Submitted to the Annals of Applied Probability

QUANTITATIVE CLTS ON THE POISSON SPACE VIA SKOROHOD ESTIMATES AND p-POINCARÉ INEQUALITIES

BY TARA TRAUTHWEIN[1,a]

[1]*Department of Statistics, University of Oxford* , [a]*tara.trauthwein@stats.ox.ac.uk*

We establish new explicit bounds on the Gaussian approximation of Poisson functionals based on novel estimates of moments of Skorohod integrals. Combining these with the Malliavin–Stein method, we derive bounds in the Wasserstein and Kolmogorov distances whose application requires minimal moment assumptions on add-one cost operators — thereby extending the results from (Last, Peccati and Schulte, 2016). Our applications include a central limit theorem (CLT) for the Online Nearest Neighbour graph, whose validity was conjectured in (Wade, 2009; Penrose and Wade, 2009). We also apply our techniques to derive quantitative CLTs for edge functionals of the Gilbert graph, of the k-Nearest Neighbour graph and of the Radial Spanning Tree. In most cases, even the qualitative CLTs are new.

CONTENTS

1. Introduction.

MSC2020 subject classifications: Primary 60D05, 60F05; secondary 60H07, 60G55, 60G57.

Keywords and phrases: Central Limit Theorem, Gilbert Graph, Kolmogorov Distance, Malliavin Calculus, Nearest Neighbour Graph, Online Nearest Neighbour Graph, Poincaré Inequality, Poisson Process, Skorohod Integral, Stein's Method, Stochastic Geometry, Radial Spanning Tree, Wasserstein Distance.

1.1. *Overview.* The aim of this paper is to establish a new collection of probabilistic inequalities, yielding quantitative CLTs for sequences of Poisson functionals under minimal moment assumptions. We will see below that our findings substantially extend and refine the second-order Poincaré inequalities proved in [40] (see also [37, 70, 71]), and heavily rely on new moment inequalities for Skorohod integrals (see Theorem 4.2) that we believe to be of independent interest. As demonstrated in Section 5, our findings are specifically tailored, but not limited, to deriving quantitative CLTs for functionals of spatial random graphs, based on Poisson inputs, in critical or near-critical regimes. Following the publication of the preprint version of this article, our results have also been applied in [9, 10] to derive CLTs for spatial integrals of solutions to a stochastic linear wave equation with space-time pure-jump Lévy white noise, thereby illustrating that applications of our findings go beyond the context of stochastic geometry.

One prominent example dealt with in the present work is the **Online Nearest Neighbour Graph** (ONNG), devised by Berger et al. in [12] as a simplified version of the FKP model of the internet graph (see [22]). The set-up is the following: Take a Poisson measure on $\mathbb{R}^d \times [0,1]$, in such a way that each point of the measure has a spatial coordinate in $\mathbb{R}^d$ and an arrival time in $[0,1]$. Within a bounded observation window, connect each point to its nearest neighbour in space which has smaller arrival time. The resulting ONNG is a tree growing in time — a simple model for an expanding network. Other graphs of interest include the **Gilbert graph**, where two points are connected if they are close enough to one another, or nearest-neighbour type graphs like the k**-Nearest Neighbour graph** and the **Radial Spanning Tree** (see Section 5).

For graphs like these, one quantity of interest is the **total edge-length**, or more generally, the α**-power weighted total edge-length**:

$$F = \sum_{e \text{ edge}} |e|^\alpha. \tag{1.1}$$

The main question is to understand how such a sum fluctuates as the graph expands. For the ONNG, convergence to the normal law was shown by Penrose [57] in the exponent range $\alpha \in \left(0, \frac{d}{4}\right)$ and conjectured in [78, 61] for the range $\alpha \in \left[\frac{d}{4}, \frac{d}{2}\right]$. This conjecture has remained open until now: as an application of our abstract bounds, we settle it in this article by providing a quantitative central limit theorem for the centred and rescaled sum of power-weighted edge-lengths with powers $\alpha \in \left(0, \frac{d}{2}\right]$ (see Theorem 5.3).

