mathbb{Y}}\right)\right)(\mathbb{Y}) - \xi(\mathbb{Y}) - 1\right]. \tag{3.5.3}$$

From (3.5.2) with $i = 1$ and $A_1 = \mathbb{Y}$, it follows that the random variable $\xi\left(\eta + \Delta\left(\mathbf{X}^{\mathbb{Y}}\right)\right)(\mathbb{Y})$ has the size bias distribution of $\xi(\mathbb{Y})$. Property (2.3.7) with $f$ being the identity function and simple algebraic computations yield

$$\begin{aligned}
\mathbb{E}\left[\xi\left(\eta + \Delta\left(\mathbf{X}^{\mathbb{Y}}\right)\right)(\mathbb{Y}) - \xi(\mathbb{Y}) - 1\right] &= \mu(\mathbb{Y})^{-1}\left\{\mathbb{E}\left[\xi(\mathbb{Y})^2\right] - \mu(\mathbb{Y})^2 - \mu(\mathbb{Y})\right\} \\
&= \mu(\mathbb{Y})^{-1}\left\{\operatorname{Var}(\xi(\mathbb{Y})) - \mu(\mathbb{Y})\right\}.
\end{aligned} \tag{3.5.4}$$

Moreover, [87, Lemma 3.5] gives

$$\operatorname{Var}(\xi(\mathbb{Y})) - \mu(\mathbb{Y}) \leq \sum_{i=1}^{k-1} \frac{1}{k!}\binom{k}{i} R \leq \frac{2^k - 1}{k!} R.$$

These inequalities combined with (3.5.3) and (3.5.4) deliver the assertion. $\square$

# Chapter 4

# Gaussian Approximation in a Birth-Growth Model with Random Growth Speed

This chapter is based on the following article:
*C. Bhattacharjee, I. Molchanov and R. Turin. Central limit theorem for a birth-growth model with Poisson arrivals and random growth speed. arXiv:2107.06792, 2021.*

## 4.1 Model and main results

We consider a generalization of the Johnson-Mehl model by introducing random growth speed for the seeds. As already described in Chapter 1, in the spatial Johnson-Mehl growth model, seeds appear at random times and locations, according to a Poisson process. Once a seed is born at time $t$, it begins to form a cell by growing radially in all directions at a constant speed $v \geq 0$, so that by time $t'$ it occupies the ball of radius $v(t' - t)$. Instead of taking $v$ constant, we assume it to be random, sampled from a probability distribution $\nu$ on $\mathbb{R}_+$. Therefore, we consider a birth-growth model in which seeds come with their own growth speed attached, hence forming a marked point process with location, time and speed components.

More precisely, we work in the space $\mathbb{X} := \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+$, $d \in \mathbb{N}$, with the Borel $\sigma$-algebra, and let $\eta$ be a

Poisson process on $\mathbb{X} := \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure $\mu := \lambda \otimes \theta \otimes \nu$, where $\lambda$ is the Lebesgue measure on $\mathbb{R}^d$, $\theta$ is a non-null locally finite measure on $\mathbb{R}_+$, and $\nu$ is a probability distribution on $\mathbb{R}_+$ with $\nu(\{0\}) < 1$. The points from $\mathbb{X}$ are written as $\boldsymbol{x} := (x, t_x, v_x)$, so that $\boldsymbol{x}$ designates a seed born in position $x$ at time $t_x$, which then grows radially in all directions with speed $v_x$. For $\boldsymbol{x} \in \mathbb{X}$, the region

$$G_{\boldsymbol{x}} = G_{x,t_x,v_x} := \big\{(y, t_y) \in \mathbb{R}^d \times \mathbb{R}_+ : t_y \geq t_x, \|y - x\| \leq v_x(t_y - t_x)\big\}$$

is the growth region of the seed $\boldsymbol{x}$. Recall that $\mathbf{N}_{\mathbb{X}}$ is the collection of all $\sigma$-finite counting measures $\mathsf{M}$ on $(\mathbb{X}, \mathcal{X})$, equipped with the smallest $\sigma$-algebra $\mathcal{N}_{\mathbb{X}}$ such that the maps $\mathsf{M} \mapsto \mathsf{M}(A)$ are measurable for all Borel $A$. We write $\boldsymbol{x} \in \mathsf{M}$ if $\mathsf{M}(\{\boldsymbol{x}\}) \geq 1$. For a given point configuration $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$, a point $\boldsymbol{x} \in \mathsf{M}$ is said to be *exposed* in $\mathsf{M}$ if it does not belong to the growth region of any other point $\boldsymbol{y} \in \mathsf{M}$, $\boldsymbol{y} \neq \boldsymbol{x}$. Given a measurable weight function $h : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ the main object of interest for us is the sum of $h$ over the exposed points $(x, t_x)$ in $\eta$. Our aim is to provide sufficient conditions for Gaussian convergence of such sums. It should be noted that the property of being exposed is not influenced by the speed component of $\boldsymbol{x}$ and also that, because of random speeds, it may happen that the cell grown from a non-exposed seed shades a subsequent seed which would be exposed otherwise. The event that a point $(x, t_x, v_x) \in \eta$ is exposed depends only on the point configuration in the region

$$L_{x,t_x} := \big\{(y, t_y, v_y) \in \mathbb{X} : \|x - y\| \leq v_y(t_x - t_y)\big\}. \tag{4.1.1}$$

The *influence set* $L_{\boldsymbol{x}} = L_{x,t_x}$ just defined, is exactly the set of points that were born before time $t_x$ and which at time $t_x$ occupy a region that covers the location $x$, thereby shading it. Note that $\boldsymbol{y} \in L_{\boldsymbol{x}}$ if and only if $\boldsymbol{x} \in G_{\boldsymbol{y}}$. Clearly, a point $\boldsymbol{x} \in \mathsf{M}$ is exposed in $\mathsf{M}$ if and only if $\mathsf{M}(L_{\boldsymbol{x}} \setminus \{x\}) = 0$. We write $(y, t_y, v_y) \preceq (x, t_x)$ or $\boldsymbol{y} \preceq \boldsymbol{x}$ if $\boldsymbol{y} \in L_{x,t_x}$ (recall that the speed component of $\boldsymbol{x}$ is irrelevant in such a relation) and so $\boldsymbol{x}$ is not an exposed point with respect to $\delta_{\boldsymbol{y}}$, where $\delta_{\boldsymbol{y}}$ denotes the Dirac measure at $\boldsymbol{y}$.

For $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ and $\boldsymbol{x} \in \mathsf{M}$, denote

$$H_{\boldsymbol{x}}(\mathsf{M}) \equiv H_{x,t_x}(\mathsf{M}) := \mathbf{1}\{\boldsymbol{x} \text{ is exposed in } \mathsf{M}\} = \mathbf{1}_{\mathsf{M}(L_{x,t_x} \setminus \{\boldsymbol{x}\})=0}.$$

We consider weight functions $h(\boldsymbol{x})$ which are products of two measurable functions $h_1 : \mathbb{R}^d \to \mathbb{R}_+$ and $h_2 : \mathbb{R}_+ \to \mathbb{R}_+$. Let

$$F(\mathsf{M}) := \int_{\mathbb{X}} h_1(x) h_2(t_x) H_{\boldsymbol{x}}(\mathsf{M})\, \mathsf{M}(\mathrm{d}\boldsymbol{x}) = \sum_{\boldsymbol{x} \in \mathsf{M}} h_1(x) h_2(t_x) H_{\boldsymbol{x}}(\mathsf{M}) \tag{4.1.2}$$

be the sum of the weights determined by $h_1(x)h_2(t_x)$ at locations and birth times of the exposed points $\boldsymbol{x} = (x, t_x, v_x)$ in $\mathsf{M}$. For example, if $h_1$ is the indicator function of a window $W \subseteq \mathbb{R}^d$ and $h_2(t) = \mathbf{1}\{t < a\}$ for some $a \in (0, \infty]$, then $F$ is the total number of exposed points located in $W$ that were born before time $a$.

The functional $F(\eta)$ is a region-stabilizing functional, in the sense of Section 2.4, and can be represented as $F(\eta) = \sum_{\boldsymbol{x}\in\eta} S(\boldsymbol{x}, \eta)$, where the score function $S$ is given by

$$S(\boldsymbol{x}, \mathsf{M}) := h_1(x)h_2(t_x)H_{\boldsymbol{x}}(\mathsf{M}), \quad \boldsymbol{x} \in \mathsf{M}. \tag{4.1.3}$$

Theorem 2.4.2 yields ready-to-use bounds on the Wasserstein and Kolmogorov distances between $F$, suitably normalized, and a standard Gaussian random variable $N$ upon validating Equation (2.4.1) and conditions (A1) and (A2) from Section 2.4.

Now we are ready to state our main results. First, we list the necessary assumptions on our model. In the sequel, we drop $\lambda$ in Lebesgue integrals and simply write $\mathrm{d}x$ instead of $\lambda(\mathrm{d}x)$.

(A) The location-weight function $h_1$ satisfies

$$\int_{\mathbb{R}^d} \max\{h_1(x), h_1(x)^8\}\, \mathrm{d}x < \infty.$$

(B) For all $a > 0$,

$$\int_0^\infty \max\{1, h_2(t)^4\}\, e^{-a\Lambda(t)}\, \theta(\mathrm{d}t) < \infty,$$

where

$$\Lambda(t) := \omega_d \int_0^t (t-s)^d \theta(\mathrm{d}s) \tag{4.1.4}$$

and $\omega_d$ is the volume of the $d$-dimensional unit Euclidean ball.

(C) The moment of $\nu$ of order $6d$ is finite, i.e., $\nu_{6d} < \infty$, where for $k \in \mathbb{N}$, we denote by $\nu_k$ the $k$-th moment of $\nu$,

$$\nu_k := \int_0^\infty v^k \nu(\mathrm{d}v).$$

Note that the function $\Lambda(t)$ given at (4.1.4) is, up to a constant, the measure of the influence set of any point $\boldsymbol{x} \in \mathbb{X}$ with time component $t_x = t$ (the measure of the influence set does not depend on the location and speed components of $\boldsymbol{x}$). Indeed, the $\mu$-content of $L_{x,t_x}$ is given by

$$\begin{aligned}\mu(L_{x,t_x}) &= \int_0^\infty \int_0^{t_x} \int_{\mathbb{R}^d} \mathbf{1}_{y \in B_{v_y(t_x - t_y)}(x)}\, \mathrm{d}y\, \theta(\mathrm{d}t_y)\nu(\mathrm{d}v_y) \\ &= \int_0^\infty \nu(\mathrm{d}v_y) \int_0^{t_x} \omega_d v_y^d (t_x - t_y)^d \theta(\mathrm{d}t_y) = \nu_d \Lambda(t_x),\end{aligned}$$

where $B_r(x)$ denotes the closed $d$-dimensional Euclidean ball of radius $r$ centered at $x \in \mathbb{R}^d$. In particular, if $\theta$ is the Lebesgue measure on $\mathbb{R}_+$, then $\Lambda(t) = \omega_d\, t^{d+1}/(d+1)$.

Define

$$h_1^{(i)} := \int_{\mathbb{R}^d} h_1(x)^i \,\mathrm{d}x, \quad i \in \mathbb{N}, \tag{4.1.5}$$

and

$$\widetilde{M}_i := \int_{\mathbb{R}^d} \widetilde{M}(x)^i \,\mathrm{d}x, \quad i = 1, 2, \qquad \text{where} \quad \widetilde{M}(x) := \max\{h_1(x)^2, h_1(x)^4\}.$$

The following theorem is our first main result.

**Theorem 4.1.1.** *Let $\eta$ be a Poisson process on $\mathbb{X}$ with intensity measure $\mu$ as above, such that the assumptions (A)–(C) hold. Then for $F := F(\eta)$ as in* (4.1.2),

$$d_W\left(\frac{F - \mathbb{E}F}{\sqrt{\operatorname{Var} F}}, N\right) \leq C\left[\frac{\widetilde{M}_2^{1/2}}{\operatorname{Var} F} + \frac{\widetilde{M}_1}{(\operatorname{Var} F)^{3/2}}\right],$$

*and*

$$d_K\left(\frac{F - \mathbb{E}F}{\sqrt{\operatorname{Var} F}}, N\right) \leq C\left[\frac{\widetilde{M}_2^{1/2} + \widetilde{M}_1^{1/2}}{\operatorname{Var} F} + \frac{\widetilde{M}_1}{(\operatorname{Var} F)^{3/2}} + \frac{\widetilde{M}_1^{5/4} + \widetilde{M}_1^{3/2}}{(\operatorname{Var} F)^2}\right]$$

*for a constant $C > 0$ which depends on $h_2$, $d$, the first $6d$ moments of $\nu$, and $\theta$.*

To derive a quantitative central limit theorem from Theorem 4.1.1, a lower bound on the variance is needed. The following proposition provides general lower and upper bounds on the variance, which are then specialized for the measure $\theta$ on $\mathbb{R}_+$ given by the density

$$\theta(\mathrm{d}t) := t^\tau \mathrm{d}t, \quad \tau \in (-1, \infty). \tag{4.1.6}$$

In the following, for $t_1, t_2 \in \mathbb{R}$, we write $t_1 \wedge t_2$ for $\min\{t_1, t_2\}$. For $a \in (0, \infty]$ and $\tau > -1$, define the function

$$l_{a,\tau}(x) := \gamma\left(\frac{\tau + 1}{d + \tau + 1}, a^{d+\tau+1} x\right) x^{-(\tau+1)/(d+\tau+1)}, \quad x > 0, \tag{4.1.7}$$

where $\gamma(p, z) := \int_0^z t^{p-1} e^{-t} \mathrm{d}t$ is the lower incomplete Gamma function.

**Proposition 4.1.2.** *Let the assumptions (A)–(C) be in force. For a Poisson process $\eta$ with intensity measure $\mu$ as above and $F := F(\eta)$ as in* (4.1.2),

$$\frac{\operatorname{Var}(F)}{h_1^{(2)}} \geq \left[\int_0^\infty h_2(t)^2 w(t) \theta(\mathrm{d}t) - 2\omega_d \nu_d \int_0^\infty \int_0^t (t - s)^d h_2(s) h_2(t) w(s) w(t) \theta(\mathrm{d}s) \theta(\mathrm{d}t)\right] \tag{4.1.8}$$

*and*

$$\frac{\operatorname{Var}(F)}{h_1^{(2)}} \leq \left[2\int_0^\infty h_2(t)^2 w(t)^{1/2} \theta(\mathrm{d}t) \right. \tag{4.1.9}$$

$$\left. + \omega_d^2 \nu_{2d} \int_{\mathbb{R}_+^2} \int_0^{t_1 \wedge t_2} (t_1 - s)^d (t_2 - s)^d h_2(t_1) h_2(t_2) w(t_1)^{1/2} w(t_2)^{1/2} \theta(\mathrm{d}s) \theta^2(\mathrm{d}(t_1, t_2))\right],$$

*where*

$$w(t) := e^{-\nu_d \Lambda(t)} = \mathbb{E}\left[H_{0,t}(\eta)\right] \tag{4.1.10}$$

*and* $h_1^{(2)}$ *is defined at* (4.1.5). *If* $\theta$ *is given by* (4.1.6) *and* $h_2(t) = \mathbf{1}\{t < a\}$ *for some* $a \in (0, \infty]$, *then*

$$C_1(d-1-\tau) < C_1' \leq \frac{\operatorname{Var}(F)}{h_1^{(2)} l_{a,\tau}(\nu_d)} \leq C_2(1 + \nu_{2d}\nu_d^{-2}) \tag{4.1.11}$$

*for constants* $C_1, C_1', C_2$ *depending on the dimension* $d$ *and* $\tau$, *and* $C_1, C_2 > 0$.

We remark here that the lower bound in (4.1.11) is useful only when $\tau \leq d-1$. But we believe that a positive lower bound still exists when $\tau > d-1$, even though our arguments in general do not apply for such $\tau$.

In the case of a deterministic speed $v$ and for $h_1 = \mathbf{1}_W$ being the indicator function of an observation window $W \subseteq \mathbb{R}^d$, Proposition 4.1.2 provides an explicit condition on $\theta$ and $h_2$ ensuring that the variance scales like the volume of the observation window in the classical Johnson–Mehl growth model. The problem of finding such a condition, explicitly formulated in [28, page 754], arose in [27], where asymptotic normality for the number of exposed seeds in a region, as the volume of the region approaches infinity, is obtained under the assumption that the variance scales properly. This was by then only shown numerically for the case when $\theta$ is the Lebesgue measure and $d = 1, 2, 3, 4$. Subsequent papers [81, 94] derived the variance scaling for $\theta$ being the Lebesgue measure and some generalizations of it, but in a slightly different formulation of the model, in which seeds that do not appear in the observation window are automatically rejected and cannot influence the growth pattern in the region $W$.

The bounds in Theorem 4.1.1 can be specified under two different scenarios. When considering a sequence of weight functions, under suitable conditions Theorem 4.1.1 provides a quantitative CLT for the corresponding functionals $(F_n)_{n\in\mathbb{N}}$. Keeping all other quantities fixed with respect to $n$, let $(h_{1,n})_{n\in\mathbb{N}}$ be a sequence of non-negative location-weight functions on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} \max\{h_{1,n}(x), h_{1,n}(x)^8\} \mathrm{d}x \leq C h_{1,n}^{(2)}, \tag{4.1.12}$$

for some constant $C$, where $h_{1,n}^{(2)}$ is defined at (4.1.5). In view of Proposition 4.1.2, this provides the following quantitative CLT.

**Theorem 4.1.3.** *For* $n \in \mathbb{N}$ *and* $\eta$ *as in Theorem 4.1.1, let* $F_n := F_n(\eta)$, *where* $F_n$ *is defined as in* (4.1.2) *with* $h_2$ *independent of* $n$ *and* $h_1 = h_{1,n}$ *for some* $(h_{1,n})_{n\in\mathbb{N}}$, *satisfying* (4.1.12) *and such that* $h_{1,n}^{(2)} \geq 1$ *for all sufficiently large* $n$. *Let the assumptions (A)–(C) be in force. If* $\theta$, $\nu$ *and* $h_2$ *satisfy*

$$\int_0^\infty h_2(t)^2 w(t)\theta(\mathrm{d}t) - 2\omega_d \nu_d \int_0^\infty \int_0^t (t-s)^d h_2(s) h_2(t) w(s) w(t) \theta(\mathrm{d}s)\theta(\mathrm{d}t) > 0\,, \tag{4.1.13}$$

*where* $w(t)$ *is given at* (4.1.10)*, then there exists a constant* $C > 0$ *such that*

$$\max\left\{d_W\left(\frac{F_n - \mathbb{E}F_n}{\sqrt{\operatorname{Var} F_n}}, N\right), d_K\left(\frac{F_n - \mathbb{E}F_n}{\sqrt{\operatorname{Var} F_n}}, N\right)\right\} \le C\left(h_{1,n}^{(2)}\right)^{-1/2},$$

*for all* $n \in \mathbb{N}$*. In particular, condition* (4.1.13) *is satisfied for* $\theta$ *given at* (4.1.6) *with* $\tau \in (-1, d-1]$ *and* $h_2(t) = \mathbf{1}\{t < a\}$ *for any* $a \in (0, \infty]$.

Note that the rate of convergence expressed in Theorem 4.1.3 is presumably optimal, as it has the same order as the inverse of square root of the variance. Let $h_{1,n} = \mathbf{1}_{W_n}$, $n \in \mathbb{N}$, be indicator functions of a sequence of windows $(W_n)_{n\in\mathbb{N}}$ with $\lambda(W_n) \to \infty$ as $n \to \infty$, and $h_2(t) = \mathbf{1}\{t < a\}$ for some $a \in (0, \infty]$. In this case (4.1.12) is trivially satisfied and, when (4.1.13) is satisfied, Theorem 4.1.3 yields a CLT for the number of exposed seeds born before time $a \in (0, \infty]$, with rate of convergence of order $1/\sqrt{\lambda(W_n)}$. This extends the CLT for the number of exposed seeds from [27] in several directions: the model is generalized to random growth speed, there is no constraint of any kind on the shape of the growing windows, and a logarithmic factor is removed from the rate of convergence.

In a different scenario, when $\theta$ has a power-law density (4.1.6) with $\tau \in (-1, d-1]$, it is possible to explicitly specify the dependence of the bound in Theorem 4.1.1 on the moments of $\nu$, as stated in the following result.

**Theorem 4.1.4.** *Let the assumptions (A) and (C) be in force. For* $\theta$ *given at* (4.1.6) *with* $\tau \in (-1, d-1]$*, consider* $F^{(a)} := F^{(a)}(\eta)$ *where* $\eta$ *is as in Theorem 4.1.1 and* $F^{(a)}$ *is defined as in* (4.1.2) *with* $h_2(t) = \mathbf{1}\{t < a\}$ *for some* $a \in (0, \infty]$*. Then, there exists a constant* $C > 0$*, depending only on* $d$ *and* $\tau$*, such that*

$$d_W\left(\frac{F^{(a)} - \mathbb{E}F^{(a)}}{\sqrt{\operatorname{Var} F^{(a)}}}, N\right) \le C l_{a,\tau}(\nu_d)^{-1/2}\left[\frac{\left(1 + \nu_{6d}\nu_d^{-6}\right)^{3/2} \widetilde{M}_2^{1/2}}{h_1^{(2)}} + \frac{\left(1 + \nu_{6d}\nu_d^{-6}\right) \widetilde{M}_1}{(h_1^{(2)})^{3/2}}\right],$$

*and*

$$\begin{aligned} d_K\left(\frac{F^{(a)} - \mathbb{E}F^{(a)}}{\sqrt{\operatorname{Var} F^{(a)}}}, N\right) \le C l_{a,\tau}(\nu_d)^{-1/2}\Bigg[&\frac{\left(1 + \nu_{6d}\nu_d^{-6}\right)^{3/2} \widetilde{M}_2^{1/2} + \left(1 + \nu_{6d}\nu_d^{-6}\right)^{1/2} \widetilde{M}_1^{1/2}}{h_1^{(2)}} \\ &+ \frac{\left(1 + \nu_{6d}\nu_d^{-6}\right) \widetilde{M}_1}{(h_1^{(2)})^{3/2}} + l_{a,\tau}(\nu_d)^{-1/4}\frac{\left(1 + \nu_{6d}\nu_d^{-6}\right)^{5/4} \widetilde{M}_1^{5/4}}{(h_1^{(2)})^2} + \frac{\left(1 + \nu_{6d}\nu_d^{-6}\right)^{3/2} \widetilde{M}_1^{3/2}}{(h_1^{(2)})^2}\Bigg], \end{aligned}$$

*where* $h_1^{(2)}$ *is defined at* (4.1.5) *and* $l_{a,\tau}$ *is defined at* (4.1.7).

As an application of Theorem 4.1.4, we consider the case when the intensity of the underlying point process grows to infinity. The quantitative central limit theorem for this case is contained in the following result.

**Corollary 4.1.5.** *Let the assumptions (A) and (C) be in force. For a location-weight function* $h_1$ *and* $h_2(t) = \mathbf{1}\{t < a\}$ *for some* $a \in (0, \infty]$*, consider* $F^{(a)}(\eta_s)$ *defined at* (4.1.2) *evaluated at the Poisson process* $\eta_s$ *with* *intensity* $s\lambda \otimes \theta \otimes \nu$ *for* $s > 0$ *and* $\theta$ *given at* (4.1.6) *with* $\tau \in (-1, d-1]$. *Then, there exists a finite constant* $C > 0$ *depending only on* $h_1$, $d$, $a$, $\tau$, $\nu_d$, *and* $\nu_{6d}$, *such that, for all* $s \geq 1$,

$$\max\left\{ d_W\left(\frac{F^{(a)}(\eta_s) - \mathbb{E}F^{(a)}(\eta_s)}{\sqrt{\operatorname{Var} F^{(a)}(\eta_s)}}, N\right), d_K\left(\frac{F^{(a)}(\eta_s) - \mathbb{E}F^{(a)}(\eta_s)}{\sqrt{\operatorname{Var} F^{(a)}(\eta_s)}}, N\right)\right\} \leq C s^{-\frac{d}{2(d+\tau+1)}}.$$

As in the case of Theorem 4.1.3, the rate obtained in Corollary 4.1.5 is presumably optimal.

*Example* 4.1.6. Let $\eta_s$ be the Poisson process with intensity $s\lambda \otimes \theta \otimes \nu$ for $s \geq 1$, and $\theta$ given at (4.1.6) with $\tau \in (-1, d-1]$. Let $F^{(W,a)}(\eta_s)$ be the number of exposed points of $\eta_s$ that are located in the space-time window $W \times [0, a)$, for some Borel $W \subseteq \mathbb{R}^d$ with $\lambda(W) \geq 1$ and $a \in (0, +\infty]$. This corresponds to the functional defined at (4.1.2), evaluated at the Poisson process $\eta_s$ and with the weight function $h(\boldsymbol{x}) = \mathbf{1}_{W \times [0,a)}(x, t_x)$. Applying Theorem 4.1.4 and following the proof of Corollary 4.1.5 (see Section 4.3) yields that

$$d_W\left(\frac{F^{(W,a)}(\eta_s) - \mathbb{E}F^{(W,a)}(\eta_s)}{\sqrt{\operatorname{Var} F^{(W,a)}(\eta_s)}}, N\right) \leq C\gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1} s\nu_d\right)^{-\frac{1}{2}} s^{-\frac{d}{2(d+\tau+1)}} \lambda(W)^{-\frac{1}{2}},$$

and

$$d_K\left(\frac{F^{(W,a)}(\eta_s) - \mathbb{E}F^{(W,a)}(\eta_s)}{\sqrt{\operatorname{Var} F^{(W,a)}(\eta_s)}}, N\right) \\ \leq C\left[\gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1} s\nu_d\right)^{-\frac{1}{2}} + \gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1} s\nu_d\right)^{-\frac{3}{4}}\right] s^{-\frac{d}{2(d+\tau+1)}} \lambda(W)^{-\frac{1}{2}},$$

for a constant $C$ that depends only on $d$, $\nu_d$, $\nu_{6d}$ and $\tau$.

## 4.2 Variance estimation

In this section, we estimate the mean and variance of the statistic $F$, thus providing a proof of Proposition 4.1.2. Recall the weight function $h(\boldsymbol{x}) := h_1(x)h_2(t_x)$ and notice that by the Mecke equation (Proposition 2.1.11), the mean of $F$ is given by

$$\mathbb{E}F(\eta) = \int_{\mathbb{X}} h(\boldsymbol{x})\mathbb{E}H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}})\mu(\mathrm{d}\boldsymbol{x}) = \int_{\mathbb{R}^d} h_1(x)\mathrm{d}x \int_0^\infty h_2(t)w(t)\theta(\mathrm{d}t) = h_1^{(1)} \int_0^\infty h_2(t)w(t)\theta(\mathrm{d}t),$$

where $w(t)$ is defined at (4.1.10). In many instances, we will use the simple inequality

$$2ab \leq a^2 + b^2, \quad a, b \geq 0. \tag{4.2.1}$$

Also notice that for $x \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \lambda(B_{r_1}(0) \cap B_{r_2}(x))\mathrm{d}x = \int_{\mathbb{R}^d} \mathbf{1}_{y \in B_{r_1}(0)} \int_{\mathbb{R}^d} \mathbf{1}_{y \in B_{r_2}(x)} \mathrm{d}x\mathrm{d}y = \omega_d^2 r_1^d r_2^d. \tag{4.2.2}$$

The multivariate Mecke equation (Proposition 2.1.12) yields that

$$\mathrm{Var}(F) = \int_{\mathbb{X}} h(\boldsymbol{x})^2 \mathbb{E} H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}}) \mu(\mathrm{d}\boldsymbol{x}) - \Big( \int_{\mathbb{X}} h(\boldsymbol{x}) \mathbb{E} H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}}) \mu(\mathrm{d}\boldsymbol{x}) \Big)^2 \\ + \int_D h(\boldsymbol{x}) h(\boldsymbol{y}) \mathbb{E}[H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{y}} + \delta_{\boldsymbol{x}}) H_{\boldsymbol{y}}(\eta + \delta_{\boldsymbol{x}} + \delta_{\boldsymbol{y}})] \mu^2(\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})),$$

where the double integration is over the region $D \subseteq \mathbb{X}$ where the points $\boldsymbol{x}$ and $\boldsymbol{y}$ are incomparable ($\boldsymbol{x} \not\preceq \boldsymbol{y}$ and $\boldsymbol{y} \not\preceq \boldsymbol{x}$), i.e.,

$$D := \big\{ (\boldsymbol{x}, \boldsymbol{y}) : \|x - y\| > \max\{ v_x(t_y - t_x), v_y(t_x - t_y) \} \big\}.$$

It is possible to get rid of one of the Dirac measures in the inner integral, since on $D$ the points are incomparable. Thus, using the translation invariance of $\mathbb{E} H_{\boldsymbol{x}}(\eta)$, we have

$$\mathrm{Var}(F) = \int_{\mathbb{R}^d} h_1(x)^2 \mathrm{d}x \int_0^\infty h_2(t)^2 w(t) \theta(\mathrm{d}t) - I_0 + I_1, \tag{4.2.3}$$

where

$$I_0 := 2 \int_{\mathbb{X}^2} \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} h_1(x) h_1(y) h_2(t_x) h_2(t_y) w(t_x) w(t_y) \mu^2(\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})),$$

and

$$I_1 := \int_D h_1(x) h_1(y) h_2(t_x) h_2(t_y) \Big[ \mathbb{E}\big[ H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}}) H_{\boldsymbol{y}}(\eta + \delta_{\boldsymbol{y}}) \big] - w(t_x) w(t_y) \Big] \mu^2(\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})).$$

Finally, we will use the following simple inequality for the incomplete gamma function

$$\min\{1, b^x\} \gamma(x, y) \le \gamma(x, by) \le \max\{1, b^x\} \gamma(x, y), \tag{4.2.4}$$

which holds for all $x \ge 0$ and $b, y > 0$.

*Proof of Proposition 4.1.2.* First, notice that the term $I_1$ in (4.2.3) is non-negative, since

$$\mathbb{E}[H_{\boldsymbol{x}}(\eta) H_{\boldsymbol{y}}(\eta)] = e^{-\mu(L_{\boldsymbol{x}} \cup L_{\boldsymbol{y}})} \ge e^{-\mu(L_{\boldsymbol{x}})} e^{-\mu(L_{\boldsymbol{y}})} = w(t_x) w(t_y).$$

Furthermore, (4.2.1) yields that

$$I_0 \le \int_{\mathbb{X}} h_1(x)^2 h_2(t_x) w(t_x) \left[ \int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} h_2(t_y) w(t_y) \mu(\mathrm{d}\boldsymbol{y}) \right] \mu(\mathrm{d}\boldsymbol{x}) \\ + \int_{\mathbb{X}} h_1(y)^2 h_2(t_y) w(t_y) \left[ \int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} h_2(t_x) w(t_x) \mu(\mathrm{d}\boldsymbol{x}) \right] \mu(\mathrm{d}\boldsymbol{y}).$$

Since $\boldsymbol{y} \preceq \boldsymbol{x}$ is equivalent to $\|y-x\| \le v_y(t_x - t_y)$, the first summand on the right-hand side above can be simplified as

$$\begin{aligned}
&\int_{\mathbb{X}} h_1(x)^2 h_2(t_x) w(t_x) \left[\int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{y}\preceq\boldsymbol{x}} h_2(t_y) w(t_y) \mu(\mathrm{d}\boldsymbol{y})\right] \mu(\mathrm{d}\boldsymbol{x}) \\
&= \int_{\mathbb{R}^d} \int_0^\infty h_1(x)^2 h_2(t_x) w(t_x) \theta(\mathrm{d}t_x) \mathrm{d}x \int_0^\infty \int_0^{t_x} \omega_d v_y^d (t_x - t_y)^d h_2(t_y) w(t_y) \theta(\mathrm{d}t_y) \nu(\mathrm{d}v_y) \\
&= \omega_d \nu_d \int_{\mathbb{R}^d} h_1(x)^2 \mathrm{d}x \int_0^\infty \int_0^t (t-s)^d h_2(s) h_2(t) w(s) w(t) \theta(\mathrm{d}s) \theta(\mathrm{d}t).
\end{aligned}$$

The second summand in the bound on $I_0$, upon interchanging integrals for the second step, turns into

$$\begin{aligned}
&\int_{\mathbb{X}} h_1(y)^2 h_2(t_y) w(t_y) \left[\int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{y}\preceq\boldsymbol{x}} h_2(t_x) w(t_x) \mu(\mathrm{d}\boldsymbol{x})\right] \mu(\mathrm{d}\boldsymbol{y}) \\
&= \int_{\mathbb{R}^d} \int_0^\infty h_1(y)^2 h_2(t_y) w(t_y) \theta(\mathrm{d}t_y) \mathrm{d}y \int_0^\infty \int_{t_y}^{\infty} \omega_d v_y^d (t_x - t_y)^d h_2(t_x) w(t_x) \theta(\mathrm{d}t_x) \nu(\mathrm{d}v_y) \\
&= \omega_d \nu_d \int_{\mathbb{R}^d} h_1(y)^2 \mathrm{d}y \int_0^\infty \int_0^t (t-s)^d h_2(s) h_2(t) w(s) w(t) \theta(\mathrm{d}s) \theta(\mathrm{d}t).
\end{aligned}$$

Combining, by (4.2.3) we obtain (4.1.8).

To prove (4.1.9), note that by the Poincaré inequality (see [60, Sec. 18.3]),

$$\mathrm{Var}(F) \le \int_{\mathbb{X}} \mathbb{E}\left(F(\eta + \delta_{\boldsymbol{x}}) - F(\eta)\right)^2 \mu(\mathrm{d}\boldsymbol{x}).$$

Observe that $\eta$ is simple and for $x \notin \eta$

$$F(\eta + \delta_{\boldsymbol{x}}) - F(\eta) = h(\boldsymbol{x}) H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}}) - \sum_{\boldsymbol{y}\in\eta} h(\boldsymbol{y}) H_{\boldsymbol{y}}(\eta) \mathbf{1}_{\boldsymbol{y}\succ\boldsymbol{x}}\,.$$

The inequality

$$-\sum_{\boldsymbol{y}\in\eta} h(\boldsymbol{x}) h(\boldsymbol{y}) H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}}) H_{\boldsymbol{y}}(\eta) \mathbf{1}_{\boldsymbol{y}\succ\boldsymbol{x}} \le 0$$

in the first step and the Mecke equation in the second step yield that

$$\begin{aligned}
&\int_{\mathbb{X}} \mathbb{E}\left|F(\eta + \delta_{\boldsymbol{x}}) - F(\eta)\right|^2 \mu(\mathrm{d}\boldsymbol{x}) \\
&\le \int_{\mathbb{X}} \mathbb{E}\left[h(\boldsymbol{x})^2 H_{\boldsymbol{x}}(\eta + \delta_{\boldsymbol{x}})\right] \mu(\mathrm{d}\boldsymbol{x}) + \int_{\mathbb{X}} \mathbb{E}\Big[\sum_{\boldsymbol{y},\boldsymbol{z}\in\eta} \mathbf{1}_{\boldsymbol{y}\succ\boldsymbol{x}} \mathbf{1}_{\boldsymbol{z}\succ\boldsymbol{x}} h(\boldsymbol{y}) h(\boldsymbol{z}) H_{\boldsymbol{y}}(\eta) H_{\boldsymbol{z}}(\eta)\Big] \mu(\mathrm{d}\boldsymbol{x}) \qquad (4.2.5) \\
&= \int_{\mathbb{X}} h(\boldsymbol{x})^2 w(t_x) \mu(\mathrm{d}\boldsymbol{x}) + \int_{\mathbb{X}^2} \mathbf{1}_{\boldsymbol{y}\succ\boldsymbol{x}} h(\boldsymbol{y})^2 w(t_y) \mu^2(\mathrm{d}(\boldsymbol{x},\boldsymbol{y})) + \int_{\mathbb{X}} \int_{D_{\boldsymbol{x}}} h(\boldsymbol{y}) h(\boldsymbol{z}) e^{-\mu(L_{\boldsymbol{y}} \cup L_{\boldsymbol{z}})} \mu^2(\mathrm{d}(\boldsymbol{y},\boldsymbol{z})) \mu(\mathrm{d}\boldsymbol{x})\,,
\end{aligned}$$

where

$$D_{\boldsymbol{x}} := \{(\boldsymbol{y},\boldsymbol{z}) \in \mathbb{X}^2 : \boldsymbol{y} \succ \boldsymbol{x}, \boldsymbol{z} \succ \boldsymbol{x}, \boldsymbol{y} \not\succ \boldsymbol{z}, \boldsymbol{z} \not\succ \boldsymbol{y}\}\,.$$

Using that $xe^{-x/2} \le 1$ for $x \ge 0$, observe that

$$\int_{\mathbb{X}^2} \mathbf{1}_{\boldsymbol{y} \succ \boldsymbol{x}} h(\boldsymbol{y})^2 w(t_y) \mu^2(\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})) = \int_{\mathbb{X}} h(\boldsymbol{y})^2 w(t_y) \mu(L_{\boldsymbol{y}}) \mu(\mathrm{d}\boldsymbol{y}) \le \int_{\mathbb{X}} h(\boldsymbol{y})^2 w(t_y)^{1/2} \mu(\mathrm{d}\boldsymbol{y}). \tag{4.2.6}$$

Next, using that $\mu(L_{\boldsymbol{y}} \cup L_{\boldsymbol{z}}) \ge (\mu(L_{\boldsymbol{y}}) + \mu(L_{\boldsymbol{z}}))/2$ and that $D_{\boldsymbol{x}} \subseteq \{\boldsymbol{y}, \boldsymbol{z} \succ \boldsymbol{x}\}$ for the first inequality, and (4.2.1) for the second one, we have

$$\begin{aligned}
&\int_{\mathbb{X}} \int_{D_{\boldsymbol{x}}} h(\boldsymbol{y}) h(\boldsymbol{z}) e^{-\mu(L_{\boldsymbol{y}} \cup L_{\boldsymbol{z}})} \mu^2(\mathrm{d}(\boldsymbol{y}, \boldsymbol{z})) \mu(\mathrm{d}\boldsymbol{x}) \\
&\le \int_{\mathbb{X}} \int_{\mathbb{X}^2} \mathbf{1}_{\boldsymbol{y}, \boldsymbol{z} \succ \boldsymbol{x}} h(\boldsymbol{y}) h(\boldsymbol{z}) w(t_y)^{1/2} w(t_z)^{1/2} \mu^2(\mathrm{d}(\boldsymbol{y}, \boldsymbol{z})) \mu(\mathrm{d}\boldsymbol{x}) \\
&\le \int_{\mathbb{R}_+^2} h_2(t_y) h_2(t_z) w(t_y)^{1/2} w(t_z)^{1/2} \int_{\mathbb{R}^{2d}} h_1(z)^2 \left( \int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{x} \prec \boldsymbol{y}, \boldsymbol{z}} \mu(\mathrm{d}\boldsymbol{x}) \right) \mathrm{d}(y, z) \theta^2(\mathrm{d}(t_y, t_z)).
\end{aligned} \tag{4.2.7}$$

By (4.2.2),

$$\begin{aligned}
\int_{\mathbb{R}^d} \int_{\mathbb{X}} \mathbf{1}_{\boldsymbol{x} \prec \boldsymbol{y}, \boldsymbol{z}} \mu(\mathrm{d}\boldsymbol{x}) \mathrm{d}y &= \int_0^{t_y \wedge t_z} \int_0^\infty \nu(\mathrm{d}v_x) \theta(\mathrm{d}t_x) \int_{\mathbb{R}^d} \lambda(B_{v_x(t_y - t_x)}(y) \cap B_{v_x(t_z - t_x)}(z)) \mathrm{d}y \\
&= \omega_d^2 \nu_{2d} \int_0^{t_y \wedge t_z} (t_y - t_x)^d (t_z - t_x)^d \theta(\mathrm{d}t_x).
\end{aligned}$$

Plugging in (4.2.7), we obtain

$$\begin{aligned}
&\int_{\mathbb{X}} \int_{D_{\boldsymbol{x}}} h(\boldsymbol{y}) h(\boldsymbol{z}) e^{-\mu(L_{\boldsymbol{y}} \cup L_{\boldsymbol{z}})} \mu^2(\mathrm{d}(\boldsymbol{y}, \boldsymbol{z})) \mu(\mathrm{d}\boldsymbol{x}) \\
&\le \omega_d^2 \nu_{2d} h_1^{(2)} \int_{\mathbb{R}_+^2} \int_0^{t_1 \wedge t_2} (t_1 - s)^d (t_2 - s)^d h_2(t_1) h_2(t_2) w(t_1)^{1/2} w(t_2)^{1/2} \theta(\mathrm{d}s) \theta^2(\mathrm{d}(t_1, t_2)).
\end{aligned}$$

This in combination with (4.2.5) and (4.2.6) proves (4.1.9).

Now we move on to prove (4.1.11). We first confirm the lower bound. Fix $\tau \in (-1, d-1]$, as otherwise the bound is trivial, and $a \in (0, \infty]$. Then

$$\Lambda(t) = \omega_d \int_0^t (t - s)^d s^\tau \mathrm{d}s = \omega_d t^{d+\tau+1} B(d+1, \tau+1) = B\, \omega_d t^{d+\tau+1},$$

where $B := B(d+1, \tau+1)$ is a value of the Beta function. Hence, $w(t) = \exp\{-B\, \omega_d \nu_d t^{d+\tau+1}\}$. Plugging in, we obtain

$$\begin{aligned}
\frac{\mathrm{Var}(F)}{h_1^{(2)}} &\ge \int_0^a e^{-B\, \omega_d \nu_d t^{d+\tau+1}} \theta(\mathrm{d}t) - 2\omega_d \nu_d \int_0^a \int_0^t (t-s)^d e^{-B\, \omega_d \nu_d (s^{d+\tau+1} + t^{d+\tau+1})} \theta(\mathrm{d}s) \theta(\mathrm{d}t) \\
&= \left( \frac{1}{B\, \omega_d \nu_d} \right)^{\frac{\tau+1}{d+\tau+1}} \left[ \int_0^b e^{-t^{d+\tau+1}} t^\tau \mathrm{d}t - \frac{2}{B} \int_0^b \int_0^t (t-s)^d e^{-(s^{d+\tau+1} + t^{d+\tau+1})} t^\tau s^\tau \mathrm{d}s \mathrm{d}t \right],
\end{aligned} \tag{4.2.8}$$

where $b := a(B\,\omega_d\nu_d)^{1/(d+\tau+1)}$. Writing $s = tu$ for some $u \in [0,1]$,

$$
\begin{aligned}
&\frac{2}{B}\int_0^b\int_0^t (t-s)^d e^{-(s^{d+\tau+1}+t^{d+\tau+1})} t^\tau s^\tau \mathrm{d}s\mathrm{d}t \\
&\le \frac{2}{B}\int_0^b t^{d+2\tau+1}\int_0^1 (1-u)^d u^\tau e^{-t^{d+\tau+1}(u^{d+\tau+1}+1)}\mathrm{d}u\mathrm{d}t < 2\int_0^b t^{d+2\tau+1}e^{-t^{d+\tau+1}}\mathrm{d}t.
\end{aligned}
$$

By substituting $t^{d+\tau+1} = z$, it is easy to check that for any $\rho > -1$,

$$
\int_0^b e^{-t^{d+\tau+1}} t^\rho \mathrm{d}t = \frac{1}{d+\tau+1}\gamma\left(\frac{\rho+1}{d+\tau+1}, b^{d+\tau+1}\right),
$$

where $\gamma$ is the lower incomplete Gamma function. In particular, using that $x\gamma(x,y) > \gamma(x+1,y)$ for $x, y > 0$ we have

$$
\int_0^b e^{-t^{d+\tau+1}} t^{d+2\tau+1}\mathrm{d}t = \frac{1}{d+\tau+1}\gamma\left(1+\frac{\tau+1}{d+\tau+1}, b^{d+\tau+1}\right) < \frac{\tau+1}{(d+\tau+1)^2}\gamma\left(\frac{\tau+1}{d+\tau+1}, b^{d+\tau+1}\right).
$$

Thus, since $\tau \in (-1, d-1]$,

$$
\begin{aligned}
&\int_0^b e^{-t^{d+\tau+1}} t^\tau \mathrm{d}t - \frac{2}{B}\int_0^b\int_0^t (t-s)^d e^{-(s^{d+\tau+1}+t^{d+\tau+1})} t^\tau s^\tau \mathrm{d}s\mathrm{d}t \\
&\qquad\qquad > \gamma\left(\frac{\tau+1}{d+\tau+1}, b^{d+\tau+1}\right)\frac{1}{d+\tau+1}\left[1-\frac{2(\tau+1)}{d+\tau+1}\right] \ge 0.
\end{aligned}
$$

By (4.2.8) and (4.2.4), we obtain the lower bound in (4.1.11).

For the upper bound in (4.1.11), for $\theta$ as in (4.1.6), arguing as above we have

$$
\int_0^a w(t)^{1/2}\theta(\mathrm{d}t) = \int_0^a e^{-B\,\omega_d\nu_d t^{d+\tau+1}/2}\theta(\mathrm{d}t) = \frac{(2/B\,\omega_d\nu_d)^{\frac{\tau+1}{d+\tau+1}}}{d+\tau+1}\gamma\left(\frac{\tau+1}{d+\tau+1}, b^{d+\tau+1}\right).
$$

Finally, substituting $s' = (B\,\omega_d\nu_d)^{\frac{1}{d+\tau+1}}s$ and similarly for $t_1$ and $t_2$, it is straightforward to see that

$$
\begin{aligned}
&\nu_{2d}\int_{[0,a)^2}\int_0^{t_1\wedge t_1}(t_1-s)^d(t_2-s)^d w(t_1)^{1/2}w(t_2)^{1/2}\theta(\mathrm{d}s)\theta^2(\mathrm{d}(t_1,t_2)) \\
&\le C\nu_{2d}\nu_d^{-2}\nu_d^{-\frac{\tau+1}{d+\tau+1}}\left(\int_{\mathbb{R}_+} t^{d+\tau}e^{-t^{d+\tau+1}/4}\mathrm{d}t\right)^2\int_0^b s'^\tau e^{-s'^{d+\tau+1}/2}\mathrm{d}s' \\
&\le C'\nu_{2d}\nu_d^{-2}\nu_d^{-\frac{\tau+1}{d+\tau+1}}\gamma\left(\frac{\tau+1}{d+\tau+1}, \frac{b^{d+\tau+1}}{2}\right)
\end{aligned}
$$

for some constants $C, C'$ depending only on $d$ and $\tau$. The upper bound in (4.1.11) now follows from (4.1.9) upon using the above computation and (4.2.4). □

## 4.3 Proofs of the results from Section 4.1

In this section, we derive our main results from Theorem 2.4.2. Recall that for $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$, the score function $S(\boldsymbol{x}, \mathsf{M})$ is defined at (4.1.3). It is straightforward to check that if $S(\boldsymbol{x}, \mathsf{M}_1) = S(\boldsymbol{x}, \mathsf{M}_2)$ for some $\mathsf{M}_1, \mathsf{M}_2 \in \mathbf{N}_{\mathbb{X}}$ with $0 \neq \mathsf{M}_1 \leq \mathsf{M}_2$ (meaning that $\mathsf{M}_2 - \mathsf{M}_1$ is a non-negative measure) and $\boldsymbol{x} \in \mathsf{M}_1$, then $S(\boldsymbol{x}, \mathsf{M}_1) = S(\boldsymbol{x}, \mathsf{M}')$ for all $\mathsf{M}' \in \mathbf{N}_{\mathbb{X}}$ such that $\mathsf{M}_1 \leq \mathsf{M}' \leq \mathsf{M}_2$, so that Equation (2.4.1) holds. Next we check assumptions (A1) and (A2) from Section 2.4.

For $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ and $x \in \mathsf{M}$, define the stabilization region

$$R(\boldsymbol{x}, \mathsf{M}) := \begin{cases} L_{x,t_x} & \text{if } \boldsymbol{x} \text{ is exposed in } \mathsf{M}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Notice that

$$\{\mathsf{M} \in \mathbf{N}_{\mathbb{X}} \colon \boldsymbol{y} \in R(\boldsymbol{x}, \mathsf{M} + \delta_{\boldsymbol{x}})\} \in \mathcal{N}_{\mathbb{X}} \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{X}\,,$$

and that

$$\mathbb{P}\left\{\boldsymbol{y} \in R(\boldsymbol{x}, \eta + \delta_{\boldsymbol{x}})\right\} = \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} e^{-\mu(L_{x,t_x})} = \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} w(t_x)\,,$$

and

$$\mathbb{P}\{\{\boldsymbol{y}, \boldsymbol{z}\} \subseteq R(\boldsymbol{x}, \eta + \delta_{\boldsymbol{x}})\} = \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} \mathbf{1}_{\boldsymbol{z} \preceq \boldsymbol{x}} e^{-\mu(L_{x,t_x})} = \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} \mathbf{1}_{\boldsymbol{z} \preceq \boldsymbol{x}} w(t_x)$$

are measurable functions of $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{X}^2$ and $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \mathbb{X}^3$ respectively, with $w(t)$ defined at (4.1.10). It is not hard to see that $R$ is monotonically decreasing in the second argument, and that for all $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ and $\boldsymbol{x} \in \mathsf{M}$, $\mathsf{M}(R(x, \mathsf{M})) \geq 1$ implies that $\boldsymbol{x}$ is exposed, hence $(\mathsf{M} + \delta_{\boldsymbol{y}})(R(x, \mathsf{M} + \delta_{\boldsymbol{y}})) \geq 1$ for all $\boldsymbol{y} \notin R(\boldsymbol{x}, \mathsf{M})$. Moreover, the function $R$ satisfies

$$S(\boldsymbol{x}, \mathsf{M}) = S\Big(\boldsymbol{x}, \mathsf{M}_{R(\boldsymbol{x}, \mathsf{M})}\Big), \quad \mathsf{M} \in \mathbf{N}_{\mathbb{X}},\ \boldsymbol{x} \in \mathsf{M}\,,$$

where $\mathsf{M}_{R(\boldsymbol{x}, \mathsf{M})}$ denotes the restriction of the measure $\mathsf{M}$ to the region $R(\boldsymbol{x}, \mathsf{M})$. Hence, assumptions (A1.1)–(A1.4) from Section 2.4 are satisfied. Further, notice that for any $p \in (0, 1]$, for all $\mathsf{M} \in \mathbf{N}_{\mathbb{X}}$ with $\mathsf{M}(\mathbb{X}) \leq 7$, we have

$$\mathbb{E}\left[S(\boldsymbol{x}, \eta + \delta_{\boldsymbol{x}} + \mathsf{M})^{4+p}\right] \leq |h_1(x) h_2(t_x)|^{4+p} w(t),$$

confirming condition (A2) from Section 2.4 with $M_p(\boldsymbol{x}) := |h_1(x) h_2(t_x)|$. For definiteness, we take $p = 1$, and define

$$\widetilde{M}(\boldsymbol{x}) := \max\{M_1(\boldsymbol{x})^2, M_1(\boldsymbol{x})^4\} \leq \widetilde{M}(x) \tilde{h}(t_x),$$

where we recall that $\widetilde{M}(x) := \max\{h_1(x)^2, h_1(x)^4\}$ and let $\tilde{h}(t_x) := \max\{h_2(t_x)^2, h_2(t_x)^4\}$. Finally, define

$$r(\boldsymbol{x}, \boldsymbol{y}) := \begin{cases} \nu_d \Lambda(t_x), & \text{if } \boldsymbol{y} \preceq \boldsymbol{x}, \\ \infty, & \text{if } \boldsymbol{y} \not\preceq \boldsymbol{x}, \end{cases}$$

so that

$$\mathbb{P}\{\boldsymbol{y} \in R(\boldsymbol{x}, \eta + \delta_{\boldsymbol{x}})\} = \mathbf{1}_{\boldsymbol{y} \preceq \boldsymbol{x}} w(t_x) = e^{-r(\boldsymbol{x}, \boldsymbol{y})}, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{X},$$

which corresponds to Equation (2.4.3) from Section 2.4. Now that we have checked all the necessary conditions, we can invoke Theorem 2.4.2. Let $\zeta := \frac{p}{40+10p} = 1/50$ and

$$g(\boldsymbol{y}) := \int_{\mathbb{X}} e^{-\zeta r(\boldsymbol{x}, \boldsymbol{y})} \mu(\mathrm{d}\boldsymbol{x}), \tag{4.3.1}$$

$$G(\boldsymbol{y}) := \widetilde{M}(y) \tilde{h}(t_y) \big(1 + g(\boldsymbol{y})^5\big), \quad \boldsymbol{y} \in \mathbb{X}. \tag{4.3.2}$$

For $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$, denote

$$q(\boldsymbol{x}, \boldsymbol{y}) := \int_{\mathbb{X}} \mathbb{P}\Big\{\{\boldsymbol{x}, \boldsymbol{y}\} \subseteq R\big(\boldsymbol{z}, \eta + \delta_{\boldsymbol{z}}\big)\Big\} \mu(\mathrm{d}\boldsymbol{z}) = \int_{\boldsymbol{x} \preceq \boldsymbol{z}, \boldsymbol{y} \preceq \boldsymbol{z}} w(t_z) \mu(\mathrm{d}\boldsymbol{z}). \tag{4.3.3}$$

For $\alpha > 0$, let

$$f_\alpha(\boldsymbol{y}) := f_\alpha^{(1)}(\boldsymbol{y}) + f_\alpha^{(2)}(\boldsymbol{y}) + f_\alpha^{(3)}(\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{X},$$

where for $\boldsymbol{y} \in \mathbb{X}$,

$$\begin{aligned} f_\alpha^{(1)}(\boldsymbol{y}) &:= \int_{\mathbb{X}} G(\boldsymbol{x}) e^{-\alpha r(\boldsymbol{x}, \boldsymbol{y})} \mu(\mathrm{d}\boldsymbol{x}) = \int_{\boldsymbol{y} \preceq \boldsymbol{x}} G(\boldsymbol{x}) w(t_x)^\alpha \mu(\mathrm{d}\boldsymbol{x}), \\ f_\alpha^{(2)}(\boldsymbol{y}) &:= \int_{\mathbb{X}} G(\boldsymbol{x}) e^{-\alpha r(\boldsymbol{y}, \boldsymbol{x})} \mu(\mathrm{d}\boldsymbol{x}) = w(t_y)^\alpha \int_{\boldsymbol{x} \preceq \boldsymbol{y}} G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}), \\ f_\alpha^{(3)}(\boldsymbol{y}) &:= \int_{\mathbb{X}} G(\boldsymbol{x}) q(\boldsymbol{x}, \boldsymbol{y})^\alpha \mu(\mathrm{d}\boldsymbol{x}). \end{aligned} \tag{4.3.4}$$

Finally, let

$$\kappa(\boldsymbol{x}) := \mathbb{P}\{S(\boldsymbol{x}, \eta + \delta_{\boldsymbol{x}}) \neq 0\} = \mathbf{1}_{\{h(\boldsymbol{x}) \neq 0\}} w(t_x), \quad \boldsymbol{x} \in \mathbb{X}.$$

For an integrable function $f : \mathbb{X} \to \mathbb{R}$, denote $\mu f := \int_{\mathbb{X}} f(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x})$. With $\beta := \frac{p}{32+4p} = 1/36$, Theorem 2.4.2 yields that $F = F(\eta)$ as in (4.1.2) satisfies

$$d_W\left(\frac{F - \mathbb{E}F}{\sqrt{\operatorname{Var} F}}, N\right) \leq C\left[\frac{\sqrt{\mu f_\beta^2}}{\operatorname{Var} F} + \frac{\mu((\kappa + g)^{2\beta} G)}{(\operatorname{Var} F)^{3/2}}\right], \tag{4.3.5}$$

and

$$d_K\left(\frac{F-\mathbb{E}F}{\sqrt{\operatorname{Var}F}},N\right)\le C\Bigg[\frac{\sqrt{\mu f_\beta^2}+\sqrt{\mu f_{2\beta}}}{\operatorname{Var}F}+\frac{\sqrt{\mu((\kappa+g)^{2\beta}G)}}{\operatorname{Var}F}$$
$$+\frac{\mu((\kappa+g)^{2\beta}G)}{(\operatorname{Var}F)^{3/2}}+\frac{(\mu((\kappa+g)^{2\beta}G))^{5/4}+(\mu((\kappa+g)^{2\beta}G))^{3/2}}{(\operatorname{Var}F)^2}\Bigg], \quad (4.3.6)$$

where $N$ is a standard normal random variable and $C\in(0,\infty)$ is a constant.