Power-weighted edge lengths have been studied in a variety of contexts, for the minimal spanning tree in e.g. [32, 57, 34, 13], as well as for other graphs in e.g. [58, 59, 61, 69]. In applications, the idea is that large positive powers favour long edges, powers close to zero weight all edges equally, and negative powers favour short edges. Negative power weights are used in spatial statistics for inverse distance weighting, an interpolation method, as well as to compute proximity matrices for spatial autocorrelation (see e.g. [46, Chapter 13.1.4], [79, Chapter 7.4] and [31, pages 65–69]).

1.2. *Main contributions.* As announced in Section 1.1, our theoretical findings refine the main bounds in [40], where the authors proved second-order Poincaré inequalities on configuration spaces — thus extending the Gaussian second-order Poincaré estimates established in [16, 51, 76, 20] to the Poisson case. The results of [40] are based on the combination of **Stein's method** [17, 50, 54] and **Malliavin Calculus** on configuration spaces [38, 43].

The starting point of Stein's method is the fact that a real-valued random variable N is standard Gaussian if and only if

$$\mathbb{E} f(N) N = \mathbb{E} f'(N) \tag{1.2}$$

for a suitable collection of functions $f : \mathbb{R} \to \mathbb{R}$. This allows one to represent the distance between the distributions of a random variable F and a standard normal N as

$$d(F, N) = \sup_{h \in \mathcal{H}} |\mathbb{E} f_h(F)F - \mathbb{E} f_h'(F)|, \tag{1.3}$$

where $\mathcal{H}$ is a suitable collection of functions depending on the choice of distance and the function f_h is the canonical solution to the differential equation

$$f_h'(x) = x f_h(x) + h(x) - \mathbb{E} h(N). \tag{1.4}$$

The crucial idea behind the results of [40] is that, for random variables $F = F(\eta)$ depending on a Poisson measure η on a space $\mathbb{X}$, one can control quantities such as (1.3) by using integration by parts formulae involving the **add-one cost operator**

$$\mathrm{D}_x F(\eta) := F(\eta + \delta_x) - F(\eta), \tag{1.5}$$

where $x \in \mathbb{X}$ and δ_x is the Dirac measure in x, as well as its iteration $\mathrm{D}^{(2)}_{x,y} = \mathrm{D}_x \mathrm{D}_y F$. As demonstrated in [40, 37, 70, 34, 71, 72] such an approach leads to flexible bounds in the Kolmogorov and Wasserstein distances, bounds that are particularly adapted for dealing with functionals displaying a form of **geometric stabilisation** — see e.g. [62, 57, 63, 37, 34] for a discussion of this concept, as well as [32] for the first seminal contribution on the topic.

One of the shortcomings of the bounds established in [40, 37, 70, 34, 72] is that their use in concrete applications typically requires one to uniformly bound over $\mathbb{X}$ the moments of order $(4 + \epsilon)$ (with $\epsilon > 0$) of $\mathrm{D} F$ and $\mathrm{D}^{(2)} F$. Such a uniform bound is not achievable in many relevant applications, e.g. for edge functionals of the ONNG in the exponent range $\alpha \in \left[\frac{d}{4}, \frac{d}{2}\right]$, the range where the central limit theorem was conjectured to hold. We substantially extend the main results from [40] in two ways:

(i) In Theorem 3.4 we establish explicit bounds in the Kolmogorov and Wasserstein distances, whose use only requires one to uniformly bound the moments of order $2 + \epsilon$ of add-one cost operators. (See also [75] for qualitative CLTs requiring both bounds on the moments of order $2 + \epsilon$ and weak stabilisation). To motivate the reader, we will now give a simple example of a consequence of Theorem 3.4.

Let ν be a centred probability measure on $\mathbb{R}$ such that

$$c := \int_{\mathbb{R}} |u|^{2+\epsilon} \nu(du) < \infty \tag{1.6}$$

for some $\epsilon > 0$. Define $\sigma := \left(\int_{\mathbb{R}} u^2 \nu(du)\right)^{1/2}$ and let η be a Poisson measure on $\mathbb{R} \times [0, \infty)$ with intensity $\nu(du) \otimes ds$. Let $T > 0$ and define

$$F_T := \int_{\mathbb{R} \times [0,T]} u \, \eta(d(u, s)). \tag{1.7}$$

Then F_T is equal in law to $\sum_{i=1}^{N(T)} X_i$, where $X_1, X_2, \ldots$ are i.i.d random variables distributed according to ν and independent of $N(T)$, a Poisson distributed random variable with parameter T.

One can compare this to the random variable

$$G_n := \sum_{i=1}^{n} X_i, \tag{1.8}$$

where the number of points is deterministically given by n. By an extension of the classical Berry–Esseen theorem given by Petrov in [64, Theorem 6, p. 115], one has the following bound on the Kolmogorov distance between the laws of $(n\sigma^2)^{-1/2}G_n$ and a standard Gaussian N:

$$d_K\left(\frac{G_n}{\sqrt{n}\sigma}, N\right) \leqslant \frac{c}{\sigma^{2+p}} n^{-p/2}, \tag{1.9}$$

where $0 < p \leqslant \min\{\epsilon, 1\}$. For the functional F_T, our Theorem 3.4 implies

$$d_W\left(\frac{F_T}{\sqrt{T}\sigma}, N\right) \leqslant \frac{2c}{\sigma^{2+p}} T^{-p/2} \tag{1.10}$$

and

$$d_K\left(\frac{F_T}{\sqrt{T}\sigma}, N\right) \leqslant \frac{(4c)^{\frac{1}{1+q/2}}}{\sigma^2} T^{-\frac{q/2}{1+q/2}} \tag{1.11}$$

for the Wasserstein and Kolmogorov distances respectively, where $0 < p \leqslant \min\{\epsilon, 1\}$ and $0 < q \leqslant \min\{\epsilon, 2\}$. The speed of convergence in Wasserstein distance corresponds exactly to the one given by Petrov. For the Kolmogorov distance we find a slightly slower speed, which is however still converging much faster than the square root of the Wasserstein distance, which is implied by the classic estimate $d_K(., N) \leqslant 2\sqrt{d_W(., N)}$ (see e.g. [50, Remark C.2.2]). We note that for the special example of F_T as defined in (1.7), the bound (1.11) can be improved further by applying techniques tailored to the unique properties of F_T. Starting from the bound we give in Theorem 3.2, one can apply a concentration inequality similar to the one developed in [17, Lemma 3.1] and derive the following result of presumably optimal rate of convergence (see also [17, Theorem 10.6]):

$$d_K\left(\frac{F_T}{\sqrt{T}\sigma}, N\right) \leqslant \frac{(20+\sqrt{2\pi})c}{4\sigma^{2+p}} T^{-p/2} + \left(\frac{2^{1+p}c}{\sigma^{2+p}}\right)^{1/p} T^{-1/2}, \tag{1.12}$$

where $0 < p \leqslant \min\{\epsilon, 1\}$.

In Section 5, we apply Theorem 3.4 to edge-statistics of the form (1.1) of the Online Nearest Neighbour Graph in the exponent range $\alpha \in \left(0, \frac{d}{2}\right)$ and of the Gilbert graph for exponents $\alpha \in \left(-\frac{d}{2}, \infty\right)$ (extending existing results from [67]); we also deal with the k-Nearest Neighbour Graph and the Radial Spanning Tree for a general class of decreasing functions $\phi : (0, \infty) \to (0, \infty)$ applied to the edge-lengths.

In all our applications, we find the same speeds in the Wasserstein and Kolmogorov distances. Roughly speaking, a $2p$-moment bound leads to a speed of convergence of $\mathrm{Var}(F)^{1/p-1}$, where $p \in (1, 2]$. If $p = 2$, we recover the speed of order 'square root of the variance', which corresponds to the speed of convergence in the classical Berry-Esseen theorem and is often presumed to be optimal in other contexts. If the variable F is integer-valued, it has been shown e.g. in [71, inequality (2.27)] that this speed cannot be improved.

If we need to lower the moment condition by choosing $p < 2$, then the resulting speed will be slower. Comparing with the above example and setting $2p = 2 + \epsilon$, it corresponds to $t^{-d\frac{\epsilon/2}{1+\epsilon/2}}$ in both Wasserstein and Kolmogorov distances. Whether or not this speed is optimal seems to be a challenging open question. As illustrated by the example (1.7), one can do better if the second add-one cost $\mathrm{D}^{(2)}_{x,y}F$ is zero. If this term is however non-zero, it hints at a much richer and more complex structure of the underlying random variable F, indicating that a speed of $t^{-\epsilon/2}$ might be too much to expect. In general, the literature on lower bounds for convergences of quantities satisfying only sub-optimal moment conditions is rather sparse and an interesting avenue for future research.