In the rest of this section, we estimate the summands on the right-hand side of the above two bounds. We start with a simple lemma.

**Lemma 4.3.1.** *For all $a\ge 0$, $b>0$ and $i\in\{0,1\}$,*

$$Q_i(a,b):=\int_0^\infty t^a\,\tilde h(t)^i\,e^{-b\Lambda(t)}\,\theta(\mathrm{d}t)<\infty. \quad (4.3.7)$$

*Proof.* Fix $i\in\{0,1\}$. Assume that $\theta([0,c])>0$ for some $c>0$, since otherwise the result holds trivially. Also, it suffices to show the finiteness of the integral over $[2c,\infty)$, as

$$\int_0^{2c} t^a\,\tilde h(t)^i\,e^{-b\Lambda(t)}\theta(\mathrm{d}t)\le (2c)^a\int_0^{2c}\tilde h(t)^i\,e^{-b\Lambda(t)}\theta(\mathrm{d}t)<\infty$$

by assumption (B). The inequality $x^{a/d}e^{-x/2}\le C$ for some finite constant $C>0$ yields that

$$\int_{2c}^\infty t^a\,\tilde h(t)^i\,e^{-b\Lambda(t)}\theta(\mathrm{d}t)\le\frac{C}{b^{a/d}}\int_{2c}^\infty\frac{t^a}{\Lambda(t)^{a/d}}\,\tilde h(t)^i\,e^{-b\Lambda(t)/2}\theta(\mathrm{d}t).$$

For $t\ge 2c$,

$$\Lambda(t)=\int_0^t(t-s)^d\theta(\mathrm{d}s)\ge\int_0^{t/2}(t-s)^d\theta(\mathrm{d}s)\ge(t/2)^d\theta([0,t/2])\ge 2^{-d}t^d\theta([0,c]).$$

Thus,

$$\int_{2c}^\infty t^a\,\tilde h(t)^i\,e^{-b\Lambda(t)}\theta(\mathrm{d}t)\le\frac{C2^a}{(b\theta([0,c]))^{a/d}}\int_{2c}^\infty\tilde h(t)^i\,e^{-b\Lambda(t)/2}\theta(\mathrm{d}t)<\infty$$

by assumption (B), yielding the result. □

To compute the bounds in (4.3.5) and (4.3.6), we need to bound $\mu f_{2\beta}$ and $\mu f_\beta^2$, with $\beta=1/36$. Nonetheless, we provide bounds on $\mu f_\alpha$ and $\mu f_\alpha^2$ for any $\alpha>0$. By Jensen's inequality, it suffices to bound $\mu f_\alpha^{(i)}$ and $\mu(f_\alpha^{(i)})^2$ for $i=1,2,3$. This is the objective of the following three lemmas.

For $g$ defined at (4.3.1)

$$g(\boldsymbol{y})=\int_{\mathbb{X}}\mathbf{1}_{\boldsymbol{y}\preceq\boldsymbol{x}}w(t_x)^\zeta\mu(\mathrm{d}\boldsymbol{x})=\int_{t_y}^\infty\int_{\mathbb{R}^d}\mathbf{1}_{x\in B_{v_y(t_x-t_y)}(y)}w(t_x)^\zeta\,\mathrm{d}x\theta(\mathrm{d}t_x)$$
$$=\omega_d v_y^d\int_{t_y}^\infty(t_x-t_y)^d w(t_x)^\zeta\,\theta(\mathrm{d}t_x)\le\omega_d v_y^d\int_0^\infty t_x^d w(t_x)^\zeta\,\theta(\mathrm{d}t_x)=\omega_d v_y^d Q_0(d,\zeta\nu_d).$$

where $Q_0$ is defined at (4.3.7). Therefore, the function $G$ defined at (4.3.2) is given by

$$G(\boldsymbol{y}) = \widetilde{M}(y)\tilde{h}(t_y)\left(1 + g(\boldsymbol{y})^5\right) \le \omega_d^5\,\widetilde{M}(y)\tilde{h}(t_y)(1 + Qv_y^{5d}), \tag{4.3.8}$$

where

$$Q \equiv Q(\nu_d) := Q_0(d, \zeta\nu_d)^5.$$

**Lemma 4.3.2.** *For any $\alpha > 0$ and $f_\alpha^{(1)}$ defined at* (4.3.4),

$$\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C_1\widetilde{M}_1 \quad \textit{and} \quad \int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) \le C_2\widetilde{M}_2,$$

*where*

$$C_1 := \frac{\omega_d^5}{\alpha}Q_1(0, \alpha\nu_d/2)(1 + Q\nu_{5d}) \quad \textit{and} \quad C_2 := \omega_d^{12}Q_1(d, \alpha\nu_d/2)^2 Q_0(0, \alpha\nu_d)(1 + Q\nu_{5d})^2\nu_{2d}.$$

*Proof.* Using $xe^{-x/2} \le 1$ for $x \ge 0$, we can write

$$\begin{aligned}
\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) &\le \int_{\mathbb{X}}\int_{\boldsymbol{y}\preceq\boldsymbol{x}} G(\boldsymbol{x})w(t_x)^\alpha\mu(\mathrm{d}\boldsymbol{x})\mu(\mathrm{d}\boldsymbol{y}) \\
&= \omega_d^5\,\nu_d\int_{\mathbb{X}}\Lambda(t_x)\widetilde{M}(x)\tilde{h}(t_x)\left(1 + Qv_x^{5d}\right)w(t_x)^\alpha\mu(\mathrm{d}\boldsymbol{x}) \\
&= \omega_d^5\,(1 + Q\nu_{5d})\widetilde{M}_1\int_0^\infty \nu_d\Lambda(t_x)w(t_x)^\alpha\tilde{h}(t_x)\theta(\mathrm{d}t_x) \\
&\le \frac{\omega_d^5}{\alpha}Q_1(0, \alpha\nu_d/2)(1 + Q\nu_{5d})\widetilde{M}_1.
\end{aligned}$$

For the second assertion, changing the order of the integrals in the second step and using (4.3.8) for the final step, we get

$$\begin{aligned}
\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) &= \int_{\mathbb{X}}\int_{\boldsymbol{y}\preceq\boldsymbol{x}_1}\int_{\boldsymbol{y}\preceq\boldsymbol{x}_2} G(\boldsymbol{x}_1)w(t_{x_1})^\alpha G(\boldsymbol{x}_2)w(t_{x_2})^\alpha\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2)\mu(\mathrm{d}\boldsymbol{y}) \\
&= \int_{\mathbb{X}}\int_{\mathbb{X}}\left(\int_{\boldsymbol{y}\preceq\boldsymbol{x}_1,\boldsymbol{y}\preceq\boldsymbol{x}_2}\mu(\mathrm{d}\boldsymbol{y})\right)G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)(w(t_{x_1})w(t_{x_2}))^\alpha\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2) \\
&= \int_{\mathbb{X}}\int_{\mathbb{X}}\mu(L_{x_1,t_{x_1}} \cap L_{x_2,t_{x_2}})G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)(w(t_{x_1})w(t_{x_2}))^\alpha\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2) \\
&\le \omega_d^{10}(1 + Q\nu_{5d})^2\iint_{\mathbb{R}_+^2\times\mathbb{R}^{2d}}\mu(L_{x_1,t_{x_1}} \cap L_{x_2,t_{x_2}})\widetilde{M}(x_1)\widetilde{M}(x_2) \\
&\qquad \times (w(t_{x_1})w(t_{x_2}))^\alpha\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\mathrm{d}(x_1, x_2)\theta^2(\mathrm{d}(t_{x_1}, t_{x_2})).
\end{aligned} \tag{4.3.9}$$

By (4.2.2), for any $t_1, t_2 \in \mathbb{R}_+$,

$$\begin{aligned}
\int_{\mathbb{R}^d}\mu(L_{0,t_1} \cap L_{x,t_2})\mathrm{d}x &= \int_0^{t_1\wedge t_2}\theta(\mathrm{d}s)\int_0^\infty\nu(\mathrm{d}v)\int_{\mathbb{R}^d}\lambda(B_{v(t_1-s)}(0) \cap B_{v(t_2-s)}(x))\mathrm{d}x \\
&= \omega_d^2\int_0^\infty v^{2d}\nu(\mathrm{d}v)\int_0^{t_1\wedge t_2}(t_1 - s)^d(t_2 - s)^d\theta(\mathrm{d}s) = \omega_d^2\nu_{2d}\int_0^{t_1\wedge t_2}(t_1 - s)^d(t_2 - s)^d\theta(\mathrm{d}s).
\end{aligned}$$

From (4.3.9), using (4.2.1) for the first inequality and the above equality in the second step, we obtain

$$
\begin{aligned}
&\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) \\
&\le \omega_d^{10}(1+Q\nu_{5d})^2\int_{\mathbb{R}^d}\widetilde{M}(x_1)^2\mathrm{d}x_1 \\
&\qquad\times\int_{\mathbb{R}_+^2}\left(\int_{\mathbb{R}^d}\mu(L_{0,t_{x_1}}\cap L_{x_2-x_1,t_{x_2}})\mathrm{d}x_2\right)(w(t_{x_1})w(t_{x_2}))^\alpha\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\theta^2(\mathrm{d}(t_{x_1},t_{x_2})) \\
&= \omega_d^{12}(1+Q\nu_{5d})^2\nu_{2d}\widetilde{M}_2 \\
&\qquad\times\int_{\mathbb{R}_+^2}\int_0^{t_{x_1}\wedge t_{x_2}}(t_{x_1}-s)^d(t_{x_2}-s)^d(w(t_{x_1})w(t_{x_2}))^\alpha\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\theta(\mathrm{d}s)\theta^2(\mathrm{d}(t_{x_1},t_{x_2})) \\
&= \omega_d^{12}(1+Q\nu_{5d})^2\nu_{2d}\widetilde{M}_2 \\
&\qquad\times\int_0^\infty\int_s^\infty\int_s^\infty(t_{x_1}-s)^d(t_{x_2}-s)^d(w(t_{x_1})w(t_{x_2}))^\alpha\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\theta(\mathrm{d}t_{x_1})\theta(\mathrm{d}t_{x_2})\theta(\mathrm{d}s) \\
&= \omega_d^{12}(1+Q\nu_{5d})^2\nu_{2d}\widetilde{M}_2\int_0^\infty\left(\int_s^\infty(t-s)^d w(t)^\alpha\tilde{h}(t)\theta(\mathrm{d}t)\right)^2\theta(\mathrm{d}s) \qquad (4.3.10)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) &\le \omega_d^{12}(1+Q\nu_{5d})^2\nu_{2d}\widetilde{M}_2\int_0^\infty\left(\int_0^\infty t^d w(t)^{\alpha/2}\tilde{h}(t)\theta(\mathrm{d}t)\right)^2 w(s)^\alpha\theta(\mathrm{d}s) \\
&\le \omega_d^{12}Q_1(d,\alpha\nu_d/2)^2 Q_0(0,\alpha\nu_d)(1+Q\nu_{5d})^2\nu_{2d}\widetilde{M}_2,
\end{aligned}
$$

where we have used the fact that $w$ is a decreasing function. □

**Lemma 4.3.3.** *For any $\alpha>0$ and $f_\alpha^{(2)}$ defined at (4.3.4),*

$$
\int_{\mathbb{X}} f_\alpha^{(2)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C_1\widetilde{M}_1 \quad \text{and} \quad \int_{\mathbb{X}} f_\alpha^{(2)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) \le C_2\widetilde{M}_2
$$

*for*

$$
\begin{aligned}
C_1 &:= \omega_d^6 Q_1(0,\alpha\nu_d/2)Q_0(d,\alpha\nu_d/2)(\nu_d+Q\nu_{6d}), \\
C_2 &:= \omega_d^{12}\, Q_1(0,\alpha\nu_d/3)^2 Q_0(2d,\alpha\nu_d/3)(\nu_d+Q\nu_{6d})^2.
\end{aligned}
$$

*Proof.* By the definition of $f_\alpha^{(2)}$,

$$
\begin{aligned}
\int_{\mathbb{X}} f_\alpha^{(2)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) &\le \omega_d^5\int_{\mathbb{X}}\left(\int_{\boldsymbol{x}\preceq\boldsymbol{y}} w(t_y)^\alpha\mu(\mathrm{d}\boldsymbol{y})\right)\widetilde{M}(x)\tilde{h}(t_x)(1+Qv_x^{5d})\mu(\mathrm{d}\boldsymbol{x}) \\
&= \omega_d^6\widetilde{M}_1\int_0^\infty\int_{t_x}^\infty w(t_y)^\alpha(t_y-t_x)^d\tilde{h}(t_x)\theta(\mathrm{d}t_y)\theta(\mathrm{d}t_x)\int_0^\infty v_x^d(1+Qv_x^{5d})\nu(\mathrm{d}v_x) \\
&\le \omega_d^6(\nu_d+Q\nu_{6d})\widetilde{M}_1\int_0^\infty w(t_x)^{\alpha/2}\tilde{h}(t_x)\theta(\mathrm{d}t_x)\int_0^\infty t_y^d w(t_y)^{\alpha/2}\theta(\mathrm{d}t_y) \\
&\le \omega_d^6 Q_1(0,\alpha\nu_d/2)Q_0(d,\alpha\nu_d/2)(\nu_d+Q\nu_{6d})\widetilde{M}_1,
\end{aligned}
$$

where in the penultimate step, we have used that $w$ is decreasing. This proves the first assertion.

For the second assertion, using (4.3.8) in the third step and (4.2.1) in the final step, we have

$$\begin{aligned}\int_{\mathbb{X}} f_\alpha^{(2)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) &= \int_{\mathbb{X}} w(t_y)^{2\alpha}\left(\int_{\boldsymbol{x}_1\preceq\boldsymbol{y}} G(\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_1)\int_{\boldsymbol{x}_2\preceq\boldsymbol{y}} G(\boldsymbol{x}_2)\mu(\mathrm{d}\boldsymbol{x}_2)\right)\mu(\mathrm{d}\boldsymbol{y})\\
&= \int_{\mathbb{X}}\int_{\mathbb{X}}\left(\int_{\boldsymbol{x}_1\preceq\boldsymbol{y},\boldsymbol{x}_2\preceq\boldsymbol{y}} w(t_y)^{2\alpha}\mu(\mathrm{d}\boldsymbol{y})\right) G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2)\\
&\le \omega_d^{10}\int_{\mathbb{X}}\int_{\mathbb{X}}(1+Qv_{x_1}^{5d})(1+Qv_{x_2}^{5d})\widetilde{M}(x_1)\widetilde{M}(x_2)\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\left(\int_{\boldsymbol{x}_1\preceq\boldsymbol{y},\boldsymbol{x}_2\preceq\boldsymbol{y}} w(t_y)^{2\alpha}\mu(\mathrm{d}\boldsymbol{y})\right)\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2)\\
&\le \omega_d^{10}\int_{\mathbb{X}}\int_{\mathbb{X}}(1+Qv_{x_1}^{5d})(1+Qv_{x_2}^{5d})\widetilde{M}(x_1)^2\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\left(\int_{\boldsymbol{x}_1\preceq\boldsymbol{y},\boldsymbol{x}_2\preceq\boldsymbol{y}} w(t_y)^{2\alpha}\mu(\mathrm{d}\boldsymbol{y})\right)\mu(\mathrm{d}\boldsymbol{x}_1)\mu(\mathrm{d}\boldsymbol{x}_2)\,. \quad (4.3.11)\end{aligned}$$

For fixed $\boldsymbol{x}_1, t_{x_2}$ and $v_{x_2}$, we have

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{\boldsymbol{x}_1\preceq\boldsymbol{y},\boldsymbol{x}_2\preceq\boldsymbol{y}} w(t_y)^{2\alpha}\mu(\mathrm{d}\boldsymbol{y})\mathrm{d}x_2 &= \int_{t_{x_1}\vee t_{x_2}}^\infty w(t_y)^{2\alpha}\left(\int_{\mathbb{R}^d}\lambda(B_{v_{x_1}(t_y-t_{x_1})}(0)\cap B_{v_{x_2}(t_y-t_{x_2})}(x))\mathrm{d}x\right)\theta(\mathrm{d}t_y)\\
&= \omega_d^2 v_{x_1}^d v_{x_2}^d\int_{t_{x_1}\vee t_{x_2}}^\infty (t_y-t_{x_1})^d(t_y-t_{x_2})^d w(t_y)^{2\alpha}\theta(\mathrm{d}t_y)\,.\end{aligned}$$

Noticing that

$$\int_{\mathbb{R}_+^2} v_{x_1}^d v_{x_2}^d(1+Qv_{x_1}^{5d})(1+Qv_{x_2}^{5d})\,\nu^2(\mathrm{d}(v_{x_1},v_{x_2})) = (\nu_d+Q\nu_{6d})^2\,,$$

and, arguing similarly as for $\mu(f_\alpha^{(2)})$ above and using (4.3.11), we obtain

$$\begin{aligned}\int_{\mathbb{X}} f_\alpha^{(2)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) &\le \omega_d^{12}\,(\nu_d+Q\nu_{6d})^2\int_{\mathbb{R}^d}\widetilde{M}(x_1)^2\mathrm{d}x_1\\
&\quad\times\int_{\mathbb{R}_+^2}\tilde{h}(t_{x_1})\tilde{h}(t_{x_2})\left(\int_{t_{x_1}\vee t_{x_2}}^\infty (t_y-t_{x_1})^d(t_y-t_{x_2})^d w(t_y)^\alpha\theta(\mathrm{d}t_y)\right)\theta^2(\mathrm{d}(t_{x_1},t_{x_2}))\\
&\le \omega_d^{12}(\nu_d+Q\nu_{6d})^2\widetilde{M}_2\int_0^\infty w(t_{x_1})^{\alpha/3}\tilde{h}(t_{x_1})\theta(\mathrm{d}t_{x_1})\int_0^\infty w(t_{x_2})^{\alpha/3}\tilde{h}(t_{x_2})\theta(\mathrm{d}t_{x_2})\int_0^\infty t_y^{2d}w(t_y)^{\alpha/3}\theta(\mathrm{d}t_y)\\
&\le \omega_d^{12}\,Q_1(0,\alpha\nu_d/3)^2 Q_0(2d,\alpha\nu_d/3)(\nu_d+Q\nu_{6d})^2\widetilde{M}_2\,,\end{aligned}$$

yielding the desired conclusion. □

Before proceeding to bound the integrals of $f^{(3)}$, notice that, since $\theta$ is a non-null measure,

$$\begin{aligned}Q'_\alpha = Q'_\alpha(\nu_d) :&= \int_0^\infty t^{d-1}e^{-\frac{\alpha\nu_d}{3}\Lambda(t)}\mathrm{d}t = \int_0^\infty t^{d-1}e^{-\frac{\alpha\omega_d\nu_d}{3}\int_0^t(t-s)^d\theta(\mathrm{d}s)}\mathrm{d}t \quad (4.3.12)\\
&\le \int_0^\infty t^{d-1}e^{-\frac{\alpha\omega_d\nu_d}{3}\int_0^{t/2}(t/2)^d\theta(\mathrm{d}s)}\mathrm{d}t = \int_0^\infty t^{d-1}e^{-\frac{\alpha\omega_d\nu_d}{3}\theta([0,t/2))(t/2)^d}\mathrm{d}t < \infty\,.\end{aligned}$$

**Lemma 4.3.4.** *For any* $\alpha \in (0,1]$ *and* $f_\alpha^{(3)}$ *defined at* (4.3.4),

$$\int_{\mathbb{X}} f_\alpha^{(3)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C_1 \widetilde{M}_1 \quad \textit{and} \quad \int_{\mathbb{X}} f_\alpha^{(3)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) \le C_2 \widetilde{M}_2,$$

*where*

$$C_1 := C\, Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)(\nu_d + Q\nu_{6d})\Big[Q'_\alpha + \nu_d Q_0(2d,\alpha\nu_d/3)\Big]\,,$$
$$C_2 := C\, Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)^2\Big[Q'^2_\alpha\Big[\nu_{2d}(1+Q\nu_{5d})^2 + (\nu_d + Q\nu_{6d})^2\Big] + Q_0(2d,\alpha\nu_d/3)^2\nu_{2d}(\nu_d + Q\nu_{6d})^2\Big]\,,$$

*and* $C$ *is a positive constant that depends only on* $d$.

*Proof.* Note that $\boldsymbol{x},\boldsymbol{y} \preceq \boldsymbol{z}$ implies

$$|x-y| \le |x-z| + |y-z| \le t_z(v_x+v_y).$$

For $q$ defined at (4.3.3), we have

$$q(\boldsymbol{x},\boldsymbol{y}) \le e^{-\nu_d\Lambda(r_0)}\int_{r_0}^\infty \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y-x))e^{-\nu_d(\Lambda(t_z)-\Lambda(r_0))}\theta(\mathrm{d}t_z)\,,$$

where

$$r_0 = r_0(\boldsymbol{x},\boldsymbol{y}) := \frac{|x-y|}{v_x+v_y} \vee t_x \vee t_y\,.$$

Therefore,

$$\begin{aligned} q(\boldsymbol{x},\boldsymbol{y})^\alpha &\le e^{-\alpha\nu_d\Lambda(r_0)}\left(1 + \int_{r_0}^\infty \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y-x))e^{-\nu_d(\Lambda(t_z)-\Lambda(r_0))}\theta(\mathrm{d}t_z)\right) \\ &\le e^{-\alpha\nu_d\Lambda(r_0)} + \int_{r_0}^\infty \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y-x))e^{-\alpha\nu_d\Lambda(t_z)}\theta(\mathrm{d}t_z)\,. \end{aligned} \tag{4.3.13}$$

Then, with $f_\alpha^{(3)}$ defined at (4.3.4),

$$\begin{aligned} \int_{\mathbb{X}} f_\alpha^{(3)}(\boldsymbol{y})\mu(\boldsymbol{y}) &\le \int_{\mathbb{X}^2} G(\boldsymbol{x})e^{-\alpha\nu_d\Lambda(r_0)}\mu^2(\mathrm{d}(\boldsymbol{x},\boldsymbol{y})) \\ &\quad + \int_{\mathbb{X}^2} G(\boldsymbol{x})\int_{r_0}^\infty \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y-x))e^{-\alpha\nu_d\Lambda(t_z)}\theta(\mathrm{d}t_z)\mu^2(\mathrm{d}(\boldsymbol{x},\boldsymbol{y})). \end{aligned} \tag{4.3.14}$$

Since $\Lambda$ is increasing,

$$\exp\{-\alpha\nu_d\Lambda(r_0(\boldsymbol{x},\boldsymbol{y}))\} \le \exp\left\{-\frac{\alpha\nu_d}{3}\left[\Lambda\left(\frac{|x-y|}{v_x+v_y}\right) + \Lambda(t_x) + \Lambda(t_y)\right]\right\}, \tag{4.3.15}$$

and, by a change of variable and passing to polar coordinates, we obtain

$$\int_{\mathbb{R}^d} e^{-\frac{\alpha\nu_d}{3}\Lambda\left(\frac{|x|}{v_x+v_y}\right)}\,\mathrm{d}x \le d\omega_d(v_x+v_y)^d\int_0^\infty \rho^{d-1}e^{-\frac{\alpha\nu_d}{3}\Lambda(\rho)}\mathrm{d}\rho = d\omega_d(v_x+v_y)^dQ'_\alpha\,. \tag{4.3.16}$$

Thus, using (4.3.8), (4.3.15) and (4.3.16), we can bound the first summand on the right-hand side of (4.3.14) as

$$
\begin{aligned}
&\int_{\mathbb{X}^2} G(\boldsymbol{x}) e^{-\alpha\nu_d\Lambda(r_0)}\mu^2(\mathrm{d}(\boldsymbol{x},\boldsymbol{y})) \\
&\le \omega_d^5 \int_0^\infty w(t_x)^{\alpha/3}\tilde{h}(t_x)\mathrm{d}t_x \int_0^\infty w(t_y)^{\alpha/3}\mathrm{d}t_y \\
&\qquad\qquad\qquad \times \int_{\mathbb{R}^d} \widetilde{M}(x)\mathrm{d}x \iint_{\mathbb{R}_+^2\times\mathbb{R}^d} (1+Qv_x^{5d}) w\left(\frac{|x-y|}{v_x+v_y}\right)^{\alpha/3} \mathrm{d}y\, \nu^2(\mathrm{d}(v_x,v_y)) \\
&\le d\omega_d^6 Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)Q'_\alpha\, \widetilde{M}_1 \int_{\mathbb{R}_+^2} (1+Qv_x^{5d})(v_x+v_y)^d\nu^2(\mathrm{d}(v_x,v_y)) \\
&\le d2^d\omega_d^6 Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)Q'_\alpha(\nu_d+Q\nu_{6d})\, \widetilde{M}_1\,,
\end{aligned}
$$

where for the final step, we have used Jensen's inequality and the fact that $\nu_d\nu_{5d} \le \nu_{6d}$, which is a consequence of positive association, since $v^d$ and $v^{5d}$ are both increasing functions of $v$. Arguing similarly for the second summand in (4.3.14), using (4.3.8) in the first and (4.2.2) in the second step, we obtain

$$
\begin{aligned}
&\int_{\mathbb{X}^2} G(\boldsymbol{x}) \int_{r_0}^\infty \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y-x)) e^{-\alpha\nu_d\Lambda(t_z)}\theta(\mathrm{d}t_z)\mu^2(\mathrm{d}(\boldsymbol{x},\boldsymbol{y})) \\
&\le \omega_d^5 \int_{\mathbb{R}^d} \widetilde{M}(x)\mathrm{d}x \int_{\mathbb{R}_+^2}\int_{\mathbb{R}_+^2} (1+Qv_x^{5d}) \int_{t_x\vee t_y}^\infty w(t_z)^\alpha \\
&\qquad \times \left(\int_{\mathbb{R}^d} \lambda(B_{v_x(t_z-t_x)}(0)\cap B_{v_y(t_z-t_y)}(y))\mathrm{d}y\right) \tilde{h}(t_x)\theta(\mathrm{d}t_z)\theta^2(\mathrm{d}(t_x,t_y))\, \nu^2(\mathrm{d}(v_x,v_y)) \\
&\le \omega_d^7\, \widetilde{M}_1 \int_{\mathbb{R}_+^2} (1+Qv_x^{5d})v_x^d v_y^d\, \nu^2(\mathrm{d}(v_x,v_y)) \int_{\mathbb{R}_+^3} t_z^{2d} w(t_z)^{\alpha/3}w(t_x)^{\alpha/3}w(t_y)^{\alpha/3}\tilde{h}(t_x)\theta^3(\mathrm{d}(t_z,t_x,t_y)) \\
&\le \omega_d^7\, Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)Q_0(2d,\alpha\nu_d/3)\nu_d(\nu_d+Q\nu_{6d})\widetilde{M}_1.
\end{aligned}
$$

This concludes the proof of the first assertion.

Next, we prove the second assertion. For ease of notation, we drop obvious subscripts and write $\boldsymbol{y} = (y,s,v)$, $\boldsymbol{x}_1 = (x_1,t_1,u_1)$ and $\boldsymbol{x}_2 = (x_2,t_2,u_2)$. Using (4.3.13), write

$$
\begin{aligned}
\int_{\mathbb{X}} f_\alpha^{(3)}(\boldsymbol{y})^2\mu(\boldsymbol{y}) &= \int_{\mathbb{X}^3} G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)q(\boldsymbol{x}_1,\boldsymbol{y})^\alpha q(\boldsymbol{x}_2,\boldsymbol{y})^\alpha\mu^3(\mathrm{d}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})) \\
&\le \int_{\mathbb{X}^3} G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)(\mathfrak{I}_1+2\mathfrak{I}_2+\mathfrak{I}_3)\mu^3(\mathrm{d}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}))\,, \qquad (4.3.17)
\end{aligned}
$$

with

$$\mathfrak{I}_1 = \mathfrak{I}_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) := \exp\Big\{ - \alpha\nu_d\big[\Lambda(r_0(\boldsymbol{x}_1, \boldsymbol{y})) + \Lambda(r_0(\boldsymbol{x}_2, \boldsymbol{y}))\big]\Big\},$$

$$\mathfrak{I}_2 = \mathfrak{I}_2(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) := e^{-\alpha\nu_d\Lambda(r_0(\boldsymbol{x}_1,\boldsymbol{y}))}\int_{s\vee t_2}^{\infty} \lambda(B_{u_2(r-t_2)}(0)\cap B_{v(r-s)}(y-x_2))e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r),$$

$$\mathfrak{I}_3 = \mathfrak{I}_3(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}) := \int_{s\vee t_2}^{\infty} \lambda(B_{u_2(r-t_2)}(0)\cap B_{v(r-s)}(y-x_2))e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r) \times \int_{s\vee t_1}^{\infty} \lambda(B_{u_1(\rho-t_1)}(0)\cap B_{v(\rho-s)}(y-x_1))e^{-\alpha\nu_d\Lambda(\rho)}\theta(\mathrm{d}\rho)\,.$$

By (4.3.16),

$$\begin{aligned}
&\iint_{\mathbb{R}^{2d}} \exp\left\{-\frac{\alpha\nu_d}{3}\left[\Lambda\left(\frac{|y|}{u_1+v}\right)+\Lambda\left(\frac{|x-y|}{u_2+v}\right)\right]\right\}\mathrm{d}x\mathrm{d}y\\
&\le \int_{\mathbb{R}^d} \exp\left\{-\frac{\alpha\nu_d}{3}\Lambda\left(\frac{|y|}{u_1+v}\right)\right\}\mathrm{d}y \int_{\mathbb{R}^d} \exp\left\{-\frac{\alpha\nu_d}{3}\Lambda\left(\frac{|x|}{u_2+v}\right)\right\}\mathrm{d}x\\
&\le d^2\omega_d^2(u_1+v)^d(u_2+v)^dQ_\alpha'^2.
\end{aligned}$$

Hence, using (4.3.8) and (4.3.15) for the first step and the inequality (4.2.1) for the second one, we have

$$\begin{aligned}
&\int_{\mathbb{X}^3} G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)\mathfrak{I}_1\ \mu^3(\mathrm{d}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}))\\
&\le \omega_d^{10}\iint_{\mathbb{R}^{2d}} \widetilde{M}(x_1)\widetilde{M}(x_2)\mathrm{d}x_1\mathrm{d}x_2\int_{\mathbb{R}_+^3} e^{-\frac{\alpha\nu_d}{3}[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)]}\tilde{h}(t_1)\tilde{h}(t_2)\theta^3(\mathrm{d}(t_1,t_2,s))\\
&\qquad\times\iint_{\mathbb{R}_+^3\times\mathbb{R}^d}(1+Qu_1^{5d})(1+Qu_2^{5d})e^{-\frac{\alpha\nu_d}{3}\left[\Lambda\left(\frac{|x_1-y|}{u_1+v}\right)+\Lambda\left(\frac{|x_2-y|}{u_2+v}\right)\right]}\mathrm{d}y\,\nu^3(\mathrm{d}(u_1,u_2,v))\\
&\le d^2\omega_d^{12}\,Q_\alpha'^2Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)^2\widetilde{M}_2\\
&\qquad\times\int_{\mathbb{R}_+^3}(u_1+v)^d(u_2+v)^d(1+Qu_1^{5d})(1+Qu_2^{5d})\nu^3(\mathrm{d}(u_1,u_2,v))\\
&\le c_1\,Q_\alpha'^2Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)^2\left[\nu_{2d}(1+Q\nu_{5d})^2+(\nu_d+Q\nu_{6d})^2\right]\widetilde{M}_2
\end{aligned}$$

for some constant $c_1$ depending only on $d$, where we have used monotonicity of $Q_0$ with respect to its second argument in the penultimate step and Jensen's inequality along with positive association for the final step.

Next, we bound the second summand in (4.3.17). Using (4.2.2) in the second step, monotonicity of $\Lambda$ and

(4.3.15) in the third step and (4.3.16) in the final one, we have

$$
\begin{aligned}
&\iint_{\mathbb{R}^{2d}} \mathfrak{I}_2(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\mathrm{d}x_2\mathrm{d}y \\
&= \int_{\mathbb{R}^d} e^{-\alpha\nu_d\Lambda(r_0(\boldsymbol{x}_1,\boldsymbol{y}))}\mathrm{d}y \int_{s\vee t_2}^{\infty}\int_{\mathbb{R}^d} \lambda(B_{u_2(r-t_2)}(0)\cap B_{v(r-s)}(y-x_2))\mathrm{d}x_2\, e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r) \\
&= \omega_d^2 u_2^d v^d \int_{\mathbb{R}^d} e^{-\alpha\nu_d\Lambda(r_0(\boldsymbol{x}_1,\boldsymbol{y}))}\mathrm{d}y \int_{s\vee t_2}^{\infty} (r-t_2)^d (r-s)^d\, e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r) \\
&\le \omega_d^2 u_2^d v^d \exp\Big\{-\frac{\alpha\nu_d}{3}\left[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)\right]\Big\} \\
&\qquad\qquad \times \int_0^{\infty} r^{2d}\, e^{-\alpha\nu_d\Lambda(r)/3}\theta(\mathrm{d}r)\int_{\mathbb{R}^d}\exp\Big\{-\frac{\alpha\nu_d}{3}\Lambda\left(\frac{|x_1-y|}{u_1+v}\right)\Big\}\,\mathrm{d}y \\
&= d\omega_d^3\, Q'_\alpha Q_0(2d,\alpha\nu_d/3)u_2^d v^d (u_1+v)^d \exp\Big\{-\frac{\alpha\nu_d}{3}\left[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)\right]\Big\}\,.
\end{aligned}
$$

Therefore, arguing similarly as before, we obtain

$$
\begin{aligned}
&\int_{\mathbb{X}^3} G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)\mathfrak{I}_2(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\ \mu^3(\mathrm{d}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})) \\
&\le \omega_d^{10}\int_{\mathbb{R}^d}\widetilde{M}(x_1)^2\mathrm{d}x_1 \iint_{\mathbb{R}_+^6}(1+Qu_1^{5d})(1+Qu_2^{5d}) \\
&\qquad\qquad \times\left(\iint_{\mathbb{R}^{2d}}\mathfrak{I}_2(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\mathrm{d}x_2\,\mathrm{d}y\right)\tilde{h}(t_1)\tilde{h}(t_2)\theta^3(\mathrm{d}(t_1,t_2,s))\nu^3(\mathrm{d}(u_1,u_2,s)) \\
&\le d\omega_d^{13}\,Q'_\alpha Q_0(2d,\alpha\nu_d/3)\widetilde{M}_2\int_{\mathbb{R}_+^3} e^{-\frac{\alpha\nu_d}{3}\left[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)\right]}\tilde{h}(t_1)\tilde{h}(t_2)\theta^3(\mathrm{d}(t_1,t_2,s)) \\
&\qquad\qquad \times\int_{\mathbb{R}_+^3} v^d(u_1+v)^d(1+Qu_1^{5d})(u_2^d+Qu_2^{6d})\nu^3(\mathrm{d}(u_1,u_2,v)) \\
&\le c_2\, Q'_\alpha Q_0(2d,\alpha\nu_d/3)Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)^2\left[\nu_d(\nu_d+Q\nu_{6d})^2+\nu_{2d}(\nu_d+Q\nu_{6d})(1+Q\nu_{5d})\right]\widetilde{M}_2
\end{aligned}
$$

for some constant $c_2$ depending only on $d$. Finally, we bound the third summand in (4.3.17). Arguing as above,

$$
\begin{aligned}
&\iint_{\mathbb{R}^{2d}}\mathfrak{I}_3(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\,\mathrm{d}x_2\,\mathrm{d}y \\
&= \int_{s\vee t_1}^{\infty}\left(\int_{\mathbb{R}^d}\lambda(B_{u_1(\rho-t_1)}(0)\cap B_{v(\rho-s)}(y-x_1))\mathrm{d}y\right)e^{-\alpha\nu_d\Lambda(\rho)}\theta(\mathrm{d}\rho) \\
&\qquad\qquad \times\int_{s\vee t_2}^{\infty}\left(\int_{\mathbb{R}^d}\lambda(B_{u_2(r-t_2)}(0)\cap B_{v(r-s)}(y-x_2))\mathrm{d}x_2\right)e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r) \\
&= \omega_d^4 u_1^d u_2^d v^{2d}\int_{s\vee t_1}^{\infty}(\rho-t_1)^d(\rho-s)^d e^{-\alpha\nu_d\Lambda(\rho)}\theta(\mathrm{d}\rho)\int_{s\vee t_2}^{\infty}(r-t_1)^d(r-s)^d e^{-\alpha\nu_d\Lambda(r)}\theta(\mathrm{d}r) \\
&\le \omega_d^4 u_1^d u_2^d v^{2d}\left(\int_0^{\infty} r^{2d}\, e^{-\alpha\nu_d\Lambda(r)/3}\theta(\mathrm{d}r)\right)^2\exp\Big\{-\frac{\alpha}{3}\nu_d\big[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)\big]\Big\} \\
&\le \omega_d^4\, Q_0(2d,\alpha\nu_d/3)^2 u_1^d u_2^d v^{2d}\exp\Big\{-\frac{\alpha}{3}\nu_d\big[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)\big]\Big\}\,.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\int_{\mathbb{X}^3} G(\boldsymbol{x}_1)G(\boldsymbol{x}_2)\mathfrak{I}_3(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\ \mu^3(\mathrm{d}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}))\\
&\le \omega_d^{10}\int_{\mathbb{R}^d}\widetilde{M}(x_1)^2\mathrm{d}x_1\iint_{\mathbb{R}^6_+}(1+Qu_1^{5d})(1+Qu_2^{5d})\\
&\qquad\qquad \times\left(\iint_{\mathbb{R}^{2d}}\mathfrak{I}_3(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\mathrm{d}x_2\,\mathrm{d}y\right)\tilde{h}(t_1)\tilde{h}(t_2)\theta^3(\mathrm{d}(t_1,t_2,s))\nu^3(\mathrm{d}(u_1,u_2,v))\\
&\le \omega_d^{14}\,Q_0(2d,\alpha\nu_d/3)^2\widetilde{M}_2\int_{\mathbb{R}^3_+}(1+Qu_1^{5d})(1+Qu_2^{5d})u_1^du_2^dv^{2d}\nu^3(\mathrm{d}(u_1,u_2,v))\\
&\qquad\qquad \times\int_{\mathbb{R}^3_+}e^{-\frac{\alpha\nu_d}{3}[\Lambda(t_1)+\Lambda(t_2)+2\Lambda(s)]}\tilde{h}(t_1)\tilde{h}(t_2)\theta^3(\mathrm{d}(t_1,t_2,s))\\
&\le \omega_d^{14}\,Q_0(2d,\alpha\nu_d/3)^2Q_0(0,\alpha\nu_d/3)Q_1(0,\alpha\nu_d/3)^2\nu_{2d}(\nu_d+Q\nu_{6d})^2\widetilde{M}_2\,.
\end{aligned}
$$

Combining the bound for the summands on the right-hand side of (4.3.17) and noticing that, by (4.2.1),

$$
\begin{aligned}
&Q'_\alpha Q_0(2d,\alpha\nu_d/3)\Big[\nu_d(\nu_d+Q\nu_{6d})^2+\nu_{2d}(\nu_d+Q\nu_{6d})(1+Q\nu_{5d})\Big]\\
&\le Q'_\alpha Q_0(2d,\alpha\nu_d/3)\Big[\sqrt{\nu_{2d}}(\nu_d+Q\nu_{6d})^2+\nu_{2d}(\nu_d+Q\nu_{6d})(1+Q\nu_{5d})\Big]\\
&\le Q'^2_\alpha\Big[\nu_{2d}(1+Q\nu_{5d})^2+(\nu_d+Q\nu_{6d})^2\Big]+Q_0(2d,\alpha\nu_d/3)^2\nu_{2d}(\nu_d+Q\nu_{6d})^2
\end{aligned}
$$

yields the desired conclusion. □

To compute the bounds in (4.3.5) and (4.3.6), we now only need to bound $\mu((\kappa+g)^{2\beta}G)$.

**Lemma 4.3.5.** *For* $\alpha\in(0,1]$,

$$\mu((\kappa+g)^\alpha G)\le\omega_d^{5+\alpha}(C_1+C_2+C_3)\widetilde{M}_1\,,$$

*where*

$$
\begin{aligned}
C_1 &:= Q_1(0,\alpha\nu_d)(1+Q\nu_{5d})\,,\\
C_2 &:= \nu_{\alpha d}Q_1(0,\alpha\zeta\nu_d/2)Q_0(d,\zeta\nu_d/2)^\alpha\,,\\
C_3 &:= \nu_{(5+\alpha)d}Q_1(0,(5+\alpha)\zeta\nu_d/2)Q_0(d,\zeta\nu_d/2)^{5+\alpha}\,.
\end{aligned}
$$

*Proof.* Define the function

$$\psi(t):=\int_t^\infty(s-t)^de^{-\zeta\nu_d\Lambda(s)}\theta(\mathrm{d}s)\,,$$

so that $g(\boldsymbol{x})=\omega_dv_x^d\psi(t_x)$ and

$$G(\boldsymbol{x})\le\omega_d^5\widetilde{M}(x)\tilde{h}(t_x)\big(1+v_x^{5d}\psi(t_x)^5\big)\,.$$

By subadditivity, it suffices to separately bound

$$\int_{\mathbb{X}} \kappa^{\alpha}(\boldsymbol{x}) G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}) \quad \text{and} \quad \int_{\mathbb{X}} g(\boldsymbol{x})^{\alpha} G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}) .$$

By (4.3.8),

$$\begin{aligned} \int_{\mathbb{X}} \kappa^{\alpha}(\boldsymbol{x}) G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}) &\leq \omega_d^5 \int_{\mathbb{X}} \widetilde{M}(x)(1+Q v_x^{5d}) e^{-\alpha \nu_d \Lambda(t_x)} \tilde{h}(t_x)\, \mathrm{d}x\, \theta(\mathrm{d}t_x)\, \nu(\mathrm{d}v_x) \\ &= \omega_d^5 Q_1(0, \alpha\nu_d)(1+Q\nu_{5d}) \widetilde{M}_1 = \omega_d^5 C_1 \widetilde{M}_1 . \end{aligned}$$

For the second integral, write

$$\begin{aligned} \int g(\boldsymbol{x})^{\alpha} G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}) &\leq \omega_d^{5+\alpha} \int_{\mathbb{R}^d} \widetilde{M}(x) \mathrm{d}x \int_0^{\infty} \int_0^{\infty} \psi(t_x)^{\alpha} \tilde{h}(t_x) v_x^{\alpha d}(1+\psi(t_x)^5 v_x^{5d}) \nu(\mathrm{d}v_x) \theta(\mathrm{d}t_x) \\ &= \omega_d^{5+\alpha} \widetilde{M}_1 \left[ \nu_{\alpha d} \int_0^{\infty} \psi(t)^{\alpha} \tilde{h}(t) \theta(\mathrm{d}t) + \nu_{(5+\alpha)d} \int_0^{\infty} \psi(t)^{5+\alpha} \tilde{h}(t) \theta(\mathrm{d}t) \right] . \end{aligned}$$

Note that, for any $a > 0$,

$$\begin{aligned} \int_0^{\infty} \psi(t)^a \tilde{h}(t) \theta(\mathrm{d}t) &= \int_0^{\infty} \left( \int_t^{\infty} (s-t)^d e^{-\zeta \nu_d \Lambda(s)} \theta(\mathrm{d}s) \right)^a \tilde{h}(t) \theta(\mathrm{d}t) \\ &\leq \int_0^{\infty} e^{-a\zeta\nu_d \Lambda(t)/2} \tilde{h}(t) \theta(\mathrm{d}t) \left( \int_0^{\infty} s^d e^{-\zeta\nu_d\Lambda(s)/2} \theta(\mathrm{d}s) \right)^a \\ &= Q_1(0, a\zeta\nu_d/2) Q_0(d, \zeta\nu_d/2)^a , \end{aligned}$$

where we have used the monotonicity of $\Lambda$ in the second step. Hence,

$$\int g(\boldsymbol{x})^{\alpha} G(\boldsymbol{x}) \mu(\mathrm{d}\boldsymbol{x}) \leq \omega_d^{5+\alpha}(C_2 + C_3) \widetilde{M}_1 .$$

Combining with the above bound yields the result. □

*Proofs of Theorems 4.1.1 and 4.1.3:* Theorem 4.1.1 follows from (4.3.5) and (4.3.6) upon using Lemmas 4.3.2, 4.3.3, 4.3.4 and 4.3.5 and including the factors involving the moments of the speed into the constants.

The assertion in Theorem 4.1.3 follows by combining Theorem 4.1.1 and Proposition 4.1.2. □

*Proof of Theorem 4.1.4.* Let $\theta$ be given at (4.1.6), and as in the proof of Proposition 4.1.2 let $\Lambda(t) = B\,\omega_d t^{d+\tau+1}$, where $B := B(d+1, \tau+1)$. By (4.3.7), for $x \in \mathbb{R}_+$ and $y > 0$,

$$Q_0(x,y) = \int_0^{\infty} t^{x+\tau} e^{-y\omega_d B\, t^{d+\tau+1}} \mathrm{d}t = \frac{(y\omega_d B)^{-\frac{x+\tau+1}{d+\tau+1}}}{d+\tau+1} \Gamma\left( \frac{x+\tau+1}{d+\tau+1} \right) = C_1(x,\tau) y^{-\frac{x+\tau+1}{d+\tau+1}} , \tag{4.3.18}$$

where

$$C_1(x,\tau) := \frac{(\omega_d B)^{-\frac{x+\tau+1}{d+\tau+1}}}{d+\tau+1} \Gamma\left( \frac{x+\tau+1}{d+\tau+1} \right) .$$

Then

$$Q = Q_0(d, \zeta\nu_d)^5 = C_1(d,\tau)\zeta^{-5}\nu_d^{-5}.$$

Since $h_2(t) = \mathbf{1}\{t < a\}$, a change of variables and (4.2.4) yield that

$$Q_1(0,y) = \int_0^a t^\tau e^{-y\omega_d B\, t^{d+\tau+1}} \mathrm{d}t \le C \frac{(y\omega_d B)^{-\frac{\tau+1}{d+\tau+1}}}{d+\tau+1} \gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1} y\right) \le C' l_{a,\tau}(y)\,, \tag{4.3.19}$$

where $C$ and $C'$ are constants depending only on $d$ and $\tau$, and $l_{a,\tau}(y)$ is defined at (4.1.7). Similarly, by (4.3.12),

$$Q'_\alpha := \frac{1}{d+\tau+1}\Gamma\left(\frac{d}{d+\tau+1}\right)(\alpha B\omega_d\nu_d/3)^{-\frac{d}{d+\tau+1}} = C_3(\alpha,\tau)\nu_d^{-\frac{d}{d+\tau+1}}\,,$$

where

$$C_3(\alpha,\tau) := \frac{1}{d+\tau+1}\Gamma\left(\frac{d}{d+\tau+1}\right)\left(\frac{\alpha B\omega_d}{3}\right)^{-\frac{d}{d+\tau+1}}.$$

By (4.3.18), for $b > 0$,

$$Q_0(x, by) = b^{-\frac{x+\tau+1}{d+\tau+1}} Q_0(x,y),$$

while from (4.3.19) and (4.2.4), we have

$$Q_1(0, by) \le \max\{1, b^{-\frac{\tau+1}{d+\tau+1}}\} Q_1(0,y)\,.$$

Recall the parameters $p = 1$, $\beta = 1/36$ and $\zeta = 1/50$. Lemmas 4.3.2–4.3.4 in combination with the above estimates along with the inequality $\nu_\delta\nu_{6d-\delta} \le \nu_{6d}$ for any $0 < \delta < 6d$ yield that there exists a constant $C$ depending only on $d$ and $\tau$ such that

$$\int_{\mathbb{X}} f_{2\beta}^{(1)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C l_{a,\tau}(\nu_d)\left(1 + \nu_{5d}\nu_d^{-5}\right)\widetilde{M}_1 \le C l_{a,\tau}(\nu_d)\left(1 + \nu_{6d}\nu_d^{-6}\right)\widetilde{M}_1\,,$$

$$\int_{\mathbb{X}} f_{2\beta}^{(2)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C l_{a,\tau}(\nu_d)\nu_d^{-1}(\nu_d + \nu_{6d}\nu_d^{-5})\widetilde{M}_1 = l_{a,\tau}(\nu_d)\left(1 + \nu_{6d}\nu_d^{-6}\right)\widetilde{M}_1\,,$$

and

$$\int_{\mathbb{X}} f_{2\beta}^{(3)}(\boldsymbol{y})\mu(\mathrm{d}\boldsymbol{y}) \le C l_{a,\tau}(\nu_d)\nu_d^{-\frac{\tau+1}{d+\tau+1}}\nu_d^{-\frac{d}{d+\tau+1}}\left(\nu_d + \nu_{6d}\nu_d^{-5}\right)\widetilde{M}_1 = C l_{a,\tau}(\nu_d)\left(1 + \nu_{6d}\nu_d^{-6}\right)\widetilde{M}_1\,.$$

Next, referring to (4.3.10) in the proof of Lemma 4.3.2, since $\tilde{h}(t) \le \tilde{h}(s)$ for $s \le t$, we can upper bound

$$\int_0^\infty \left(\int_s^\infty (t-s)^d w(t)^\alpha \tilde{h}(t)\theta(\mathrm{d}t)\right)^2 \theta(\mathrm{d}s) \le \int_0^\infty \left(\int_0^\infty t^d w(t)^{\alpha/2}\theta(\mathrm{d}t)\right)^2 \tilde{h}(s) w(s)^\alpha \theta(\mathrm{d}s)$$

yielding an alternative upper bound

$$\int_{\mathbb{X}} f_\alpha^{(1)}(\boldsymbol{y})^2\mu(\mathrm{d}\boldsymbol{y}) \le C_2\widetilde{M}_2,$$

where

$$C_2 := \omega_d^{12} Q_0(d, \alpha\nu_d/2)^2 Q_1(0, \alpha\nu_d)(1 + Q\nu_{5d})^2 \nu_{2d}.$$

Thus, there exists a constant $C$ depending only on $d$ and $\tau$ such that

$$\int_{\mathbb{X}} f_\beta^{(1)}(\boldsymbol{y})^2 \mu(\mathrm{d}\boldsymbol{y}) \le C l_{a,\tau}(\nu_d) \nu_{2d} \nu_d^{-2} \left(1 + \nu_{5d}\nu_d^{-5}\right)^2 \widetilde{M}_2 \le C l_{a,\tau}(\nu_d) \nu_{6d} \nu_d^{-6} \left(1 + \nu_{6d}\nu_d^{-6}\right)^2 \widetilde{M}_2 \,,$$