(ii) The case of an edge functional of the type (1.1) with $\alpha = \frac{d}{2}$ for the ONNG is of a different nature. The variance contains an additional logarithmic factor (as conjectured and partially shown in [78], and fully established in Theorem 5.3) and a $2+\epsilon$ moment bound proves too strong a condition. To deal with this particular case, we develop in Theorem 3.3 an estimate of the Wasserstein distance that depends on a time parameter. Taking η to be a Poisson measure on a space $\mathbb{X}\times[0,1]$ with intensity $\lambda\otimes ds$, the estimate contains moments of $\mathbb{E}\big[\,\mathrm{D}_{(x,s)}\,F\big|\eta_{|\mathbb{X}\times[0,s)}\big]$ and $\mathbb{E}\big[\,\mathrm{D}^{(2)}_{(x,s),(y,u)}\,F\big|\eta_{|\mathbb{X}\times[0,s\vee u)}\big]$ instead of $\mathrm{D}_{(x,s)}\,F$ and $\mathrm{D}^{(2)}_{(x,s),(y,u)}\,F$. This distinction is crucial and leads to a quantitative central limit theorem in the critical case $\alpha = \frac{d}{2}$, stated in Theorem 5.3.

The underlying result making these improved estimates possible is a new inequality, stated in Theorem 4.2, providing a bound on the pth moment of a Skorohod integral $\delta(h)$ (see Section 2 for a definition), with $p\in[1,2]$. Even more generally, a bound is provided for quantities of the type $|\mathbb{E}\phi(\delta(h))|$, where $\phi:\mathbb{R}\to\mathbb{R}$ is differentiable with $(p-1)$-Hölder continuous derivative for some $p\in(1,2]$. To the best of our knowledge, no comparable inequality exists to date and this result is of independent interest. In particular, Theorem 4.2 implies generalisations of the classical **Poincaré inequality** according to which for a Poisson functional $F\in L^2(\mathbb{P}_\eta)$, it holds that

$$\mathbb{E}\left[F^2\right]-\mathbb{E}[F]^2\leqslant\mathbb{E}\int\left(\mathrm{D}_x\,F\right)^2\lambda(dx). \tag{1.13}$$

As shown in Corollary 4.3 and Remark 4.4, the Poincaré inequality remains true (up to a multiplying constant) if the exponent 2 is replaced by p, thus for $p\in[1,2]$:

$$\mathbb{E}|F|^p-|\mathbb{E}F|^p\leqslant 2^{2-p}\mathbb{E}\int_{\mathbb{X}}|\mathrm{D}_x\,F|^p\,\lambda(dx). \tag{1.14}$$

If the functional F is non-negative or centred, this inequality follows from modified log-Sobolev type inequalities shown in [15] and also [1]. We stress, however, that the bound in Theorem 4.2 is much more general and not directly deducible from [15, 1].

If we work over the time-augmented space $\mathbb{X}\times[0,1]$, it has been shown by Last and Penrose in [41, Thm. 1.5] that

$$\mathrm{Var}(F)=\mathbb{E}\int_{\mathbb{X}}\int_0^1\mathbb{E}[\mathrm{D}_{(x,t)}\,F|\eta_{|\mathbb{X}\times[0,t)}]^2dt\lambda(dx). \tag{1.15}$$

Correspondingly, the inequality in Corollary 4.3 can be refined even further by conditioning on the information up to time t:

$$\mathbb{E}|F|^p-|\mathbb{E}F|^p\leqslant 2^{2-p}\mathbb{E}\int_{\mathbb{X}}\int_0^1\left|\mathbb{E}[\mathrm{D}_{(x,t)}\,F|\eta_{|\mathbb{X}\times[0,t)}]\right|^p dt\lambda(dx). \tag{1.16}$$

This inequality is of great importance for improving estimates of both Wasserstein and Kolmogorov distances. Another consequence of Theorem 4.2 is a more technical inequality given in Corollary 4.7, this one being crucial when refining the bound on the Kolmogorov distance.