Similarly, using that $Q_1(0, y) \le Q_0(0, y)$, Lemmas 4.3.3–4.3.5 yield

$$\int_{\mathbb{X}} f_\beta^{(2)}(\boldsymbol{y})^2 \mu(\mathrm{d}\boldsymbol{y}) \le C l_{a,\tau}(\nu_d) \left(1 + \nu_{6d}\nu_d^{-6}\right)^2 \widetilde{M}_2 \,,$$

$$\begin{aligned}\int_{\mathbb{X}} f_\beta^{(3)}(\boldsymbol{y})^2 \mu(\mathrm{d}\boldsymbol{y}) &\le C l_{a,\tau}(\nu_d) \nu_d^{-\frac{2(\tau+1)}{d+\tau+1}} \widetilde{M}_2 \left[\nu_d^{-\frac{2d}{d+\tau+1}} \nu_{2d}(1 + \nu_{6d}\nu_d^{-6})^2 + \nu_{2d}\nu_d^2 \nu_d^{-\frac{2(2d+\tau+1)}{d+\tau+1}} (1 + \nu_{6d}\nu_d^{-6})^2\right] \\ &\le C l_{a,\tau}(\nu_d) \nu_{6d} \nu_d^{-6} (1 + \nu_{6d}\nu_d^{-6})^2 \widetilde{M}_2\end{aligned}$$

and

$$\begin{aligned}\mu((\kappa + g)^{2\beta} G) &\le C l_{a,\tau}(\nu_d) \left(\left(1 + \nu_{5d}\nu_d^{-5}\right) + \nu_{2\beta d}\nu_d^{-2\beta} + \nu_{(5+2\beta)d}\nu_d^{-5-2\beta}\right) \widetilde{M}_1 \\ &\le C l_{a,\tau}(\nu_d)(1 + \nu_{6d}\nu_d^{-6}) \widetilde{M}_1 \,,\end{aligned}$$

where $C$ is a constant depending only on $d$ and $\tau$ that may vary from line to line. Plugging the above estimates in (4.3.5) and (4.3.6) and using Proposition 4.1.2 to lower bound the variance yield the desired bounds. □

*Proof of Corollary 4.1.5:* Define the Poisson process $\eta^{(s)}$ with intensity measure $\mu^{(s)} := \lambda \otimes \theta \otimes \nu^{(s)}$, where $\nu^{(s)}(A) := \nu(s^{-1/d}A)$ for all Borel sets $A$. It is straightforward to see that the set of locations of exposed points of $\eta_s$ has the same distribution as of those of $\eta^{(s)}$, multiplied by $s^{-1/d}$, i.e., the set $\{x : \boldsymbol{x} \in \eta_s \text{ is exposed}\}$ coincides in distribution with $\{s^{-1/d}x : \boldsymbol{x} \in \eta^{(s)} \text{ is exposed}\}$. Hence, the functional $F^{(a)}(\eta_s)$ has the same distribution as $F_s^{(a)}(\eta^{(s)})$, where $F_s^{(a)}$ is defined as in (4.1.2) for the weight function

$$h(\boldsymbol{x}) = h_{1,s}(x) h_2(t_x) = h_1(s^{-1/d}x) \mathbf{1}\{t_x < a\} \,.$$

It is easy to check that for $k \in \mathbb{N}$, the $k$-th moment of $\nu^{(s)}$ is given by $\nu_k^{(s)} = s^{k/d}\nu_k$, the quantities constructed from $h_{1,s}$ become $\widetilde{M}_i^{(s)} = s\widetilde{M}_i$, $i = 1, 2$, and $h_{1,s}^{(2)} = s h_1^{(2)}$. Finally noticing that, for $s \ge 1$,

$$l_{a,\tau}(\nu_d^{(s)}) = \gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1} s\nu_d\right) (s\nu_d)^{-\frac{\tau+1}{d+\tau+1}} \ge \gamma\left(\frac{\tau+1}{d+\tau+1}, a^{d+\tau+1}\nu_d\right) (s\nu_d)^{-\frac{\tau+1}{d+\tau+1}},$$

the result follows directly from Theorem 4.1.4. □

# Part II

# Set-valued Sublinear Expectations

# Chapter 5

# Convex Bodies Generated by Sublinear Expectations of Random Vectors

This chapter is based on the following article:

*I. Molchanov and R. Turin. Convex bodies generated by sublinear expectations of random vectors. Adv. in Appl. Math., 131:Paper No. 102251, 31, 2021.*

## 5.1 Sublinear expectations of random variables

In this first section, we give the definition of sublinear expectation of random variables, explaining their dual representation and presenting several examples. We mention the particularly important Kusuoka representation which expresses any law-determined sublinear expectation in terms of integrated quantiles and describe a novel construction (called the maximum extension) suitable to produce parametric families of sublinear expectations from each given one.

### 5.1.1 Definition and dual representation

Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a nonatomic probability space, and let $L^p(\mathbb{R}^d)$ denote the family of all $p$-integrable random vectors in $\mathbb{R}^d$, with $p \in [1, \infty]$. Endow $L^p(\mathbb{R}^d)$ with the $\sigma(L^p, L^q)$-topology, which is the weak-star topology based on the pairing of $L^p(\mathbb{R}^d)$ and $L^q(\mathbb{R}^d)$ with $p^{-1} + q^{-1} = 1$, see [3, Sec. 5.14]. Denote $\mathbb{R}_+ = [0, \infty)$.

The following definition amends the standard definition of sublinear expectations of random variables (see, e.g., [79]) by including the extra lower semicontinuity property, which is often additionally imposed.

**Definition 5.1.1.** A *sublinear expectation* is a function $\mathsf{e} : L^p(\mathbb{R}) \to (-\infty, \infty]$ with $p \in [1, \infty]$, satisfying the following properties for all $\beta, \beta' \in L^p(\mathbb{R})$:

i) monotonicity: $\mathsf{e}(\beta) \leq \mathsf{e}(\beta')$ if $\beta \leq \beta'$ a.s.;

ii) translation equivariance: $\mathsf{e}(\beta + a) = \mathsf{e}(\beta) + a$ for all $a \in \mathbb{R}$, and $\mathsf{e}(0) = 0$;

iii) positive homogeneity: $\mathsf{e}(c\beta) = c\mathsf{e}(\beta)$ for all $c > 0$;

iv) subadditivity: $\mathsf{e}(\beta + \beta') \leq \mathsf{e}(\beta) + \mathsf{e}(\beta')$,

v) lower semicontinuity in $\sigma(L^p, L^q)$, that is,

$$\mathsf{e}(\beta) \leq \liminf_{n\to\infty} \mathsf{e}(\beta_n)$$

for each sequence $\{\beta_n, n \geq 1\}$ converging to $\beta$ in the weak-star topology $\sigma(L^p, L^q)$.

The sublinear expectation $\mathsf{e}$ is often referred to as *numerical* one, in contrast with the set-valued expectation introduced in Section 5.2.1. The translation equivariance property implies that $\mathsf{e}(a) = a$ for each deterministic $a$. The sublinear expectation $\mathsf{e}$ is said to be *finite* if it takes a finite value on all $\beta \in L^p(\mathbb{R})$.

*Example* 5.1.2 (Relation to coherent risk measures). For $\beta \in L^p(\mathbb{R})$, define $\mathsf{r}(\beta) = \mathsf{e}(-\beta)$. The obtained antimonotonic and subadditive function is called a *coherent risk measure* of $\beta$, see [33] and [39, Def. 4.5]. The negative of the risk is said to be a utility function, see [33].

A random variable $\beta$ is said to be acceptable if its risk is at most zero. If $\beta$ is the financial position at the terminal time, its risk $\mathsf{r}(\beta)$ yields the smallest amount $a$ of capital which should be reserved at the initial time to render $\beta + a$ acceptable; this amount may be negative if $\mathsf{r}(\beta) < 0$, and then capital can be released or invested. The subadditivity property of the risk (equivalently, of $\mathsf{e}$) is the manifestation of the financial principle, saying that diversification decreases the risk. Many results from the theory of risk measures can be easily reformulated for sublinear expectations. For instance, from the theory of risk measures, it is known that the lower semicontinuity property always holds if $p \in [1, \infty)$ and $\mathsf{e}$ takes only finite values, see [50].

While the following result is well known for risk measures [39, Cor. 4.18] and sublinear expectations [79, Th. 1.2.1], we provide its proof for completeness.

**Theorem 5.1.3.** *A functional* $\mathsf{e} : L^p(\mathbb{R}) \to (-\infty, \infty]$ *is a sublinear expectation if and only if*

$$\mathsf{e}(\beta) = \sup_{\gamma \in \mathcal{M}_{\mathsf{e}}, \mathbf{E}\gamma = 1} \mathbf{E}(\gamma\beta), \tag{5.1.1}$$

*where* $\mathcal{M}_{\mathsf{e}}$ *is a convex* $\sigma(L^q, L^p)$*-closed cone in* $L^q(\mathbb{R}_+)$*.*

*Proof. Sufficiency* is easy to confirm by a direct check of the properties.

*Necessity.* Let $\mathcal{A}$ be the family of $\beta \in L^p(\mathbb{R})$, such that $\mathsf{e}(\beta) \le 0$. The sublinearity property yields that $\mathcal{A}$ is a convex cone. The lower semicontinuity property implies that this cone is weak-star closed. The polar cone to $\mathcal{A}$ is defined as

$$\mathcal{A}^o = \{\gamma \in L^q(\mathbb{R}) : \mathbf{E}(\gamma\beta) \le 0 \text{ for all } \ \beta \in \mathcal{A}\}. \tag{5.1.2}$$

Since $-\mathbf{1}_A \in \mathcal{A}$ for the indicator of any event $A$, all random variables from $\mathcal{A}^o$ are a.s. non-negative. The bipolar theorem from functional analysis (see, e.g., [3, Th. 5.103]) yields that $(\mathcal{A}^o)^o = \mathcal{A}$. Hence,

$$\begin{aligned} \mathsf{e}(\beta) &= \inf\{a \in \mathbb{R} : (\beta - a) \in \mathcal{A}\} \\ &= \inf\{a \in \mathbb{R} : \mathbf{E}((\beta - a)\gamma) \le 0 \text{ for all } \ \gamma \in \mathcal{A}^o\} \\ &= \inf\{a \in \mathbb{R} : \mathbf{E}(\gamma\beta) \le a\mathbf{E}(\gamma) \text{ for all } \ \gamma \in \mathcal{A}^o\}. \end{aligned}$$

Thus, (5.1.1) holds with $\mathcal{M}_{\mathsf{e}} = \mathcal{A}^o$. □

Representation (5.1.1) is called the *dual* representation of $\mathsf{e}$. It is easy to see that each $\gamma$ in (5.1.1) can be chosen to be a function of $\beta$, namely, the conditional expectations $\mathbf{E}(\gamma|\beta)$.

A sublinear expectation is said to be *law-determined* (often named law invariant) if it attains the same value on identically distributed random variables, and this is the case for all examples considered in this work. In terms of the representation (5.1.1), this means that, for each $\gamma \in \mathcal{M}_{\mathsf{e}}$, the set $\mathcal{M}_{\mathsf{e}}$ contains all random variables sharing the same distribution with $\gamma$.

A sublinear expectation is said to be *continuous from below* if it is continuous on all almost surely convergent increasing sequences of random variables in $L^p(\mathbb{R})$. It follows from [50] that each finite sublinear expectation on $L^p(\mathbb{R})$ with $p \in [1, \infty)$ is continuous from below. Every law-determined continuous from below sublinear expectation on a nonatomic probability space is *dilatation monotonic*, meaning that

$$\mathsf{e}(\mathbf{E}(\beta|\mathfrak{A})) \le \mathsf{e}(\beta) \tag{5.1.3}$$

for each sub-$\sigma$-algebra $\mathfrak{A}$ of $\mathfrak{F}$, see [39, Cor. 4.59]. In particular, $\mathbf{E}\beta \le \mathsf{e}(\beta)$ for all $\beta \in L^p(\mathbb{R})$.

### 5.1.2 Average quantiles and the Kusuoka representation

For a fixed value of $\alpha \in (0,1]$ and $\beta \in L^1(\mathbb{R})$, define

$$\mathsf{e}_\alpha(\beta) = \frac{1}{\alpha}\int_{1-\alpha}^{1} q_t(\beta)\mathrm{d}t, \tag{5.1.4}$$

where

$$q_t(\beta) = \sup\{s \in \mathbb{R} : \mathbb{P}\{\beta \le s\} < t\} = \inf\{s \in \mathbb{R} : \mathbb{P}\{\beta \le s\} \ge t\} \tag{5.1.5}$$

is the $t$-quantile of $\beta$. Because of integration, the choice of a particular quantile in case of multiplicities is immaterial. This sublinear expectation is subsequently called the *average quantile*. In particular, $\mathsf{e}_1(\beta) = \mathbf{E}\beta$ is the mean. If $\beta$ has a nonatomic distribution, then $\mathsf{e}_\alpha(\beta) = \mathbf{E}(\beta|\beta \ge q_{1-\alpha}(\beta))$.

The value of $\mathsf{r}(\beta) = \mathsf{e}_\alpha(-\beta)$ is obtained by averaging the quantiles of $\beta$ at levels between 0 and $\alpha$. This risk measure is well studied in finance and widely applied in practice under the name of the average Value-at-Risk or expected shortfall, see, e.g., [2]. By computing the dual cone at (5.1.2) or rephrasing the representation of the risk measure $\mathsf{e}_\alpha(-\beta)$ from [50, Th. 4.1], one can derive the following dual representation

$$\mathsf{e}_\alpha(\beta) = \sup_{\gamma \in L^\infty([0,\alpha^{-1}]), \mathbf{E}\gamma = 1} \mathbf{E}(\gamma\beta). \tag{5.1.6}$$

This immediately yields that the average quantiles satisfy all properties imposed in Definition 5.1.1.

Average quantiles form a building block for all other law-determined sublinear expectations. The following result for risk measures is known as the *Kusuoka representation*: it was first obtained by Kusuoka [57] in case $p = \infty$ and can also be found in [39, Cor. 4.58] and [33, Th. 32]; the $L^p$-variant follows from the Orlicz space version proved in [43]. For its validity, it is essential that the probability space is nonatomic.

**Theorem 5.1.4.** *Each law-determined sublinear expectation on $L^p(\mathbb{R})$ with $p \in [1,\infty]$ can be represented as*

$$\mathsf{e}(\beta) = \sup_{\nu \in \mathcal{P}_{\mathsf{e}}} \int_{(0,1]} \mathsf{e}_\alpha(\beta)\nu(\mathrm{d}\alpha), \tag{5.1.7}$$

*where $\mathcal{P}_{\mathsf{e}}$ is the family of probability measures $\nu$ on $(0,1]$ such that $\int_{(0,1]} \mathsf{e}_\alpha(\beta)\nu(\mathrm{d}\alpha) \le 0$ whenever $\mathsf{e}(\beta) \le 0$.*

It is possible to show that $\mathsf{e}$ is finite on $L^p(\mathbb{R})$ if and only if the function $t \mapsto \int_{(t,1]} s^{-1}\nu(\mathrm{d}s)$ is $q$-integrable on $(0,1]$ with respect to the Lebesgue measure for all $\nu \in \mathcal{P}_{\mathsf{e}}$. If $\mathsf{e}$ is finite and $p \in [1,\infty)$, one can provide a constructive representation of $\mathcal{P}_{\mathsf{e}}$ in terms of the extremal points of the set $\mathcal{M}_{\mathsf{e}}^1 = \{\gamma \in \mathcal{M}_{\mathsf{e}} : \mathbf{E}\gamma = 1\}$, where $\mathcal{M}_e$ is defined in (5.1.1). The case $p = \infty$ requires extra arguments, since a norm bounded set in $L^1$ is not necessarily weakly compact, hence, the supremum in (5.1.1) is not necessarily attained. Since $(\Omega, \mathcal{F}, \mathbb{P})$ is nonatomic, we can assume without loss of generality that $\Omega$ is the interval $[0,1]$ equipped with its Borel

$\sigma$-algebra and the Lebesgue measure $\mathbb{P}$. Let $\gamma : [0,1) \to [0,\infty)$ be a nondecreasing right-continuous function that is extremal in $\mathcal{M}^1_{\mathsf{e}}$. Define the probability measure $\nu_\gamma$ on $(0,1]$ by letting

$$\nu_\gamma((0,\alpha)) = \int_{[1-\alpha,1)} (\gamma(t) - \gamma(1-\alpha))\mathrm{d}t.$$

and $\nu(\{1\}) = \gamma(0)$. It is shown in [98] that $\mathcal{P}_{\mathsf{e}}$ can be chosen to be the set of $\nu_\gamma$ for the family of all right-continuous nondecreasing functions $\gamma$ which are extremal in $\mathcal{M}^1_{\mathsf{e}}$.

### 5.1.3 Examples of sublinear expectations

A simple example of a sublinear expectation is provided by the essential supremum

$$\mathsf{e}(\beta) = \operatorname{ess\,sup} \beta,$$

which is finite for all $\beta \in L^\infty(\mathbb{R})$. If $\alpha \downarrow 0$, then the average quantile $\mathsf{e}_\alpha(\beta)$ increases to the (possibly, infinite) value $\mathsf{e}_0(\beta)$, which is equal to the essential supremum of $\beta$. Next, we discuss more involved constructions of sublinear expectations.

*Example* 5.1.5 (Spectral sublinear expectation). Let $\varphi : (0,1] \to \mathbb{R}_+$ be a nonincreasing function such that $\int_0^1 \varphi(t)\mathrm{d}t = 1$, $\varphi$ is called a spectral function. Then

$$\mathsf{e}_{\int\varphi}(\beta) = \int_{(0,1]} q_{1-t}(\beta)\varphi(t)\mathrm{d}t \tag{5.1.8}$$

is called a spectral sublinear expectation, see [1] for the closely related definition of the spectral risk measure. By Fubini's theorem, $\mathsf{e}_{\int\varphi}$ admits the following equivalent representation

$$\mathsf{e}_{\int\varphi}(\beta) = \int_{(0,1]} \mathsf{e}_\alpha(\beta)\nu(\mathrm{d}\alpha), \tag{5.1.9}$$

where $\mathsf{e}_\alpha$ is given by (5.1.4) and $\nu$ is the probability measure on $(0,1]$ with

$$\varphi(t) = \int_{(t,1]} s^{-1}\nu(\mathrm{d}s), \quad t \in (0,1]. \tag{5.1.10}$$

Conversely, for any probability measure $\nu$ on $(0,1]$, (5.1.9) yields a spectral sublinear expectation. The set $\mathcal{P}_{\mathsf{e}}$ in the Kusuoka representation of $\mathsf{e}_{\int\varphi}(\beta)$ consists of the single probability measure $\nu$, so the right-hand side (5.1.7) is the supremum over a family of spectral sublinear expectations.

*Example* 5.1.6 (One-sided moments). The $L^p$-norm $\|\beta\|_p$ satisfies all properties of a sublinear expectation but the monotonicity and translation equivariance. It is possible to come up with a norm-based sublinear expectation on $L^p(\mathbb{R})$ with $p \in [1,\infty)$ by letting

$$\mathsf{e}_{p,a}(\beta) = \mathbf{E}\beta + a\big(\mathbf{E}(\beta - \mathbf{E}\beta)_+^p\big)^{1/p} \tag{5.1.11}$$

with $a \in [0,1]$, where $x_+ = \max(x,0)$ denotes the positive part of $x \in \mathbb{R}$. The corresponding risk measure was introduced in [36]. Note that $\mathsf{e}_{p,a}(\beta) = \frac{a}{2}\|\beta\|_p$ if $\beta$ is symmetric. Translation equivariance and positive homogeneity of $\mathsf{e}_{p,a}$ are obvious. The subadditivity of the second term follows from $(t+s)_+ \leq (t)_+ + (s)_+$ and the subadditivity of the $L^p$-norm. To prove the monotonicity, we first observe that since $\mathsf{e}_{p,a}$ is subadditive, we only need to show that $\mathsf{e}_{p,a}(\gamma) \leq 0$ for any almost surely negative integrable $\gamma$. Indeed, substituting $(\gamma - \mathbf{E}\gamma)_+ \leq -\mathbf{E}\gamma$ in (5.1.11) implies that $\mathsf{e}_{p,a}(\gamma) \leq \mathbf{E}\gamma - a\mathbf{E}\gamma \leq 0$.

The sublinear expectation given by (5.1.11) admits the dual representation (5.1.1) with the cone $\mathcal{M}_{\mathsf{e}}$ generated by the family of random variables $\gamma = 1 + a(\zeta - \mathbf{E}\zeta)$ for all $\zeta \in L^q(\mathbb{R}_+)$ with $\|\zeta\|_q \leq 1$, see [33, p. 46]. The family $\mathcal{P}_{\mathsf{e}}$ from (5.1.7) is explicitly known only for $p = 1$; it consists of probability measures obtained as $(1-at)\delta_1 + at\delta_t$, which is the weighted sum of the Dirac measures at 1 and $t$ for $t \in [0,1]$. Then

$$\mathsf{e}_{1,a}(\beta) = \sup_{t\in[0,1]} \Big[(1-at)\mathbf{E}\beta + at\mathsf{e}_t(\beta)\Big] = \mathbf{E}\beta + a \sup_{t\in[0,1]} t\mathsf{e}_t(\beta - \mathbf{E}\beta). \tag{5.1.12}$$

Recall in this relation that

$$t\mathsf{e}_t(\beta) = \int_{1-t}^{1} q_s(\beta)\mathrm{d}s,$$

so that the supremum on the right-hand side of (5.1.12) is indeed the expectation of $(\beta - \mathbf{E}\beta)_+$.

*Example* 5.1.7 (Expectile). Following [15], define the *expectile* $\mathsf{e}_{[\tau]}(\beta)$ of a random variable $\beta \in L^1(\mathbb{R})$ at level $\tau \in (0,1)$ as the (necessarily, unique) solution $x \in \mathbb{R}$ of

$$\tau\mathbf{E}(\beta - x)_+ = (1-\tau)\mathbf{E}(x-\beta)_+.$$

If $\tau \in [1/2, 1)$, then the expectile is a sublinear expectation, see [15]. For $\tau = 1/2$, we obtain the mean of $\beta$. For $\tau \in [1/2,1)$, the dual representation holds with $\mathcal{M}_{\mathsf{e}}$ being the set of $\gamma \in L^\infty(\mathbb{R}_+)$ such that the ratio between the essential supremum and the essential infimum of $\gamma$ is at most $\tau/(1-\tau)$. The Kusuoka representation holds with

$$\mathsf{e}_{[\tau]}(\beta) = \sup_{t\in[0,2-1/\tau]} \Big[(1-t)\mathbf{E}\beta + t\mathsf{e}_{\frac{(1-\tau)t}{(2\tau-1)(1-t)}}(\beta)\Big].$$

### 5.1.4 Maximum extension

Let $\mathsf{e}$ be a law-determined sublinear expectation on $L^p(\mathbb{R})$ with $p \in [1,\infty]$. The following construction suggests a way of extending $\mathsf{e}$ to a monotone parametric family of sublinear expectations. For a fixed $m \geq 1$, define

$$\mathsf{e}^{\vee m}(\beta) = \mathsf{e}(\max(\beta_1,\ldots,\beta_m)), \tag{5.1.13}$$

where $\beta_1, \dots, \beta_m$ are independent copies of $\beta \in L^p(\mathbb{R})$. All properties in Definition 5.1.1 are straightforward and we refer to this sublinear expectation as the *maximum extension* of $\mathsf{e}$. Let us stress that this extension applies only to law-determined sublinear expectations.

It is possible to obtain a family of such expectations $\mathsf{e}^{\vee(\lambda)}$ continuously parametrised by $\lambda \in (0,1]$. For this, $m$ is replaced by a geometrically distributed random variable $N$ with parameter $\lambda$, that is, $\mathbb{P}\{N = k\} = (1-\lambda)^{k-1}\lambda$, $k \geq 1$. Define

$$\mathsf{e}^{\vee(\lambda)} = \mathsf{e}(\max(\beta_1, \dots, \beta_N)), \quad \lambda \in (0,1].$$

This family of sublinear expectations interpolates between $\mathsf{e}^{\vee(1)}(\beta) = \mathsf{e}(\beta)$ and $\mathsf{e}^{\vee(0)}(\beta)$ which is set to be $\operatorname{ess\,sup} \beta$.

*Example* 5.1.8. The maximum extension can be applied to the average quantile risk measure $\mathsf{e}_\alpha$; the result is denoted by $\mathsf{e}_\alpha^{\vee m}$. For $\alpha = 1$, we obtain the *expected maximum*

$$\mathsf{e}_1^{\vee m}(\beta) = \mathbf{E}\max(\beta_1, \dots, \beta_m). \tag{5.1.14}$$

Note that

$$\mathsf{e}_1^{\vee m}(\beta) = \int_0^1 q_t(\max\{\beta_1, \dots, \beta_m\})\mathrm{d}t = \int_0^1 q_{t^{\frac{1}{m}}}(\beta)\mathrm{d}t = m\int_0^1 t^{m-1} q_t(\beta)\mathrm{d}t\,.$$

For $m \geq 2$, $\mathsf{e}_1^{\vee m}$ is the spectral sublinear expectation given at (5.1.8) with $\varphi(t) = m(1-t)^{m-1}$, equivalently, (5.1.9) with $\nu(\mathrm{d}t) = m(m-1)t(1-t)^{m-2}\mathrm{d}t$. Similar calculations yield that

$$\begin{aligned}\mathsf{e}_\alpha^{\vee m}(\beta) = \mathsf{e}_\alpha(\max(\beta_1, \dots, \beta_m)) = \frac{m(m-1)}{\alpha}\int_0^{1-(1-\alpha)^{1/m}} & t(1-t)^{m-2}\mathsf{e}_t(\beta)\mathrm{d}t \\ & + \frac{m}{\alpha}(1-\alpha)^{(m-1)/m}(1-(1-\alpha)^{1/m})\mathsf{e}_{1-(1-\alpha)^{1/m}}(\beta)\,. \end{aligned} \tag{5.1.15}$$

## 5.2 Set-valued maps

The present section presents our construction of convex closed sets $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ generated by a random vector $\xi$ and a given sublinear expectation $\mathsf{e}$. Subsection 5.2.2 describes a generalisation based on relaxing some properties of the underlying numerical sublinear expectations, namely, replacing them with gauge functions. This construction yields centroid bodies [67] and half-space depth-trimmed regions [73], the latter are closely related to convex floating bodies introduced in [97] and their weighted variant from [16].

### 5.2.1 Set-valued sublinear expectations

Fix a law-determined sublinear expectation $\mathsf{e}$ on $L^p(\mathbb{R})$, $p \in [1,\infty]$. For a $p$-integrable probability measure $\mu$ on $\mathbb{R}^d$, equivalently, for a random vector $\xi \in L^p(\mathbb{R}^d)$ with distribution $\mu$, define

$$h(u) = \mathsf{e}(\langle \xi, u\rangle), \quad u \in \mathbb{R}^d, \tag{5.2.1}$$

where $\langle \xi, u\rangle$ denotes the scalar product in $\mathbb{R}^d$. The function $h$ is subadditive

$$h(u+u') = \mathsf{e}(\langle \xi, u+u'\rangle) \le \mathsf{e}(\langle \xi, u\rangle) + \mathsf{e}(\langle \xi, u'\rangle) = h(u) + h(u'),$$

and homogeneous

$$h(cu) = \mathsf{e}(\langle \xi, cu\rangle) = c\mathsf{e}(\langle \xi, u\rangle) = ch(u), \quad c \ge 0.$$

Furthermore, $h$ is lower semicontinuous, since $\langle \xi, u_n\rangle \to \langle \xi, u\rangle$ in $\sigma(L^p, L^q)$ if $u_n \to u$ as $n \to \infty$ and $\mathsf{e}$ is assumed to be lower semicontinuous. These three properties identify support functions of convex closed sets, see [92, Th. 1.7.1]. Therefore, there exists a (possibly, unbounded) convex closed set $F$ such that its support function

$$h(F,u) = \sup\{\langle x, u\rangle :\ x \in F\}$$

is given by (5.2.1). This set is denoted by $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ or $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\mu)$. The construction can be summarised by the equality

$$h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi), u) = \mathsf{e}(\langle \xi, u\rangle), \quad u \in \mathbb{R}^d. \tag{5.2.2}$$

The following result shows that $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ is a set-valued sublinear function of $\xi$, called a *set-valued sublinear expectation* generated by $\mathsf{e}$. In other instances, we pass to $\boldsymbol{\mathcal{E}}_{\mathsf{e}}$ the sub- and superscripts of $\mathsf{e}$, e.g., $\boldsymbol{\mathcal{E}}_{[\tau]}$ is obtained by choosing $\mathsf{e}$ to be the expectile $\mathsf{e}_{[\tau]}$.

For convex closed sets $F, F'$, their (closed) Minkowski sum $F + F'$ is the closure of $\{x + x' :\ x \in F, x' \in F'\}$, and the dilation of $F$ by $c > 0$ is $cF = \{cx :\ x \in F\}$.

**Theorem 5.2.1.** *Fix $p \in [1,\infty]$ and a law-determined sublinear expectation $\mathsf{e}$ defined on $L^p(\mathbb{R})$. The corresponding map $\boldsymbol{\mathcal{E}}_{\mathsf{e}}$ (given at (5.2.2)) from $L^p(\mathbb{R}^d)$ to the family of convex closed sets in $\mathbb{R}^d$ satisfies the following properties:*

*i) monotonicity: if $\xi \in F$ a.s. for a convex closed $F$, then $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) \subseteq F$;*

*ii) singleton preserving: $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(a) = \{a\}$ for all deterministic $a$;*

*iii) affine equivariance $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(A\xi + a) = A\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) + a$ for all matrices $A$ and $a \in \mathbb{R}^d$;*

*iv) subadditivity:* $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi+\eta) \subseteq \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)+\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\eta)$*;*

*v) lower semicontinuity of support functions, that is,* $h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi),u) \leq \liminf_{n\to\infty} h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi_n),u)$ *for all* $u \in \mathbb{R}^d$ *if* $\xi_n \to \xi$ *in* $\sigma(L^p, L^q)$*;*

*vi) if* $\mathsf{e}(\beta)$ *is finite for all* $\beta \in L^p(\mathbb{R})$*, then the map* $\xi \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ *is continuous in the Hausdorff metric (see, [92, Sec, 1.8]) with respect to the norm on* $L^p$*;*

*vii) if* $\mathsf{e}$ *is continuous from below, then* $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ *contains the expectation* $\mathbf{E}\xi$*.*

*Proof.* Property (i) holds since $\langle \xi, u\rangle \leq h(F,u)$ and in view of the monotonicity property of $\mathsf{e}$. Property (ii) directly follows from the construction, and, for the affine equivariance, note that

$$h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(A\xi+a),u) = \mathsf{e}(\langle \xi, A^\top u\rangle) + \langle a,u\rangle = h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi),A^\top u) + \langle a,u\rangle = h(A\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)+a,u).$$

The subadditivity follows from

$$h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi+\eta),u) = \mathsf{e}(\langle \xi+\eta,u\rangle) \leq \mathsf{e}(\langle \xi,u\rangle) + \mathsf{e}(\langle \eta,u\rangle) = h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi),u) + h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\eta),u).$$

If $\xi_n \to \xi$ in $\sigma(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d))$, then $\langle \xi_n, u\rangle \to \langle \xi, u\rangle$ in $\sigma(L^p(\mathbb{R}), L^q(\mathbb{R}))$. By the lower semicontinuity of $\mathsf{e}$,

$$\mathsf{e}(\langle \xi,u\rangle) \leq \liminf_{n\to\infty} \mathsf{e}(\langle \xi_n,u\rangle).$$

This implies the lower semicontinuity of the support functions.

Property (vi) follows from the Extended Namioka Theorem, which says that every finite sublinear expectation is continuous with respect to the norm topology, see [18]. Recall that sublinear expectations on $L^\infty$ is also Lipschitz. Hence, $\mathsf{e}(\langle \xi_n,u\rangle) \to \mathsf{e}(\langle \xi,u\rangle)$ if $\xi_n \to \xi$ in $L^p$. The convergence of support functions implies the convergence of the corresponding sets in the Hausdorff metric, see [92, Th. 1.8.15].

Finally, (vii) is a consequence of the dilatation monotonicity property (5.1.3). □

*Example* 5.2.2. If $\mathsf{e}$ is the essential supremum, then $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ equals the closed convex hull of the support of $\xi$.

If $p=\infty$, then an easy argument shows that the map $\xi \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ between $L^\infty(\mathbb{R}^d)$ and the family of convex compact sets in $\mathbb{R}^d$ is 1-Lipschitz, that is, the Hausdorff distance between $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ and $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\eta)$ is at most $\|\xi-\eta\|_\infty$ for all $\xi,\eta \in L^\infty(\mathbb{R}^d)$. Indeed,

$$h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi),u) - h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\eta),u) = \mathsf{e}(\langle \xi,u\rangle) - \mathsf{e}(\langle \eta,u\rangle) \leq \mathsf{e}(\langle \eta,u\rangle + \|\xi-\eta\|_\infty) - \mathsf{e}(\langle \eta,u\rangle) = \|\xi-\eta\|_\infty$$

for all unit $u \in \mathbb{R}^d$.

If $\xi, \eta \in L^p(\mathbb{R}^d)$ and $\mathbf{E}(\eta|\xi) = 0$ a.s., then the dilatation monotonicity property (5.1.3) implies that

$$\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi + \eta) \supseteq \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\mathbf{E}(\xi + \eta|\xi)) = \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi).$$

Hence, if $\xi_1, \xi_2, \dots$ is a sequence of i.i.d. centred $p$-integrable random vectors, then $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi_1 + \dots + \xi_n)$, $n \geq 1$, is a growing sequence of nested convex sets in $\mathbb{R}^d$.

*Remark* 5.2.3. If $\xi$ is dominated by $\eta$ in the *convex order*, meaning that $\mathbf{E}f(\xi) \leq \mathbf{E}f(\eta)$ for all convex functions $f$, then $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) \subseteq \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\eta)$, see [39, Cor. 4.59]. In particular, the sequence $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi_n)$, $n \geq 1$, grows if $(\xi_n)_{n\geq 0}$ is a martingale.

*Example* 5.2.4. Let $\langle \xi, u \rangle$ be distributed as $\zeta \|u\|_L$, where $\zeta$ is a random variable and $\|\cdot\|_L$ is a certain norm on $\mathbb{R}^d$ with $L$ being the unit ball; then $\xi$ is called pseudo-isotropic, see, e.g., [44]. In this case, $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) = cL^o$, where

$$L^o = \{u : \; h(L, u) \leq 1\} \tag{5.2.3}$$

is the *polar set* to $L$ and $c = \mathsf{e}(\zeta) = \mathsf{e}(\langle \xi, u \rangle)$ for any given $u \in \partial L$. For instance, this is the case if $\xi$ is symmetric $\alpha$-stable with $\alpha \in (1, 2]$; then $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ is expressed in terms of the associated convex body of $\xi$, see [70]. If $\xi$ is Gaussian, then $L^o$ is the ellipsoid determined by the covariance matrix of $\xi$ and translated by the mean of $\xi$.

The dual representation of $\mathsf{e}$ given by Theorem 5.1.3 immediately implies the following result.

**Corollary 5.2.5.** *The set-valued sublinear expectation generated by* $\mathsf{e}$ *can be represented as*

$$\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) = \operatorname{cl}\{\mathbf{E}(\xi\gamma) : \; \gamma \in \mathcal{M}_{\mathsf{e}}, \mathbf{E}\gamma = 1\}, \tag{5.2.4}$$

*where* cl *denotes the topological closure in* $\mathbb{R}^d$ *and* $\mathcal{M}_{\mathsf{e}}$ *is the family of probability measures from* (5.1.1).

The convexity of $\mathcal{M}_{\mathsf{e}}$ implies that the set on the right-hand side of (5.2.4) is convex. This set can be written as the intersection of the cone $\{(\mathbf{E}\gamma, \mathbf{E}(\xi\gamma)) : \; \gamma \in \mathcal{M}_{\mathsf{e}}\}$ with the set $\{1\} \times \mathbb{R}^d$ and then projected on its last $d$-components.

*Remark* 5.2.6. It is possible to construct a variant of the set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ by applying the underlying sublinear expectation $\mathsf{e}$ to the positive part $(\langle \xi, u \rangle)_+$ of the scalar product of $\xi$ and $u$. The obtained function is the support function of a convex closed set, which may be considered a sublinear expectation of the segment $[0, \xi]$, see [72] for a study of sublinear expectations with set-valued arguments.

### 5.2.2 Less regular maps

One might also consider a variant of the sublinear expectation which is a positive homogeneous, subadditive and lower semicontinuous function $\mathsf{g} : L^p(\mathbb{R}) \to (-\infty, \infty]$ and so is not necessarily monotone or translation equivariant. We refer to this function as a *convex gauge.* The most important example is the $L^p$-norm, so that $\mathsf{g}(\beta) = \|\beta\|_p$, which is convex but not translation equivariant.

For a lower semicontinuous convex gauge $\mathsf{g}$, we define $\boldsymbol{\mathcal{G}}(\xi)$ as the convex closed set such that

$$h(\boldsymbol{\mathcal{G}}(\xi), u) = \mathsf{g}(\langle \xi, u\rangle), \quad u \in \mathbb{R}^d.$$

It is easily seen that $\mathsf{g}(\langle \xi, u\rangle)$ is indeed a support function.

*Example* 5.2.7. Let $\mathsf{g}(\beta) = \|\beta\|_p$. For $\xi \in L^p(\mathbb{R}^d)$, the convex body $\boldsymbol{\mathcal{G}}(\xi)$ is the $L^p$-centroid of $\xi$ (or of its distribution $\mu$). These convex bodies have been introduced in [83] for $p = 1$ and in [67] for a general $p$, and further thoroughly studied, see, e.g., [37, 46, 77].

In some cases, $\mathsf{g}$ fails to be convex. For instance, this is the case for $L^p$-norm with $p \in (0, 1)$. Another important case arises when $\mathsf{g}(\beta)$ is the quantile function $q_t(\beta)$ given by (5.1.5) for a fixed $t \in (0, 1)$, which is known to be not necessarily subadditive in $\beta$. In the absence of subadditivity, it is natural to consider the largest convex set whose support function is dominated by the quantile function of $\langle x, u\rangle$, namely, let

$$D_\delta(\xi) = \bigcap_{u\in\mathbb{R}^d} \left\{x \in \mathbb{R}^d : \ \langle x, u\rangle \le q_{1-\delta}(\langle \xi, u\rangle)\right\}. \tag{5.2.5}$$

The set $D_\delta(\xi)$ is called the *depth-trimmed region* of $\xi$. The support function of $D_\delta(\xi)$ may be strictly less than $q_{1-\delta}(\langle \xi, u\rangle)$, for example, if $\xi$ is uniformly distributed on a triangle on the plane, see [63]. The set $D_\delta(\xi)$ is necessarily empty if $\xi$ is nonatomic and $\delta \in (1/2, 1]$.

The set $D_\delta(\xi)$ is related to the *Tukey (or half-space) depth* (see [103]), which associates to a point $x$ the smallest $\mu$-content of a half-space containing $x$, where $\mu$ is the distribution of $\xi$. The depth-trimmed region of $\xi$ is the excursion set of the Tukey depth, so that

$$D_\delta(\xi) = \bigcap_{\mu(H)>1-\delta} H\,, \tag{5.2.6}$$

where $H$ runs through the collection of all closed half-spaces. If $\xi$ has *contiguous support* (that is, the support of $\langle \xi, u\rangle$ is connected for every $u$), then (5.2.5) holds with $q$ being any other quantile function in case of multiplicities, and the intersection in (5.2.6) can be taken over half-spaces $H$ with $\mu(H) \ge 1 - \delta$, see [22, 54].

*Example* 5.2.8. Let $\xi$ be uniformly distributed on a convex body $K$. Then $D_\delta(\xi)$ is the *convex floating body* of $K$, see [97] and [108]. A variant of this concept for nonuniform distributions on $K$ has been studied in [16].

Recall that a random vector $\xi$ with distribution $\mu$ is said to have *k-concave distribution*, with $k \in [-\infty, \infty]$, if

$$\mu(\theta A + (1-\theta)B) \geq \begin{cases} \min\{\mu(A), \mu(B)\} & \text{if } k = -\infty, \\ \mu(A)^{\theta}\mu(B)^{(1-\theta)} & \text{if } k = 0, \\ (\theta\mu(A)^k + (1-\theta)\mu(B)^k)^{1/k} & \text{otherwise,} \end{cases}$$

for all Borel sets $A$ and $B$ and $\theta \in [0,1]$. In case of $k = 0$, the measure $\mu$ is called *log-concave*. The next theorem establishes some conditions under which $q_\delta(\langle \xi, u \rangle)$ is a support function; it is a direct consequence of [19, Th. 6.1].

**Theorem 5.2.9.** *Let $\xi$ be a symmetric $k$-concave random vector with $k \geq -1$ and such that the support of $\xi$ is full-dimensional. Then*

$$h(D_\delta(\xi), u) = q_{1-\delta}(\langle \xi, u \rangle), \quad u \in \mathbb{R}^d,$$

*for all $\delta \in (0, 1/2)$.*

## 5.3 Convex bodies generated by average quantiles

One of the most important sublinear expectations is based on using weighted integrals of the quantile function. The corresponding convex bodies are studied in this section, where we show their close connection to metronoids [48] and zonoid-trimmed regions [56]. The Kusuoka representation of numerical sublinear expectations yields Theorem 5.3.4, which provides a representation of a general convex set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ (derived from $\xi$ using a sublinear expectation $\mathsf{e}$) in terms of Aumann integrals of metronoids. We further provide a uniqueness result for the distribution of $\xi$ on the basis of a family of convex bodies generated by it, and also a concentration result for random convex sets constructed from the empirical distribution of $\xi$.

### 5.3.1 Metronoids and zonoid-trimmed regions

For $\xi \in L^1(\mathbb{R}^d)$ and $\alpha \in (0,1]$, denote by $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ the convex set generated by the average quantile sublinear expectation $\mathsf{e}_\alpha$ given by (5.1.4). Such convex sets are hereafter called *average quantile sets*. In particular, $\boldsymbol{\mathcal{E}}_1(\xi) = \mathbf{E}\xi$. Since $\mathsf{e}_\alpha$ is finite on $L^1(\mathbb{R})$, the set $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ is compact. Noticing that $q_t(-\beta) = -q_{1-t}(\beta)$, it is easy to see that $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ has nonempty interior for all $\alpha \in (0,1)$, hence, is a convex body. The set $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ increases as $\alpha$ decreases to zero with limit $\boldsymbol{\mathcal{E}}_0(\xi)$, being the convex hull of the support of $\xi$.

The following result relates average quantile sets and the *zoniod-trimmed regions* introduced in [56] as

$$Z_\alpha(\xi) = \big\{\mathbf{E}(\xi f(\xi)) : \ f : \mathbb{R}^d \to [0, \alpha^{-1}] \, \text{measurable and} \, \mathbf{E}f(\xi) = 1\big\}.$$

**Proposition 5.3.1.** *For all $\alpha \in (0,1]$, $\boldsymbol{\mathcal{E}}_\alpha(\xi) = Z_\alpha(\xi)$.*

*Proof.* Representation (5.1.6) yields that

$$h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) = \mathsf{e}_\alpha(\langle \xi, u\rangle) = \sup_{\gamma \in L^\infty([0,\alpha^{-1}]),\, \mathbf{E}\gamma = 1} \langle \mathbf{E}(\gamma\xi), u\rangle \,.$$

Noticing that any $\gamma$ in the last expression can be replaced with $\mathbf{E}(\gamma|\xi)$ yields that

$$\sup_{\substack{\gamma \in L^\infty([0,\alpha^{-1}]) \\ \mathbf{E}\gamma = 1}} \langle \mathbf{E}(\gamma\xi), u\rangle = \sup_{\substack{f\colon \mathbb{R}^d \to [0,\alpha^{-1}] \\ \mathbf{E}f(\xi)=1}} \langle \mathbf{E}(\xi f(\xi)), u\rangle. \qquad \square$$

Let $\mu$ be a locally finite Borel measure on $\mathbb{R}^d$. Denote by $L^1_\mu([0,1])$ the family of functions $f : \mathbb{R}^d \to [0,1]$ such that $\int x f(x)\mu(\mathrm{d}x)$ exists. The set

$$\mathrm{M}(\mu) = \left\{ \int_{\mathbb{R}^d} x f(x)\mu(\mathrm{d}x) : \; f \in L^1_\mu([0,1]), \int_{\mathbb{R}^d} f \mathrm{d}\mu = 1 \right\}$$

has the support function

$$h(\mathrm{M}(\mu), u) = \sup_{f \in L^1_\mu([0,1]), \int f \mathrm{d}\mu = 1} \int \langle x, u\rangle f(x)\mu(\mathrm{d}x), \quad u \in \mathbb{R}^d.$$

The set $\mathrm{M}(\mu)$ was introduced in [48] and called the *metronoid* of $\mu$. This definition applies also for possibly infinite measures $\mu$, e.g., if $\mu$ is the Lebesgue measure, then $\mathrm{M}(\mu) = \mathbb{R}^d$, since each point $x \in \mathbb{R}^d$ can be obtained by letting $f$ be the indicator of the unit ball centred at $x$ normalised by the volume of the unit ball. Furthermore, $\mathrm{M}(\mu)$ is empty if the total mass of $\mu$ is less than one, and $\mathrm{M}(\mu)$ is the singleton $\int x\mu(\mathrm{d}x)$ if $\mu$ is an integrable probability measure. The following result establishes a relation between metronoids and average quantile sets.

**Proposition 5.3.2.** *Let $\mu$ be an integrable probability measure on $\mathbb{R}^d$. Then $\mathrm{M}(\alpha^{-1}\mu) = \boldsymbol{\mathcal{E}}_\alpha(\mu)$ for any $\alpha \in (0,1]$.*

*Proof.* Consider a random vector $\xi$ with distribution $\mu$. By (5.1.6), for every $u \in \mathbb{R}^d$ the support function of $\mathrm{M}(\alpha^{-1}\mu)$ is

$$\begin{aligned} h(\mathrm{M}(\alpha^{-1}\mu), u) &= \sup_{0 \le f \le 1,\, \int f \mathrm{d}\mu = \alpha} \int \langle x, u\rangle f(x) \alpha^{-1}\mu(\mathrm{d}x) \\ &= \sup_{0 \le f \le \alpha^{-1},\, \mathbf{E}f(\xi) = 1} \mathbf{E}\big(\langle \xi, u\rangle f(\xi)\big) \\ &= \sup_{\gamma \in L^\infty([0,\alpha^{-1}]),\, \mathbf{E}\gamma = 1} \mathbf{E}\left(\langle \xi, u\rangle \gamma\right) \\ &= \mathsf{e}_\alpha(\langle \xi, u\rangle) = h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u)\,, \end{aligned}$$

where in the second equality $f\alpha^{-1}$ was replaced by $f$ and later $f(\xi)$ by $\gamma$. $\square$

*Example* 5.3.3. Let $\xi$ have a discrete distribution with atoms at $x_1, \dots, x_n$ of probabilities $p_1, \dots, p_n$. Then $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ is the polytope

$$\boldsymbol{\mathcal{E}}_\alpha(\xi) = \left\{ \sum_{i=1}^n \lambda_i p_i x_i : \ \lambda_1, \dots, \lambda_n \in [0, \alpha^{-1}], \ \sum_{i=1}^n \lambda_i p_i = 1 \right\},$$

see [48, Prop. 2.3], where this is proved for metronoids.

### 5.3.2 A representation of general $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$

Fix $\xi \in L^1(\mathbb{R}^d)$ and consider the average quantile sets $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ as a set-valued function of $\alpha \in (0,1]$. Let $\nu$ be a probability measure on $(0,1]$, which appears in the spectral sublinear expectation (5.1.9) from Example 5.1.5. The closed *Aumann integral* (see [7]) of the set-valued function $\alpha \mapsto \boldsymbol{\mathcal{E}}_\alpha(\xi)$ is the convex closed set $\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi)$, whose support function at any direction $u$ equals the integral of the support function, so that

$$h(\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi), u) = \int_{(0,1]} h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u)\nu(\mathrm{d}\alpha), \quad u \in \mathbb{R}^d. \tag{5.3.1}$$

Recognising the right-hand side as $\mathsf{e}_{\int\varphi}(\langle \xi, u\rangle)$, it is immediately seen that $\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi)$ is the set-valued sublinear expectation generated by the spectral numerical one from Example 5.1.5. Equivalently, $\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi)$ equals the closure of the set of integrals of all measurable integrable functions $f(\alpha)$, $\alpha \in (0,1]$, such that $f(\alpha) \in \boldsymbol{\mathcal{E}}_\alpha(\xi)$ for all $\alpha$, see [7] and [71, Sec. 2.1.2]. This is reflected by writing

$$\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi) = \mathrm{cl} \int_{(0,1]} \boldsymbol{\mathcal{E}}_\alpha(\xi)\nu(\mathrm{d}\alpha). \tag{5.3.2}$$

Since $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ increases to the closed convex hull of the support of $\xi$ as $\alpha \downarrow 0$, the set $\boldsymbol{\mathcal{E}}_{\int\varphi}(\xi)$ is not necessarily bounded.

The following result provides a representation of the set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ constructed using a general law-determined sublinear expectation $\mathsf{e}$. It confirms that the average quantile sets (equivalently, metronoids) are building blocks for a general $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$. Denote by $\mathrm{conv}\, A$ the closed convex hull of a set $A$ in $\mathbb{R}^d$.

**Theorem 5.3.4.** *For each $\xi \in L^p(\mathbb{R}^d)$ and a set-valued sublinear expectation $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ generated by a law-determined sublinear expectation $\mathsf{e}$, we have*

$$\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi) = \mathrm{conv} \bigcup_{\nu \in \mathcal{P}_{\mathsf{e}}} \int_{(0,1]} \boldsymbol{\mathcal{E}}_\alpha(\xi)\nu(\mathrm{d}\alpha),$$

*where $\mathcal{P}_{\mathsf{e}}$ is the family probability measures $\nu$ on $(0,1]$ from the Kusuoka representation of $\mathsf{e}$, see* (5.1.7).

*Proof.* By Theorem 5.1.4,

$$\mathsf{e}(\langle\xi,u\rangle)=\sup_{\nu\in\mathcal{P}_{\mathsf{e}}}\int_{(0,1]}\mathsf{e}_\alpha(\langle\xi,u\rangle)\nu(\mathrm{d}\alpha)=\sup_{\nu\in\mathcal{P}_{\mathsf{e}}}\int_{(0,1]}h(\boldsymbol{\mathcal{E}}_\alpha(\xi),u)\nu(\mathrm{d}\alpha).$$

The proof is completed by noticing (5.3.1), using the notation (5.3.2) and the fact that the supremum of support functions is the support function of the closed convex hull of the involved sets. □

### 5.3.3 Average quantile sets as integrated depth-trimmed regions

Under the symmetry and log-concavity assumptions on $\xi$, the average quantile sets $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ can be characterised as set-valued integrals of the depth-trimmed regions (equivalently, weighted floating bodies) $D_\delta(\xi)$ introduced in (5.2.5). Similarly to (5.3.1), the closed Aumann integral of the function $t\mapsto D_t(\xi)$ with respect to a measure $\nu$ on $[0,1]$ is defined as the convex set whose support function equals the integral of the support functions of $D_t(\xi)$, that is,

$$h\Big(\int_0^1 D_t(\xi)\nu(\mathrm{d}t),u\Big)=\int_0^1 h(D_t(\xi),u)\nu(\mathrm{d}t),\quad u\in\mathbb{R}^d.$$

If the measure $\nu$ attaches positive mass to the set of $t\in[0,1]$ where $D_t(\xi)$ is empty, the integral is set to be the empty set.

The following result establishes relationships between average quantile sets (or metronoids) and depth-trimmed regions. Its second part generalises [49, Th. 1.1], which concerns the case of $\xi$ supported by a convex body.

**Theorem 5.3.5.** *Let $\xi\in L^1(\mathbb{R}^d)$. Then*

$$D_\alpha(\xi)\subseteq\frac{1}{\alpha}\int_0^\alpha D_t(\xi)\mathrm{d}t\subseteq\boldsymbol{\mathcal{E}}_\alpha(\xi).\tag{5.3.3}$$

*If $\xi$ has a log-concave distribution, then*

$$D_{\frac{e-1}{e}\alpha}(\xi)\subseteq\boldsymbol{\mathcal{E}}_\alpha(\xi)\subseteq D_{\frac{\alpha}{e}}(\xi)\tag{5.3.4}$$

*for every $\alpha\in(0,1]$.*

*Proof.* By definition of the average quantile set,

$$h(\boldsymbol{\mathcal{E}}_\alpha(\xi),u)=\frac{1}{\alpha}\int_{1-\alpha}^1 q_t(\langle\xi,u\rangle)\mathrm{d}t\geq\frac{1}{\alpha}\int_{1-\alpha}^1 h(D_{1-t}(\xi),u)\mathrm{d}t=\frac{1}{\alpha}\int_0^\alpha h(D_t(\xi),u)\mathrm{d}t,$$

where the inequality follows from (5.2.5). Finally, (5.3.3) follows from the monotonicity of $D_t(\xi)$.