The proof of Theorem 4.2 relies on a new version of Itô's formula, shown in Theorem 4.1. In contrast to the classical Itô formula for Poisson point processes as given in [30, Theorem II.5.1], our version does not assume the process to be a semi-martingale or the integrand to be predictable. In turn, we only use the term corresponding to the integral with respect to a compensated Poisson measure. A detailed discussion of differences and similarities with the classical Itô formula and comparable results in the literature is provided in Section 4.1. We believe this result also to be of independent interest, as to the best of our knowledge no such formula for anticipative integrands and general Poisson processes exists in the literature.

REMARK 1.1. As demonstrated in the forthcoming Section 5, the principal achievement of this paper is the derivation of probabilistic bounds requiring minimal moment assumptions, that one can directly apply to a variety of models without implementing truncation or smoothing procedures. That said, it is plausible to expect that alternate bounds to some of those derived in Section 5 could be derived by combining the results of [40] with a truncation procedure similar to the ones implemented e.g. in [80, proof of Theorem 1.1] or [52, proof of Corollary 3.2]. In order to keep the length of this paper within reasonable limits, the comparison between the two approaches (in situations where both apply) will be discussed elsewhere.

Outline of the paper: In Section 2 we provide a short background on Poisson processes and Malliavin Calculus, with a more detailed account to be found in Appendix 6. In Section 3, we discuss second-order p-Poincaré inequalities, while Section 4 contains our version of Itô's formula and the new estimates for Skorohod integrals. Applications are discussed in Section 5, in particular the ONNG is dealt with in Section 5.1. The later sections contain the proofs of our principal results: Section 7 contains the proof of Theorem 4.2, and the second-order Poincaré inequalities (Theorems 3.2, 3.3 and 3.4) are shown in Section 8. All other proofs can be found in the Supplementary Material [74]: Theorem 4.1 is shown in [74, Section 1], the Corollaries 4.3 and 4.7 are proven in Sections 2 and 3 of [74] respectively, whereas the proofs for the Gilbert graph, the k-Nearest Neighbour graphs, the Radial Spanning Tree and the Online Nearest Neighbour graph can be found in Sections 4, 5, 6, and 7 of [74] respectively.

2. Framework and notations. We provide here an overview of the most relevant (in the context of this article) properties of Poisson point processes and elements of Poisson Malliavin calculus, as well as a very short introduction to Stein's method. Further definitions and properties that will be necessary for the proofs can be found in Appendix 6. We refer the reader to [38, 43] for an exhaustive discussion of the material presented below.

Poisson random measure. Let $(\mathbb{W}, \mathcal{W}, \nu)$ be a σ-finite measure space and let $\mathbf{N}_{\mathbb{W}}$ be the set of $\mathbb{N}_0 \cup \{\infty\}$-valued measures on $(\mathbb{W}, \mathcal{W})$. Define the σ-algebra $\mathcal{N}_{\mathbb{W}}$ on $\mathbf{N}_{\mathbb{W}}$ as the smallest σ-algebra such that for all $W \in \mathcal{W}$, the map $\mathbf{N}_{\mathbb{W}} \ni \xi \mapsto \xi(W) \in \mathbb{N} \cup \{\infty\}$ is measurable. If it is clear from context which space we refer to, we will write $\mathbf{N}$ and $\mathcal{N}$ instead of $\mathbf{N}_{\mathbb{W}}$ and $\mathcal{N}_{\mathbb{W}}$.

A **Poisson random measure** with intensity ν is a $(\mathbf{N}, \mathcal{N})$-valued random element χ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

- for all $W \in \mathcal{W}$ and all $k \in \mathbb{N}_0$, we have $\mathbb{P}(\chi(W) = k) = \exp(-\nu(W))\frac{\nu(W)^k}{k!}$, with the convention that $\chi(W) = \infty$ (resp. $\chi(W) = 0$) $\mathbb{P}$-a.s. if $\nu(W) = \infty$ (resp. $\nu(W) = 0$);
- for $W_1, \dots, W_n \in \mathcal{W}$ disjoint, the random variables $\chi(W_1), \dots, \chi(W_n)$ are mutually independent.