Fix $u \in \mathbb{R}^d$. Consider $\beta = \langle \xi, u \rangle$ and note that the distribution $\nu$ of $\beta$ is log-concave by the invariance of the log-concavity property under projection. For (5.3.4), it suffices to show that

$$q_{\left(1-\frac{e-1}{e}\alpha\right)}(\beta) \le \mathsf{e}_\alpha(\beta) \le q_{\left(1-\frac{1}{e}\alpha\right)}(\beta). \tag{5.3.5}$$

Being the projection of a log-concave vector, $\beta$ is either deterministic or absolutely continuous with connected support. In the first case (5.3.5) becomes trivial, thus we can assume that $\beta$ is absolutely continuous with connected support. In particular, $q$ in (5.3.5) can be equivalently chosen to be the left- or the right-quantile function. Observe that, for measurable sets $A$ and $B$, convex $C$ and $\theta \in [0,1]$,

$$\begin{aligned}
\nu\left(A \cap C\right)^{\theta} \nu\left(B \cap C\right)^{1-\theta} &\le \nu\left(\theta\left(A \cap C\right) + \left(1-\theta\right)\left(B \cap C\right)\right) \\
&= \nu\left(\left(\theta A \cap \theta C\right) + \left(\left(1-\theta\right)B \cap \left(1-\theta\right)C\right)\right) \\
&\le \nu\left(\left(\theta A + \left(1-\theta\right)B\right) \cap \left(\theta C + \left(1-\theta\right)C\right)\right) \\
&= \nu\left(\left(\theta A + \left(1-\theta\right)B\right) \cap C\right) .
\end{aligned}$$

Therefore, the probability measure obtained by restricting $\nu$ to the interval $(q_{1-\alpha}(\beta), \infty)$ and normalising by the factor $\alpha{-}1$ is log-concave, and we consider a random variable $X$ with such distribution. It follows from the theory of risk measures (see, e.g., [102, Prop. 2.1]), that for the case of absolutely continuous random variables, supremum in the characterisation of $\mathsf{e}_\alpha(\beta)$ in (5.1.6) is attained at $\gamma = \alpha^{-1}\mathbf{1}_{\{\beta > q_{1-\alpha}(\beta)\}}$, which implies

$$\mathbf{E}X = \alpha^{-1}\mathbf{E}\left(\beta \mathbf{1}_{\{\beta > q_{1-\alpha}(\beta)\}}\right) = \mathsf{e}_\alpha(\beta) \,. \tag{5.3.6}$$

It follows from [19, Eq. (5.7)] that for any log-concave random variable $X$,

$$e^{-1} \le \mathbb{P}\left\{X > \mathbf{E}X\right\} \le 1 - e^{-1} \,. \tag{5.3.7}$$

Therefore, (5.3.6) and (5.3.7) yield that

$$e^{-1} \le \alpha^{-1}\nu(\mathsf{e}_\alpha(\beta), \infty) \le 1 - e^{-1} \,.$$

Hence,

$$e^{-1}\alpha \le \mathbb{P}\left\{\beta > \mathsf{e}_\alpha(\beta)\right\} \le \left(1 - e^{-1}\right)\alpha \,,$$

which implies (5.3.5), given that $\beta$ has connected support. □

### 5.3.4 A uniqueness result for maximum extensions

A single set $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ surely does not characterise the distribution of $\xi$. However, families of such sets can be sufficient to recover the distribution of $\xi$.

*Example* 5.3.6. Assume that $\xi, \eta \in L^1(\mathbb{R}^d)$ and consider the average quantile sets $\boldsymbol{\mathcal{E}}_\alpha(\xi)$ and $\boldsymbol{\mathcal{E}}_\alpha(\eta)$. If $\boldsymbol{\mathcal{E}}_\alpha(\xi) = \boldsymbol{\mathcal{E}}_\alpha(\eta)$ for all $\alpha \in (0, 1/2]$, then $\xi$ and $\eta$ have the same distribution. This follows from Proposition 5.3.1 and [56, Th. 5.6].

Since the definition of $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ is based on the univariate sublinear expectation $\mathsf{e}$ applied to the projections of $\xi$, the following result is a straightforward application of the Cramér–Wold theorem, see, e.g., [51, Cor. 5.5].

**Proposition 5.3.7.** *A family of sets $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$, $\mathsf{e} \in E$, generated by sublinear expectations $\mathsf{e}$ from a certain family $E$ uniquely identifies the distribution of $\xi \in L^p(\mathbb{R}^d)$ if and only if the family of the underlying univariate sublinear expectations $\mathsf{e}(\beta)$, $\mathsf{e} \in E$, uniquely identifies the distribution of any $\beta \in L^p(\mathbb{R})$.*

Natural families of sublinear expectations arise by applying the maximum extension to a given sublinear expectation.

*Example* 5.3.8. Consider the expected maximum sublinear expectation $\mathsf{e}_1^{\vee m}$ given by (5.1.14). Then the convex body $\boldsymbol{\mathcal{E}}_1^{\vee m}(\xi)$ is the expectation $\mathbf{E}P_m$ of the random polytope $P_m$ obtained as the convex hull of $m$ independent copies of $\xi$, see [71, Sec. 2.1]. It is well known that the sequence $\mathsf{e}_1^{\vee m}(\beta)$, $m \geq 1$, uniquely identifies the distribution of $\beta \in L^1(\mathbb{R})$, see [47] and [42]. As a consequence, the nested sequence $\mathbf{E}P_m$, $m \geq 1$, of convex bodies uniquely determines the distribution of $\xi$, see [106].

Applying the maximum extension (5.1.13) to the spectral sublinear expectation $\mathsf{e}_{\int\varphi}(\cdot)$ yields the sublinear expectation $\mathsf{e}_{\int\varphi}^{\vee m}(\cdot)$ and the corresponding sequence of nested convex bodies $\boldsymbol{\mathcal{E}}_{\int\varphi}^{\vee m}(\xi)$, $m \geq 1$.

**Theorem 5.3.9.** *Let $\xi, \eta \in L^1(\mathbb{R}^d)$. For any constant $c \geq 0$, consider the spectral function $\varphi(t) = (c+1)(1-t)^c$. If*

$$\boldsymbol{\mathcal{E}}_{\int\varphi}^{\vee m}(\xi) = \boldsymbol{\mathcal{E}}_{\int\varphi}^{\vee m}(\eta), \quad m \geq 1,$$

*then $\xi$ and $\eta$ have the same distribution.*

*Proof.* In view of Proposition 5.3.7, it suffices to prove this result for two random variables $\beta$ and $\gamma$. For any integer $m \geq 1$, we have

$$\int_0^1 q_{1-t}\left(\max(\beta_1, \ldots, \beta_m)\right) \varphi(t)\mathrm{d}t = \int_0^1 q_{1-t}\left(\max(\gamma_1, \ldots, \gamma_m)\right) \varphi(t)\mathrm{d}t,$$

where $\beta_i, \gamma_i$, $i = 1, \ldots, m$, are independent copies of $\beta, \gamma$, respectively. By a change of variables,

$$\begin{aligned}\int_0^1 q_{1-t}\left(\max(\beta_1, \ldots, \beta_m)\right) \varphi(t)\mathrm{d}t &= (c+1)\int_0^1 q_t\left(\max(\beta_1, \ldots, \beta_m)\right) t^c\mathrm{d}t \\ &= (c+1)\int_0^1 q_{t^{\frac{1}{m}}}(\beta) t^c\mathrm{d}t \\ &= m(c+1)\int_0^1 q_s(\beta) s^{cm+m-1}\mathrm{d}s.\end{aligned}$$

Therefore,

$$\int_0^1 f(s)s^{(c+1)(m-1)}\mathrm{d}s = 0\,, \quad m \geq 1\,,$$

with

$$f(s) = s^c\left(q_s(\beta) - q_s(\gamma)\right) \in L^1([0,1])\,.$$

The family

$$\mathcal{A} = \left\{c_0 + \sum_{i=1}^{n} c_i x^{(c+1)m_i}\,:\, n, m_1, \dots, m_n \in \mathbb{N},\, c_0, \dots, c_n \in \mathbb{R}\right\}$$

is an algebra of continuous functions separating the points on $[0,1]$. By linearity of the Lebesgue integral

$$\int_0^1 f(s)a(s)\mathrm{d}s = 0$$

for all $a \in \mathcal{A}$. The Stone–Weierstrass theorem (see, e.g., [38, Th. 4.45]) yields that

$$\int_0^1 f(s)g(s)\mathrm{d}s = 0$$

for all continuous functions $g$ on $[0,1]$. Therefore, $f$ vanishes almost everywhere, so the proof is complete. □

### 5.3.5 Concentration of empirical average quantile sets

Let $\xi \in L^p(\mathbb{R}^d)$ with distribution $\mu$. Consider the empirical random measure constructed by $n$ independent copies $\xi_1, \dots, \xi_n$ of $\xi$ as

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i}\,, \quad n \geq 1\,, \tag{5.3.8}$$

where $\delta_x$ is the one point mass measure at $x \in \mathbb{R}^d$. The average quantile convex body $\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n)$ generated by $\hat{\mu}_n$ is a random convex set, which approximates the body $\boldsymbol{\mathcal{E}}_\alpha(\mu)$ as $n$ grows to infinity. In fact, the sequence $\{\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n),\, n \geq 1\}$ almost surely converges to $\boldsymbol{\mathcal{E}}_\alpha(\mu)$ in the Hausdorff metric, as directly follows from [56, Th. 5.2] and Proposition 5.3.1. The following theorem provides probabilistic bounds for this convergence.

**Theorem 5.3.10.** *Let $\mu$ be a probability measure with bounded support of diameter $R$, and let $r$ be the largest radius of a centred Euclidean ball contained in the average quantile set $\boldsymbol{\mathcal{E}}_\alpha(\mu)$ for some $\alpha \in (0,1)$. For all $\varepsilon > 0$ and $n \in \mathbb{N}$,*

$$\mathbb{P}\Big\{(1-\varepsilon)\boldsymbol{\mathcal{E}}_\alpha(\mu) \subseteq \boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n) \subseteq (1+\varepsilon)\boldsymbol{\mathcal{E}}_\alpha(\mu)\Big\} \geq 1 - 6^{d+1}(1+1/\varepsilon)^d \exp\left\{-\frac{\alpha\varepsilon^2 r^2 n}{44R^2}\right\}.$$

We use the following auxiliary result.

**Lemma 5.3.11** (see [41, Lemma 5.2]). *Let $K$ be a convex body which contains the origin in its interior. For each $\delta \in (0, 1/2)$, there exists a set $\mathcal{N} \subseteq \partial K$ with cardinality at most $(3/\delta)^d$ such that each $v \in \partial K$ satisfies*

$$v = w_0 + \sum_{i=1}^{\infty} \delta_i w_i$$

*for $w_i \in \mathcal{N}$, $i \geq 0$, and $\delta_i \in [0, \delta^i]$, $i \geq 1$.*

*Proof of Theorem 5.3.10.* On a (possibly enlarged) probability space $\Omega \times \Omega'$, let $\xi$ be a $\mu$-distributed random vector and let $\hat{\xi}_n$ take one of the values $\xi_1, \ldots, \xi_n$ with equal probabilities. For any fixed $u \in \mathbb{R}^d$,

$$h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) = \frac{1}{\alpha} \int_{1-\alpha}^{1} q_t(\langle \xi, u \rangle) \mathrm{d}t$$

and

$$h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), u) = \frac{1}{\alpha} \int_{1-\alpha}^{1} q_t(\langle \hat{\xi}_n, u \rangle) \mathrm{d}t \,.$$

Clearly, $\langle \hat{\xi}_n, u \rangle$ is distributed according to the empirical distribution function generated by the sample $\langle \xi_i, u \rangle$, $i = 1, \ldots, n$. Thus, the right-hand sides of the two equations are, respectively, the conditional value at risk of $\beta = \langle \xi, u \rangle$ and its sample-based estimator, see [20, 107]. Note that the support of $\beta$ is a subset of an interval of length $R$. By [107, Th. 3.1], for any $\eta > 0$,

$$\mathbb{P}\left\{h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), u) \leq h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) - \eta\right\} \leq 3 \exp\left\{-\frac{\alpha \eta^2 n}{5R^2}\right\}$$

and

$$\mathbb{P}\left\{h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), u) \geq h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) + \eta\right\} \leq 3 \exp\left\{-\frac{\alpha \eta^2 n}{11R^2}\right\} .$$

Noticing that the second bound is larger than the first one and that $h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) \geq r$ by the imposed condition, we obtain

$$\mathbb{P}\Big\{(1 - \varepsilon/2) h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) \leq h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), u) \leq (1 + \varepsilon/2) h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u)\Big\} \geq 1 - 6 \exp\left\{-\frac{\alpha \varepsilon^2 r^2 n}{44R^2}\right\} . \quad (5.3.9)$$

Let $\mathcal{N} \subseteq \partial \boldsymbol{\mathcal{E}}_\alpha(\xi)^o$ be a set from Lemma 5.3.11, with $\delta = \frac{\varepsilon}{2+2\varepsilon}$, where $\boldsymbol{\mathcal{E}}_\alpha(\xi)^o$ is the polar set to $\boldsymbol{\mathcal{E}}_\alpha(\xi)$, see (5.2.3). Since $h(\boldsymbol{\mathcal{E}}_\alpha(\xi), u) = 1$ for all $u \in \partial \boldsymbol{\mathcal{E}}_\alpha(\xi)^o$, the union bound applied to (5.3.9) yields that

$$(1 - \varepsilon/2) \leq h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), w) \leq (1 + \varepsilon/2) \quad \text{for all} \;\; w \in \mathcal{N} \quad (5.3.10)$$

with probability at least

$$1 - 6 \left(\frac{6 + 6\varepsilon}{\varepsilon}\right)^d \exp\left\{-\frac{\alpha \varepsilon^2 r^2 n}{44R^2}\right\} .$$

For any $v \in \partial\boldsymbol{\mathcal{E}}_\alpha(\xi)^o$ and some sequences $w_i \in \mathcal{N}$ and $\delta_i \geq 0$, $i \geq 1$, the sublinearity of $h$, Lemma 5.3.11 and (5.3.10) imply that

$$\begin{aligned} h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), v) &= h\left(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), w_0 + \sum_{i=1}^{\infty} \delta_i w_i\right) \\ &\leq (1+\varepsilon/2) \sum_{i=0}^{\infty} \left(\frac{\varepsilon}{2+2\varepsilon}\right)^i \\ &= (1+\varepsilon/2)\, \frac{1}{1-\left(\frac{\varepsilon}{2+2\varepsilon}\right)} = (1+\varepsilon)\, h(\boldsymbol{\mathcal{E}}_\alpha(\xi), v) \end{aligned}$$

and

$$\begin{aligned} h(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), v) &= h\left(\boldsymbol{\mathcal{E}}_\alpha(\hat{\mu}_n), w_0 + \sum_{i=1}^{\infty} \delta_i w_i\right) \\ &\geq (1-\varepsilon/2) - (1+\varepsilon/2) \sum_{i=1}^{\infty} \left(\frac{\varepsilon}{2+2\varepsilon}\right)^i \\ &= (1-\varepsilon/2) - (1+\varepsilon/2)\, \frac{\left(\frac{\varepsilon}{2+2\varepsilon}\right)}{1-\left(\frac{\varepsilon}{2+2\varepsilon}\right)} = (1-\varepsilon)\, h(\boldsymbol{\mathcal{E}}_\alpha(\xi), v)\,, \end{aligned}$$

which deliver the desired assertion. □

## 5.4 Floating-like bodies

In this section, we specialize our general construction to the case when $\xi$ is uniformly distributed on a convex body $K$ (that is, a compact convex set in $\mathbb{R}^d$ with nonempty interior), and so $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ yields a transform $K \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K) = \boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$. We derive several properties of this transformation for general $\mathsf{e}$, in particular, establish the continuity of such maps in the Hausdorff metric.

In special cases, our construction yields $L^p$-centroid bodies (see [67] and [92, Sec. 10.8]) and Ulam floating bodies recently introduced in [49]. The latter form a particularly important special setting, which is confirmed by showing that all transformations $K \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ can be expressed in terms of Ulam floating bodies. For instance, Corollary 5.4.8 provides a representation of the centroid body of an origin symmetric $K$ as the convex hull of dilated Ulam floating bodies of $K$. In this course, results for sublinear expectations yield a new insight into the well-known aforementioned constructions of convex bodies, deliver some new relations between them, and provide a general source of nonlinear transformations of convex bodies. Finally, we formulate several conjectures.

### 5.4.1 Sublinear transform

Consider the set-valued sublinear expectation $\boldsymbol{\mathcal{E}}_{\mathsf{e}}$ generated by a law-determined numerical sublinear expectation $\mathsf{e}$. Let $\xi$ be a random vector uniformly distributed on a convex body $K \subseteq \mathbb{R}^d$. Recall that $K$ is assumed to have a nonempty interior. In the following, we write $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ instead of $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(\xi)$ and refer to $K \mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ as a *sublinear transform* of $K$ generated by the numerical sublinear expectation $\mathsf{e}$. We also refer to $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ as a *floating-like body*.

Denoting by $\mathcal{K}$ the family of convex bodies in $\mathbb{R}^d$, the sublinear transform is a map $\boldsymbol{\mathcal{E}}_{\mathsf{e}} : \mathcal{K} \to \mathcal{K}$. It is easy to see that $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K) \subseteq K$ for all $K$. If $\xi$ is uniformly distributed on $K$ and $A$ is a nondegenerate matrix, then $A\xi$ is uniformly distributed on $AK$. Thus,

$$\boldsymbol{\mathcal{E}}_{\mathsf{e}}(AK + a) = A\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K) + a, \quad a \in \mathbb{R}^d.$$

The sublinear transform $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(B)$ of a centred Euclidean ball $B$ is another centred Euclidean ball, which is contained in $B$. Furthermore, the sublinear transform of an ellipsoid is also an ellipsoid.

The sublinear transform is not necessarily monotone for inclusion, see Example 5.4.6. In view of Remark 5.2.3, $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K) \subseteq \boldsymbol{\mathcal{E}}_{\mathsf{e}}(L)$ for all sublinear transforms $\boldsymbol{\mathcal{E}}_{\mathsf{e}}$ if

$$\frac{1}{V_d(K)} \int_K f(x)\mathrm{d}x \le \frac{1}{V_d(L)} \int_L f(x)\mathrm{d}x$$

for all convex functions $f : \mathbb{R}^d \to \mathbb{R}$, where $V_d(\cdot)$ denotes the $d$-dimensional Lebesgue measure. The latter condition implies that $K$ and $L$ share the same barycentre.

If $K_n \to K$ in the Hausdorff metric as $n \to \infty$ and $\xi_n, \xi$ are uniformly distributed on $K_n, K$, respectively, then $\xi_n \to \xi$ in $\sigma(L^p, L^q)$ for any $p \in [1, \infty]$ by the dominated convergence theorem. By Theorem 5.2.1(v), $h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K), u) \le \liminf h(\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K_n), u)$.

The continuity of the sublinear map in the Hausdorff metric follows from the next result, which we find interesting in its own right. Denote by $\mathrm{diam}(K)$ the diameter of $K$ and by $K \triangle L$ the symmetric difference of $K$ and $L$.

**Theorem 5.4.1.** *Assume that $p \in [1, \infty)$. For any two convex bodies $K$ and $L$, there exist random vectors $\xi$ and $\eta$ uniformly distributed on $K$ and $L$, respectively, such that*

$$\|\xi - \eta\|_p \le \left(\frac{V_d(K \triangle L)}{\max(V_d(L), V_d(K))}\right)^{\frac{1}{p}} \mathrm{diam}(K \cup L)\,. \tag{5.4.1}$$

*Proof.* It suffices to prove the statement for $p = 1$. Indeed,

$$\|\xi - \eta\|_p \le \mathrm{diam}(K \cup L)^{(p-1)/p} \|\xi - \eta\|_1^{1/p}.$$

Consider Monge's optimal transport problem of finding

$$\mathfrak{C}(\mu,\nu) = \inf_{T_\sharp\mu=\nu} \int_{\mathbb{R}^d} \|x - T(x)\| \mathrm{d}\mu(x)\,, \tag{5.4.2}$$

where $\mu$ and $\nu$ are the uniform distributions on $K$ and $L$, respectively, and $T_\sharp\mu$ denotes the push-forward of the measure $\mu$ by $T$. It is known from the theory of optimal mass transportation (see, e.g., [4] or [104]) that the infimum in (5.4.2) is attained on an optimal transport map $T$. Moreover, under our assumptions, [86, Th. B] yields the equivalence between Monge's transport problem and its alternative formulation by Kantorovich. Namely,

$$\mathfrak{C}(\mu,\nu) = \min_{\gamma\in\Pi(\mu,\nu)} \iint_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\| \mathrm{d}\gamma(x,y)\,,$$

where $\Pi(\mu,\nu)$ denotes the family of probability measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginals $\mu$ and $\nu$. In other words, $\mathfrak{C}(\mu,\nu)$ is the 1-Wasserstein distance between $\mu$ and $\nu$. The dual representation of Kantorovich's problem (e.g. [104, Th. 1.14]) yields that

$$\min_{\gamma\in\Pi(\mu,\nu)} \iint_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\| \mathrm{d}\gamma(x,y) = \max_{f\in\mathrm{Lip}_1} \left\{ \int_{\mathbb{R}^d} f(x)\mathrm{d}\mu(x) - \int_{\mathbb{R}^d} f(x)\mathrm{d}\nu(x) \right\}, \tag{5.4.3}$$

where $\mathrm{Lip}_1$ is the family of 1-Lipschitz functions on $\mathbb{R}^d$.

By adding a constant to $f$, one can restrict the maximisation in (5.4.3) to the set of 1-Lipschitz functions with values in $[0, \mathrm{diam}(K\cup L)]$. Then

$$\begin{aligned}
\int_{\mathbb{R}^d} f(x)\mathrm{d}\mu(x) - \int_{\mathbb{R}^d} f(x)\mathrm{d}\nu(x) &= \frac{1}{V_d(K)}\int_K f(x)\mathrm{d}x - \frac{1}{V_d(L)}\int_L f(x)\mathrm{d}x \\
&= \frac{V_d(L)-V_d(K)}{V_d(K)V_d(L)} \int_{K\cap L} f(x)dx + \frac{1}{V_d(K)}\int_{K\setminus L} f(x)\mathrm{d}x - \frac{1}{V_d(L)}\int_{L\setminus K} f(x)\mathrm{d}x \\
&\le \left(\frac{V_d(L\setminus K)}{V_d(K)}\frac{V_d(K\cap L)}{V_d(L)} + \frac{V_d(K\setminus L)}{V_d(K)}\right)\mathrm{diam}(K\cup L) \\
&\le \left(\frac{V_d(L\setminus K)}{V_d(K)} + \frac{V_d(K\setminus L)}{V_d(K)}\right)\mathrm{diam}(K\cup L) \\
&= \frac{V_d(K\triangle L)}{V_d(K)}\mathrm{diam}(K\cup L)\,.
\end{aligned}$$

Changing the order of summands, one obtains a similar bound with $V_d(K)$ replaced by $V_d(L)$, hence the result. □

**Theorem 5.4.2.** *Let* $\mathsf{e}$ *be a sublinear expectation defined on* $L^p(\mathbb{R})$ *for some* $p\in[1,\infty)$ *and having finite values. Then the map* $K\mapsto \boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ *is continuous in the Hausdorff metric.*

*Proof.* Note that the convergence of convex bodies (with nonempty interiors) in the Hausdorff metric is equivalent to their convergence in the symmetric difference metric, see [99]. If $K_n\to K$ in the Hausdorff metric, then

$\cup_n K_n$ is bounded and $\inf_n V_d(K_n)$ is strictly positive. By Theorem 5.4.1, it is possible to find a sequence of random vectors $\{\xi_n, n \geq 1\}$ such that $\xi_n$ is uniformly distributed on $K_n$ and $\xi_n$ converges in $L^p$ to a random vector $\xi$ uniformly distributed on $K$. The result follows from Theorem 5.2.1(vi). $\square$

*Example* 5.4.3. The construction of the sublinear transform can be amended by replacing the underlying sublinear expectation $\mathsf{e}$ with a (not necessarily convex) gauge function. For example, if the gauge function is a quantile, one obtains the set $D_\alpha(K)$, which is the *convex floating body* of $K$, see [8] and [97].

### 5.4.2 Ulam floating bodies

Consider the sublinear transform $K \mapsto \boldsymbol{\mathcal{E}}_\alpha(K)$ generated by the average quantile sublinear expectation $\mathsf{e}_\alpha$. Note that $\boldsymbol{\mathcal{E}}_1(K) = \{x_K\}$ is the barycentre of $K$ (the expectation of $\xi$ uniformly distributed in $K$), and $\boldsymbol{\mathcal{E}}_0(K) = K$.

The metronoid $\mathrm{M}(\mu)$ of the measure $\mu$ with density $\delta^{-1}\mathbf{1}_K$ is called the *Ulam floating body* of $K$ at level $\delta$ and is denoted by $\mathrm{M}_\delta(K)$, see [49]. This measure $\mu$ is the uniform probability distribution on $K$ scaled by $\delta^{-1}V_d(K)$. Proposition 5.3.2 yields that

$$\boldsymbol{\mathcal{E}}_\alpha(K) = \mathrm{M}_{\alpha V_d(K)}(K). \tag{5.4.4}$$

Affine equivariance of sublinear transforms implies that $M_\delta(cK) = cM_{\delta c^{-d}}(K)$. Since the uniform probability distribution on $K$ is log-concave, (5.3.4) yields a relationship between convex floating bodies of $K$ (denoted by $D_\alpha(K)$) and Ulam floating bodies, proved in [49, Th. 1.1].

The following result for $\alpha \in (0, 1/2)$ follows from Theorem 5.2.9, see also [69]. Together with (5.4.4), it implies that Ulam floating bodies can be obtained as Aumann integrals of convex floating bodies. The case $\alpha = 1/2$ follows by continuity.

**Corollary 5.4.4.** *For each origin symmetric convex body $K$ and $\alpha \in (0, 1/2]$, we have*

$$\boldsymbol{\mathcal{E}}_\alpha(K) = \frac{1}{\alpha}\int_0^\alpha D_t(K)\mathrm{d}t.$$

Hence, $\alpha\boldsymbol{\mathcal{E}}_\alpha(K)$ grows in $\alpha$ for $\alpha \in (0, 1/2]$, equivalently, the dilated Ulam floating body $tM_t(K)$ grows for $t \in (0, V_d(K)/2]$.

The next result follows from Theorem 5.3.4; it implies that Ulam floating bodies are building blocks for all sublinear transforms.

**Corollary 5.4.5.** *For each law-determined sublinear expectation $\mathsf{e}$, the corresponding sublinear transform $\boldsymbol{\mathcal{E}}_\mathsf{e}$ can be represented as*

$$\boldsymbol{\mathcal{E}}_\mathsf{e}(K) = \mathrm{conv} \bigcup_{\nu \in \mathcal{P}_\mathsf{e}} \int_{(0,1]} \boldsymbol{\mathcal{E}}_\alpha(K)\nu(\mathrm{d}\alpha), \tag{5.4.5}$$

*where* $\nu$ *runs through a family* $\mathcal{P}_{\mathsf{e}}$ *of probability measures on* $(0,1]$ *that yields the Kusuoka representation of* $\mathsf{e}$*, see* (5.1.7)*.*

It is possible to replace $\boldsymbol{\mathcal{E}}_\alpha$ with $M_{\alpha V_d(K)}$ on the right-hand side of (5.4.5). While the integration domain in (5.4.5) excludes 0, it is always possible to approximate $\boldsymbol{\mathcal{E}}_0(K) = K$ by a sequence $\boldsymbol{\mathcal{E}}_{\alpha_n}(K)$ as $\alpha_n \downarrow 0$. Thus, the Kusuoka representation can be equivalently written using probability measures on $[0,1]$.

*Example* 5.4.6. The map $K \mapsto \boldsymbol{\mathcal{E}}_\alpha(K)$ is not necessarily monotone. An easy counterexample is provided by two segments $[0,1]$ and $[0,2]$ on the line. However, the monotonicity fails even for origin symmetric convex bodies. Consider two convex bodies on the plane: $L = [-a,a] \times [-\varepsilon,\varepsilon]$ with $a+\varepsilon \le 1$ and the $\ell_1$-ball $K$. We show that for suitable values of $a$ and $\alpha$, the support function of $\boldsymbol{\mathcal{E}}_\alpha(L)$ is not smaller than the support function of $\boldsymbol{\mathcal{E}}_\alpha(K)$ in direction $u = (1,0)$. Let $\beta = \langle \xi, u\rangle$ for $\xi$ uniformly distributed in $K$. Note that $\gamma = \langle \eta, u\rangle$ is uniformly distributed on $[-a,a]$ if $\eta$ is uniform on $L$. The quantile functions are

$$q_t(\beta) = 1 - \sqrt{2(1-t)}, \quad q_t(\gamma) = (2t-1)a, \quad t \in [1/2, 1].$$

For $\alpha \in [0, 1/2]$,

$$\mathsf{e}_\alpha(\beta) = 1 - 2\sqrt{2}\alpha^{1/2}/3$$

and

$$\mathsf{e}_\alpha(\gamma) = a(1-\alpha).$$

If $\alpha = 1/2$, then $\mathsf{e}_\alpha(\beta) < \mathsf{e}_\alpha(\gamma)$ if $\frac{2}{3} < a < 1$, meaning that $\boldsymbol{\mathcal{E}}_\alpha(L)$ is not necessarily a subset of $\boldsymbol{\mathcal{E}}_\alpha(K)$.

The monotonicity of Ulam floating body transform (which easily follows from Proposition 2.1 of [49]) implies that, after normalising by volume, $\boldsymbol{\mathcal{E}}_\alpha$ becomes monotone, namely,

$$\boldsymbol{\mathcal{E}}_{\alpha/V_d(K)}(K) \subseteq \boldsymbol{\mathcal{E}}_{\alpha/V_d(L)}(L), \quad 0 \le \alpha \le V_d(K),$$

if $K \subseteq L$.

If the family $\mathcal{P}_{\mathsf{e}}$ in (5.4.5) consists of a single measure $\nu$, we obtain a convex body $\boldsymbol{\mathcal{E}}_{\int\varphi}(K)$ generated by the spectral sublinear expectation $\mathsf{e}_{\int\varphi}$, where $\varphi$ is the spectral function related to $\nu$ by (5.1.10). Recall that the maximum extension of the average quantile is a spectral sublinear expectation, see Example 5.1.8.

*Example* 5.4.7. Consider the sublinear expectation $\mathsf{e}_1^{\vee m}$ given by (5.1.14). Note that

$$\max(\langle u, \xi_1\rangle, \dots, \langle u, \xi_m\rangle) = h(P_m, u),$$

where $P_m = \operatorname{conv}(\xi_1, \dots, \xi_m)$ is the convex hull of independent copies of $\xi$. Then $\mathbf{E}h(P_m, u)$ is the support function of the expectation $\mathbf{E}P_m$ of the random polytope $P_m$, see [71, Sec. 2.1]. Therefore, $\boldsymbol{\mathcal{E}}_1^{\vee m}(K) = \mathbf{E}P_m$.

Asymptotic properties of these expected polytopes and their relation to floating bodies have been studied in [41], see also [40]. If $m = 1$, then $\boldsymbol{\mathcal{E}}_1(K) = \{x_K\}$ is the barycentre of $K$. The calculation in Example 5.1.8 yields that

$$\mathbf{E}P_m = \boldsymbol{\mathcal{E}}_1^{\vee m}(K) = m(m-1)\int_{(0,1]} \boldsymbol{\mathcal{E}}_\alpha(K)\alpha(1-\alpha)^{m-2}\mathrm{d}\alpha$$
$$= m(m-1)\int_{(0,1]} \mathrm{M}_{\alpha V_d(K)}(K)\alpha(1-\alpha)^{m-2}\mathrm{d}\alpha.$$

Hence, the expected random polytope equals the weighted integral of Ulam floating bodies.

More generally, $\boldsymbol{\mathcal{E}}_\alpha^{\vee m}(K)$ is obtained by applying (5.1.15) as follows

$$\boldsymbol{\mathcal{E}}_\alpha^{\vee m}(K) = \frac{m(m-1)}{\alpha}\int_0^{1-(1-\alpha)^{1/m}} t(1-t)^{m-2}\boldsymbol{\mathcal{E}}_t(K)\mathrm{d}t + \frac{m}{\alpha}(1-\alpha)^{(m-1)/m}(1-(1-\alpha)^{1/m})\boldsymbol{\mathcal{E}}_{1-(1-\alpha)^{1/m}}(K).$$

### 5.4.3 Centroid bodies and the expectile transform

If $\mathsf{e}_{p,a}$ is defined by (5.1.11) for $p \in [1,\infty)$, then the corresponding floating-like body $\boldsymbol{\mathcal{E}}_{p,a}(K)$ has the support function

$$h(\boldsymbol{\mathcal{E}}_{p,a}(K),u) = \langle x_K, u\rangle + a\Big(\mathbf{E}(\langle \xi - x_K, u\rangle)_+^p\Big)^{1/p}, \tag{5.4.6}$$

where $\xi$ is uniformly distributed on $K$ and $x_K = \mathbf{E}\xi$ is the barycentre of $K$.

If $K$ is origin-symmetric, then $x_K = 0$ and

$$\boldsymbol{\mathcal{E}}_{p,1}(K) = c\Gamma_p K,$$

where $c > 0$ is an explicit constant depending on $p$ and dimension and $\Gamma_p K$ is the *$L^p$-centroid body* of $K$, see [65] for $p = 1$ and [67] for general $p$. For a not necessarily origin symmetric $K$, this convex body is defined as

$$h(\Gamma_p K, u) = \left(\frac{1}{c_{d,p}V_d(K)}\int_K |\langle u, y\rangle|^p \mathrm{d}y\right)^{1/p},$$

where $c_{d,p}$ is a constant chosen to ensure that this transformation does not change the unit Euclidean ball, see [92, Eq. (10.72)]. For $p = 1$, $a = 1$ and an origin symmetric $K$,

$$\boldsymbol{\mathcal{E}}_{1,1}(K) = \frac{1}{2}\Gamma K,$$

where $\Gamma K$ is the classical *centroid body* of $K$, see [92, Eq. (10.67)] and [65]. The dual representation of $\mathsf{e}_{1,1}$ from Example 5.1.6 yields that

$$\Gamma K = 2\boldsymbol{\mathcal{E}}_{1,1}(K) = \operatorname{conv}\{\mathbf{E}(\gamma\xi) :\ \gamma \in [0,2]\}.$$

The right-hand side is the expectation of the random convex body $[0, 2\xi]$ being the segment in $\mathbb{R}^d$ with end-points at the origin and $2\xi$, see [71, Sec. 2.1].

The *asymmetric $L^p$-moment body* $\mathrm{M}_p^+ K$ introduced in [46] (see also [92, Eq. (10.76)]) has the support function proportional to

$$\int_K (\langle x, y\rangle)_+^p \mathrm{d}y.$$

Thus,

$$\boldsymbol{\mathcal{E}}_{p,a}(K) = x_K + c_1 a \,\mathrm{M}_p^+(K - x_K)$$

for a constant $c_1$ depending on $p \in [1, \infty)$ and dimension.

Corollary 5.2.5 and the dual representation of $\mathsf{e}_{p,a}$ from Example 5.1.6 (see also [33, p. 46]) yield that the asymmetric $L^p$-moment bodies with $p \in [1, \infty)$ can be represented in terms of

$$\boldsymbol{\mathcal{E}}_{p,a}(K) = x_K + a\,\mathrm{cl}\big\{\mathbf{E}((\gamma - \mathbf{E}\gamma)\xi) :\ \gamma \in L^q(\mathbb{R}_+), \|\gamma\|_q \le 1\big\}.$$

Furthermore, Corollary 5.4.5 shows that each $L^p$-centroid body of an origin symmetric $K$ equals the convex hull of a family of integrated Ulam floating bodies of $K$. This representation can be made very explicit in case $p = 1$; it follows from Theorem 5.1.4 combined with the results presented in Example 5.1.6. Namely,

$$\boldsymbol{\mathcal{E}}_{1,a}(K) = x_K + a\,\mathrm{conv} \bigcup_{t \in [0,1]} t\boldsymbol{\mathcal{E}}_t(K - x_K). \tag{5.4.7}$$

The following result specialises the above relationship for centroid bodies.

**Corollary 5.4.8.** *If $K$ is an origin symmetric convex body, then its centroid body $\Gamma K$ satisfies*

$$\Gamma K = \frac{2}{V_d(K)}\,\mathrm{conv} \bigcup_{t \in [0, V_d(K)]} t\,\mathrm{M}_t(K). \tag{5.4.8}$$

Since $K$ is origin symmetric, $t\boldsymbol{\mathcal{E}}_t(K) = \int_0^t D_s(K)\mathrm{d}s$ grows in $t \in (0, 1/2]$, see Corollary 5.4.4. Thus, the union in (5.4.8) can be reduced to $t \in [V_d(K)/2, 1]$.

*Example* 5.4.9. The definition of the Orlicz centroid bodies from [66] can be also incorporated in our setting using the sublinear expectation

$$\mathsf{e}(\beta) = \inf\{\lambda > 0 :\ \mathbf{E}\psi(\beta/\lambda) \le 1\},$$

where $\psi : \mathbb{R} \to [0, \infty)$ is a convex function with $\psi(0) = 0$ and such that $\psi$ is strictly increasing on the positive half-line or strictly decreasing on the negative half-line. This sublinear expectation is the norm of $\beta$ in the corresponding Orlicz space.

*Example* 5.4.10. Consider the expectile $\mathsf{e}_{[\tau]}$ defined in Example 5.1.7 with parameter $\tau \in (0, 1/2]$. In view of the results presented in Example 5.1.7, the corresponding floating-like body $\boldsymbol{\mathcal{E}}_{[\tau]}(K)$ can be represented as

$$\boldsymbol{\mathcal{E}}_{[\tau]}(K) = x_K + \operatorname{conv} \bigcup_{t \in [0, V_d(K)]} \frac{t(2\tau - 1)}{t(2\tau - 1) + (1 - \tau)V_d(K)} \big( \mathrm{M}_t(K) - x_K \big). \tag{5.4.9}$$

Representations (5.4.8) and (5.4.9) suggest looking at the transform of convex bodies given by

$$K \mapsto x_K + \operatorname{conv} \bigcup_{t \in [0, V_d(K)]} \psi(t) \big( \mathrm{M}_t(K) - x_K \big)$$

for a function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$. As demonstrated above, this transform relates the centroid body transform and the expectile transform to the Ulam floating body transform.

### 5.4.4 Open problems related to the sublinear transform

Several calculated examples suggest that $\boldsymbol{\mathcal{E}}_\alpha(K + L) \subseteq \boldsymbol{\mathcal{E}}_\alpha(K) + \boldsymbol{\mathcal{E}}_\alpha(L)$, and we conjecture that this is the case. It is easy to see that this holds on the line for a general sublinear transform.

It was shown in [45] that the equality of two symmetric $p$-centroid bodies for $p$ not being an even integer yields the equality of the corresponding sets. This question is open for Ulam floating bodies, see [49], not to say also for general floating-like bodies.

It is obvious that $\boldsymbol{\mathcal{E}}_{\mathsf{e}}(K)$ is a dilate of $K$ if $K$ is an ellipsoid. This question has been explored for convex floating bodies, see [109] and references therein. However, the case of Ulam floating body seems to be open, as well as the case of general sublinear transform.

There is a substantial theory of conditional (dynamic) sublinear expectations, e.g., constructed using backwards stochastic differential equations, see [79]. By applying conditional sublinear expectations to $\xi$ uniformly distributed in $K$, one comes up with stochastic processes whose values are convex bodies. Further investigation of such processes is left for future work.

# Bibliography

[1] C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Banking Finance*, 26:1505–1518, 2002.

[2] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *J. Banking Finance*, 26:1487–1503, 2002.

[3] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis.* Springer, Berlin, 2006.

[4] L. Ambrosio. Lecture notes on optimal transport problems. In *Mathematical Aspects of Evolving Interfaces (Funchal, 2000)*, volume 1812 of *Lecture Notes in Math.*, pages 1–52. Springer, Berlin, 2003.

[5] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[6] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statist. Sci.*, 5(4):403–434, 1990.

[7] R. J. Aumann. Integrals of set-valued functions. *J. Math. Anal. Appl.*, 12:1–12, 1965.

[8] I. Bárány and D. G. Larman. Convex bodies, economic cap coverings, random polytopes. *Mathematika*, 35:274–291, 1988.

[9] A. D. Barbour. Stein's method and Poisson process convergence. *J. Appl. Probab.*, 25A:175–184, 1988.

[10] A. D. Barbour. Multivariate Poisson-binomial approximation using Stein's method. In *Stein's method and applications*, volume 5 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 131–142. Singapore Univ. Press, Singapore, 2005.

[11] A. D. Barbour and T. C. Brown. Stein's method and point process approximation. *Stochastic Process. Appl.*, 43(1):9–31, 1992.

[12] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1992.

[13] A. D. Barbour and A. Xia. On Stein's factors for Poisson approximation in Wasserstein distance. *Bernoulli*, 12(6):943–954, 2006.

[14] Y. Baryshnikov and J. E. Yukich. Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.*, 15(1A):213–253, 2005.

[15] F. Bellini, B. Klar, A. Müller, and E. Rosazza Gianin. Generalized quantiles as risk measures. *Insurance Math. Econom.*, 54:41–48, 2014.

[16] F. Besau, M. Ludwig, and E. M. Werner. Weighted floating bodies and polytopal approximation. *Trans. Amer. Math. Soc.*, 370(10):7129–7148, 2018.

[17] C. Bhattacharjee and I. Molchanov. Gaussian approximation for sums of region-stabilizing scores. arXiv:2101.05103, 2021.

[18] S. Biagini and M. Frittelli. On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. In *Optimality and Risk — Modern Trends in Mathematical Finance*, pages 1–28. Springer, Berlin, 2009.

[19] S. G. Bobkov. Convex bodies and norms associated to convex measures. *Probab. Theory Related Fields*, 147:303–332, 2010.

[20] D. B. Brown. Large deviations bounds for estimating conditional value-at-risk. *Oper. Res. Lett.*, 35:722–730, 2007.

[21] T. C. Brown and A. Xia. Stein's method and birth-death processes. *Ann. Probab.*, 29(3):1373–1403, 2001.

[22] V.-E. Brunel. Concentration of the empirical level sets of Tukey's halfspace depth. *Probab. Theory Related Fields*, 173:1165–1196, 2019.

[23] L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975.

[24] L. H. Y. Chen. Stein's method: some perspectives with applications. In *Probability towards 2000 (New York, 1995)*, volume 128 of *Lect. Notes Stat.*, pages 97–122. Springer, New York, 1998.

[25] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[26] L. H. Y. Chen and A. Xia. Stein's method, Palm theory and Poisson process approximation. *Ann. Probab.*, 32(3B):2545–2569, 2004.

[27] S. N. Chiu and M. P. Quine. Central limit theory for the number of seeds in a growth model in $\mathbf{R}^d$ with inhomogeneous Poisson arrivals. *Ann. Appl. Probab.*, 7(3):802–814, 1997.

[28] S. N. Chiu and M. P. Quine. Central limit theorem for germination-growth models in $\mathbb{R}^d$ with non-Poisson locations. *Adv. in Appl. Probab.*, 33(4):751–755, 2001.

[29] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications.* John Wiley & Sons, Ltd., Chichester, third edition, 2013.

[30] L. Decreusefond, M. Schulte, and C. Thäle. Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.*, 44(3):2147–2197, 2016.

[31] L. Decreusefond and A. Vasseur. Stein's method and Papangelou intensity for Poisson or Cox process approximation. arXiv:1807.02453, 2018.

[32] P. Deheuvels and D. Pfeifer. Poisson approximations of multinomial distributions and point processes. *J. Multivariate Anal.*, 25(1):65–89, 1988.

[33] F. Delbaen. *Monetary Utility Functions.* Osaka University Press, Osaka, 2012.

[34] P. Eichelsbacher, M. Raič, and T. Schreiber. Moderate deviations for stabilizing functionals in geometric probability. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):89–128, 2015.

[35] V. Čekanavičius and P. Vellaisamy. Compound Poisson approximations in $\ell_p$-norm for sums of weakly dependent vectors. *J. Theor. Probab.*, 2020. DOI:10.1007/s10959-020-01042-9.

[36] T. Fischer. Risk capital allocation by coherent risk measures based on one-sided moments. *Insurance Math. Econom.*, 32:135–146, 2003.

[37] B. Fleury, O. Guédon, and G. Paouris. A stability result for mean width of $L_p$-centroid bodies. *Adv. Math.*, 214(2):865–877, 2007.

[38] G. B. Folland. *Real Analysis.* John Wiley & Sons, Inc., New York, 2 edition, 1999.

[39] H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time.* De Gruyter, Berlin, 2 edition, 2004.

[40] D. Fresen. A multivariate Gnedenko law of large numbers. *Ann. Probab.*, 41:3051–3080, 2013.

[41] D. J. Fresen and R. A. Vitale. Concentration of random polytopes around the expected convex hull. *Electron. Commun. Probab.*, 19:no. 59, 8, 2014.

[42] J. Galambos and S. Kotz. *Characterizations of Probability Distributions*, volume 675 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978.

[43] N. Gao, D. Leung, C. Munari, and F. Xanthos. Fatou property, representations, and extensions of law-invariant risk measures on general Orlicz spaces. *Finance Stoch.*, 22(2):395–415, 2018.

[44] R. Ger, M. Keane, and J. K. Misiewicz. On convolutions and linear combinations of pseudo-isotropic distributions. *J. Theoret. Probab.*, 13:977–995, 2000.

[45] L. Guo, G. Leng, and Y. Lin. A stability result for $p$-centroid bodies. *Bull. Korean Math. Soc.*, 55:139–148, 2018.

[46] C. Haberl and F. E. Schuster. General $L_p$ affine isoperimetric inequalities. *J. Differential Geom.*, 83:1–26, 2009.

[47] W. Hoeffding. On the distribution of the expected values of the order statistics. *Ann. Math. Statist.*, 24:93–100, 03 1953.

[48] H. Huang and B. A. Slomka. Approximations of convex bodies by measure-generated sets. *Geom. Dedicata*, 200:173–196, 2019.

[49] H. Huang, B. A. Slomka, and E. M. Werner. Ulam floating bodies. *J. Lond. Math. Soc. (2)*, 100:425–446, 2019.

[50] M. Kaina and L. Rüschendorf. On convex risk measures on $L^p$-spaces. *Math. Meth. Oper. Res.*, 69:475–495, 2009.

[51] O. Kallenberg. *Foundations of Modern Probability*. Springer-Verlag, New York, second edition, 2002.

[52] O. Kallenberg. *Random measures, theory and applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017.

[53] A. N. Kolmogorov. On the statistical theory of the crystallization of metals. *Bull. Acad. Sci. USSR, Math. Ser*, 1(3):355–359, 1937.

[54] L. Kong and I. Mizera. Quantile tomography: using quantiles with multivariate data. *Statist. Sinica*, 22:1589–1610, 2012.

[55] V. S. Koroljuk and Y. V. Borovskich. *Theory of $U$-statistics*, volume 273 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994.

[56] G. Koshevoy and K. Mosler. Zonoid trimming for multivariate distributions. *Ann. Statist.*, 25:1998–2017, 1997.

[57] S. Kusuoka. On law invariant coherent risk measures. In *Advances in Mathematical Economics, Vol. 3*, volume 3 of *Adv. Math. Econ.*, pages 83–95. Springer, Tokyo, 2001.

[58] R. Lachièze-Rey, M. Schulte, and J. E. Yukich. Normal approximation for stabilizing functionals. *Ann. Appl. Probab.*, 29(2):931–993, 2019.

[59] G. Last and M. Otto. Disagreement coupling of Gibbs processes with an application to Poisson approximation. arXiv:2104.00737, 2021.

[60] G. Last and M. Penrose. *Lectures on the Poisson Process.* Cambridge University Press, Cambridge, 2018.

[61] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.

[62] A. J. Lee. *$U$-statistics*, volume 110 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York, 1990.

[63] K. Leichtweiß. Zur Affinoberfläche konvexer Körper. *Manuscripta Math.*, 56:429–464, 1986.

[64] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[65] E. Lutwak. Centroid bodies and dual mixed volumes. *Proc. London Math. Soc. (3)*, 60(2):365–391, 1990.

[66] E. Lutwak, D. Yang, and G. Zhang. Orlicz centroid bodies. *J. Differential Geom.*, 84:365–387, 2010.

[67] E. Lutwak and G. Zhang. Blaschke–Santaló inequalities. *J. Differential Geom.*, 45:1–16, 1997.

[68] E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.

[69] M. Meyer and S. Reisner. A geometric property of the boundary of symmetric convex bodies and convexity of flotation surfaces. *Geom. Dedicata*, 37(3):327–337, 1991.

[70] I. Molchanov. Convex and star shaped sets associated with multivariate stable distributions. I. Moments and densities. *J. Multiv. Anal.*, 100:2195–2213, 2009.

[71] I. Molchanov. *Theory of Random Sets*. Springer, London, 2 edition, 2017.

[72] I. Molchanov and A. Mühlemann. Nonlinear expectations of random sets. *Finance Stoch.*, 25:5–41, 2021.

[73] S. Nagy, C. Schütt, and E. M. Werner. Halfspace depth and floating body. *Stat. Surv.*, 13:52–118, 2019.

[74] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probab. Theory Related Fields*, 145(1-2):75–118, 2009.

[75] S. Y. Novak. Poisson approximation. *Probab. Surv.*, 16:228–276, 2019.

[76] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, Ltd., Chichester, second edition, 2000.

[77] G. Paouris and E. M. Werner. Relative entropy of cone measures and $L_p$ centroid bodies. *Proc. Lond. Math. Soc. (3)*, 104(2):253–286, 2012.

[78] F. Papangelou. The conditional intensity of general point processes and an application to line processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 28:207–226, 1973/74.

[79] S. Peng. *Nonlinear Expectations and Stochastic Calculus under Uncertainty*. Springer, Berlin, 2019.

[80] M. D. Penrose and J. E. Yukich. Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.*, 11(4):1005–1041, 2001.

[81] M. D. Penrose and J. E. Yukich. Limit theory for random sequential packing and deposition. *Ann. Appl. Probab.*, 12(1):272–301, 2002.

[82] M. D. Penrose and J. E. Yukich. Weak laws of large numbers in geometric probability. *Ann. Appl. Probab.*, 13(1):277–303, 2003.

[83] C. M. Petty. Centroid surfaces. *Pacific J. Math.*, 11:1535–1547, 1961.

[84] F. Pianoforte and M. Schulte. Poisson approximation with applications to stochastic geometry. arXiv:2104.02528, 2021.

[85] F. Pianoforte and R. Turin. Multivariate Poisson and Poisson process approximations with applications to Bernoulli sums and $U$-statistics. arXiv:2105.01599, 2021.

[86] A. Pratelli. On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation. *Ann. Inst. H. Poincaré Probab. Statist.*, 43:1–13, 2007.

[87] M. Reitzner and M. Schulte. Central limit theorems for $U$-statistics of Poisson point processes. *Ann. Probab.*, 41(6):3879–3909, 2013.

[88] B. Roos. On the rate of multivariate Poisson convergence. *J. Multivariate Anal.*, 69(1):120–134, 1999.

[89] B. Roos. Poisson approximation of multivariate Poisson mixtures. *J. Appl. Probab.*, 40(2):376–390, 2003.

[90] B. Roos. Refined total variation bounds in the multivariate and compound Poisson approximation. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):337–360, 2017.

[91] N. Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:210–293, 2011.

[92] R. Schneider. *Convex Bodies. The Brunn–Minkowski Theory.* Cambridge University Press, Cambridge, 2 edition, 2014.

[93] R. Schneider and W. Weil. *Stochastic and integral geometry.* Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.

[94] T. Schreiber and J. E. Yukich. Variance asymptotics and central limit theorems for generalized growth processes with applications to convex hulls and maximal points. *Ann. Probab.*, 36(1):363–396, 2008.

[95] D. Schuhmacher. Stein's method and Poisson process approximation for a class of Wasserstein metrics. *Bernoulli*, 15(2):550–568, 2009.

[96] D. Schuhmacher and K. Stucki. Gibbs point process approximation: total variation bounds using Stein's method. *Ann. Probab.*, 42(5):1911–1951, 2014.

[97] C. Schütt and E. Werner. The convex floating body. *Math. Scand.*, 66:275–290, 1990.

[98] A. Shapiro. On Kusuoka representation of law invariant risk measures. *Math. Oper. Res.*, 38(1):142–152, 2013.

[99] G. C. Shephard and R. J. Webster. Metrics for sets of convex bodies. *Mathematika*, 12:73–88, 1965.

[100] R. L. Smith. Extreme value theory for dependent sequences via the Stein-Chen method of Poisson approximation. *Stochastic Process. Appl.*, 30(2):317–327, 1988.

[101] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602, 1972.

[102] M. Tadese and S. Drapeau. Relative bound and asymptotic comparison of expectile with respect to expected shortfall. *Insurance Math. Econom.*, 93:387–399, 2020.

[103] J. W. Tukey. Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 523–531, 1975.

[104] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, Providence, RI, 2003.

[105] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.

[106] R. A. Vitale. Expected convex hulls, order statistics, and Banach space probabilities. *Acta Appl. Math.*, 9:97–102, 1987.

[107] Y. Wang and F. Gao. Deviation inequalities for an estimator of the conditional value-at-risk. *Oper. Res. Lett.*, 38:236–239, 2010.

[108] E. Werner. The $p$-affine surface area and geometric interpretations. *Rend. Circ. Mat. Palermo (2) Suppl.*, (70, part II):367–382, 2002.

[109] E. Werner and D. Ye. On the homothety conjecture. *Indiana Univ. Math. J.*, 60(1):1–20, 2011.

[110] C. Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific Publishing Co., Inc., River Edge, NJ, 2002.