Existence and uniqueness of such a measure is shown in [43, Chapter 3]. We denote by $\mathbb{P}_\chi$ the law of χ in $(\mathbf{N}, \mathcal{N})$ and we say that χ is a $(\mathbb{W}, \nu)$-**Poisson measure**.

In view of the σ-finiteness of $(\mathbb{W}, \nu)$ and using [43, Corollary 6.5] we can and will assume throughout the paper that the Poisson measure χ is **proper**, i.e. that there exist independent random elements $X_1, X_2, \dots \in \mathbb{W}$ and an independent $\mathbb{N}_0 \cup \{\infty\}$-valued random variable κ such that $\mathbb{P}$-a.s.

$$\chi = \sum_{n=1}^{\kappa} \delta_{X_n}, \tag{2.1}$$

where δ_w is the Dirac mass at the point $w \in \mathbb{W}$. All our results only depend on the law of χ, hence this assumption has no impact on them. In this context, we will often identify χ with its support, i.e. with the random collection of points $\{X_1, X_2, \dots\}$. Note that in the case where ν has atoms, these points might appear with higher multiplicities.

Poisson functionals. For $p \geqslant 0$, denote by $L^p(\mathbb{P}_\chi)$ the set of random variables F such that there is a measurable function $f : \mathbf{N} \to \mathbb{R}$ such that $F = f(\chi)$ $\mathbb{P}$-a.s. and, if $p > 0$, such that $\mathbb{E}|F|^p < \infty$. We call F a **Poisson functional** and f a **representative** of F. All results that follow do not depend on the choice of the representative f and hence, throughout the article, we will use the symbol F indiscriminately to represent both f and F.

Add-one cost and Malliavin derivative. For a Poisson functional $F \in L^0(\mathbb{P}_\chi)$ and $w \in \mathbb{W}$, define the **add-one cost operator** of F as

$$\mathrm{D}_w F := F(\chi + \delta_w) - F(\chi), \tag{2.2}$$

and inductively set $\mathrm{D}^{(n)}_{w_1,\dots,w_n} F := \mathrm{D}_{w_n} \mathrm{D}^{(n-1)}_{w_1,\dots,w_{n-1}} F$ for $n \geqslant 1$ and $w_1, \dots, w_n \in \mathbb{W}$, where $\mathrm{D}^{(0)} F = F$ and $\mathrm{D}^{(1)} F = \mathrm{D} F$. It can be shown that, given a representative f of F, the function $(w_1, \dots, w_n, \chi) \mapsto \mathrm{D}^{(n)}_{w_1,\dots,w_n} f(\chi)$ is measurable and symmetric in $w_1, \dots, w_n$ (cf. [38, p. 5]). We denote by $\operatorname{dom}\mathrm{D}$ the set of all $F \in L^2(\mathbb{P}_\chi)$ such that

$$\mathbb{E}\int_{\mathbb{W}} (\mathrm{D}_w F)^2 \nu(dw) < \infty. \tag{2.3}$$

The restriction of the operator D to $\operatorname{dom}\mathrm{D}$ is called the **Malliavin derivative** of F (see [38, Theorem 3]). Note that $\mathrm{D} F$ is well-defined for $F \in L^0(\mathbb{P}_\eta)$, and, if $F \in L^1(\mathbb{P}_\eta)$, condition (2.3) is sufficient for F to be in $\operatorname{dom}\mathrm{D}$ (as follows from the $L^1(\mathbb{P}_\eta)$-Poincaré inequality as stated in [38, Cor. 1]).
For $F, G \in L^0(\mathbb{P}_\chi)$ and for all $w \in \mathbb{W}$, we have the following formula for the add-one cost of a product:

$$\mathrm{D}_w(FG) = (\mathrm{D}_w F)G + F(\mathrm{D}_w G) + (\mathrm{D}_w F)(\mathrm{D}_w G). \tag{2.4}$$

Chaotic decomposition. For a function $g \in L^2(\mathbb{W}^n, \nu^{(n)})$ and $n \in \mathbb{N}$, denote by $\mathrm{I}_n(g)$ the nth **Wiener–Itô integral** of g (see [38, equation (25) on page 8] for a detailed definition). Then for $F \in L^2(\mathbb{P}_\chi)$, we have the **Wiener–Itô chaos expansion**

$$F = \sum_{n=0}^{\infty} \mathrm{I}(f_n), \tag{2.5}$$

where $f_n(w_1, \dots, w_n) = \frac{1}{n!}\mathbb{E}\,\mathrm{D}^{(n)}_{w_1,\dots,w_n} F$, and $f_0 = \mathbb{E}F$ with $I_0(f_0) = f_0$, and the series converges in $L^2(\mathbb{P}_\chi)$ (cf. [38, Theorem 2]). If in addition $F \in \operatorname{dom}\mathrm{D}$, then $\mathrm{D} F \in L^2(\mathbb{P}_\chi \otimes \nu)$ and it holds $\mathbb{P}_\chi-$a.s. and for $\nu-$a.e. $w \in \mathbb{W}$ that

$$\mathrm{D}_w F = \sum_{n=1}^{\infty} n\, \mathrm{I}_{n-1}(f_n(w, .)), \tag{2.6}$$

cf. [38, Theorem 3].

Mecke's formula. Denote by $L^p(\mathbf{N} \times \mathbb{W})$ the quotient set of all measurable functions $h : \mathbf{N} \times \mathbb{W} \to \mathbb{R}$ such that, if $p > 0$, one has $\mathbb{E}\int_W |h(\chi, w)|^p \nu(dw) < \infty$.

Next, we introduce the so-called **Mecke formula** (cf. e.g. [38, formula (7)]), which holds for $h \in L^1(\mathbf{N} \times \mathbb{W})$ and for $h : \mathbf{N} \times \mathbb{W} \to [0, \infty)$ measurable:

$$\mathbb{E} \int_{\mathbb{W}} h(\chi, w)\chi(dw) = \mathbb{E} \int_{\mathbb{W}} h(\chi + \delta_w, w)\nu(dw). \tag{2.7}$$

In particular, combined with the fact that χ is assumed to be proper, this implies that for a function $h \in L^1(\mathbf{N} \times \mathbb{W})$, the integral

$$\int_{\mathbb{W}} h(\chi - \delta_w, w)\chi(dw) \tag{2.8}$$

is well-defined.

Skorohod integrals. If $h \in L^2(\mathbf{N} \times \mathbb{W})$, then for ν-a.e. $w \in \mathbb{W}$, we have $h(., w) \in L^2(\mathbb{P}_\chi)$ and thus by (2.5) we can write

$$h(\chi, w) = \sum_{n=0}^{\infty} \mathrm{I}_n(h_n(w, .)), \tag{2.9}$$

with $h_n(w, w_1, \ldots, w_n) = \frac{1}{n!}\mathbb{E}\,\mathrm{D}^{(n)}_{w_1,\ldots,w_n} h(\chi, w)$ (cf. also [38, equation (42)]). We say that $h \in \operatorname{dom}\delta$ if $h \in L^2(\mathbf{N} \times \mathbb{W})$ and

$$\sum_{n=0}^{\infty}(n+1)! \int_{\mathbb{W}^{n+1}} \tilde{h}_n^2 d\nu^{n+1} < \infty, \tag{2.10}$$

where $\tilde{h}_n$ is the symmetrisation of h_n given by

$$\tilde{h}_n(w_1, \ldots, w_{n+1}) = \frac{1}{n+1} \sum_{k=1}^{n+1} h_n(w_k, w_1, \ldots, w_{k-1}, w_{k+1}, \ldots, w_{n+1}). \tag{2.11}$$

For $h \in \operatorname{dom}\delta$ we define the **Skorohod integral** of h by

$$\delta(h) := \sum_{n=0}^{\infty} \mathrm{I}_{n+1}(h_n), \tag{2.12}$$

which converges in $L^2(\mathbb{P}_\chi)$. Note that, by [38, Theorem 5], the following condition is sufficient for $h \in L^2(\mathbf{N} \times \mathbb{W})$ to be in $\operatorname{dom}\delta$:

$$\mathbb{E} \int_{\mathbb{W}} \int_{\mathbb{W}} (\mathrm{D}_x h(\chi, y))^2 \, \nu(dx)\nu(dy) < \infty. \tag{2.13}$$

If $h \in L^1(\mathbf{N} \times \mathbb{W}) \cap \operatorname{dom}\delta$, then by [38, Theorem 6], we have $\mathbb{P}$-a.s.

$$\delta(h) = \int_{\mathbb{W}} h(\chi - \delta_w, w)\chi(dw) - \int_{\mathbb{W}} h(\chi, w)\nu(dw), \tag{2.14}$$

where the RHS is well-defined for any $h \in L^1(\mathbf{N} \times \mathbb{W})$ by (2.8).

Extension to a marked space. It will often be convenient to endow the space $\mathbb{W}$ with marks representing time. As we are only interested in the law of the Poisson functionals in question, we can always suppose that the $(\mathbb{W}, \nu)$-Poisson measure χ is the marginal of a $(\mathbb{W} \times [0,1], \nu \otimes ds)$-Poisson measure η. Indeed, $\eta(. \times [0,1])$ has the same law as χ when both are regarded as random elements in $\mathbf{N}_{\mathbb{W}}$. For a functional $F \in L^0(\mathbb{P}_\chi)$, define

$$G_F(\eta) := F(\eta(. \times [0,1])). \tag{2.15}$$

Then $G_F(\eta)$ has the same law under $\mathbb{P}_\eta$ as $F(\chi)$ under $\mathbb{P}_\chi$. Moreover, for any $(x,s) \in \mathbb{W} \times [0,1]$,

$$\mathrm{D}_{(x,s)} G_F(\eta) = \mathrm{D}_x F(\eta(. \times [0,1])), \tag{2.16}$$

which is equal in law to $\mathrm{D}_x F(\chi)$.

Predictability. We call a measurable function $h : \mathbf{N}_{\mathbb{W}\times[0,1]} \times \mathbb{W} \times [0,1] \to \mathbb{R}$ **predictable** if for all $(y,s) \in \mathbb{W} \times [0,1]$ and all $\nu \in \mathbf{N}_{\mathbb{W}\times[0,1]}$

$$h(\nu, y, s) = h(\nu_{|\mathbb{W}\times[0,s)}, y, s). \tag{2.17}$$

This definition of predictability appears e.g. in [42, equation (2.5)], where it is argued that this version of predictability is comparable to predictability in the classical sense (as defined e.g. in [30, Definition I.5.2]). It is also shown in [42, Proposition 2.4] that if $h \in L^2(\mathbf{N} \times \mathbb{W} \times [0,1])$ satisfies (2.17), then $h \in \operatorname{dom}\delta$.

Conditional expectations and Clark–Ocône formula. Let η be a $(\mathbb{W} \times [0,1], \nu \otimes ds)$-Poisson measure. Using that the measures $\eta_{|\mathbb{W}\times[0,s)}$ and $\eta_{|\mathbb{W}\times[s,1]}$ are independent, one can define a version of conditional expectation for any non-negative or integrable random variable $G \in L^0(\mathbb{P}_\eta)$ by

$$\mathbb{E}[G|\eta_{|\mathbb{W}\times[0,s)}] := \int G(\eta_{|\mathbb{W}\times[0,s)} + \xi)\Pi_s(d\xi), \tag{2.18}$$

where Π_s is the law of $\eta_{|\mathbb{W}\times[s,1]}$. If it is finite, the conditional expectation $\mathbb{E}[G|\eta_{|\mathbb{W}\times[0,s)}]$ is predictable (cf. [41, formula (2.3)] and the discussion thereafter). In particular, the following summarizes results obtained in [42] (see in particular the discussion on page 1591, equation (2.16) and Theorem 2.1):
Let $F \in L^2(\mathbb{P}_\eta)$. Then the function $h : \mathbf{N} \times \mathbb{W} \times [0,1] \to \mathbb{R}$ defined by $h(\eta, x, s) := \mathbb{E}[\mathrm{D}_{(x,s)} F|\eta_{|\mathbb{W}\times[0,s)}]$ is well-defined, finite and predictable. Moreover, $h \in L^2(\mathbf{N} \times \mathbb{W} \times [0,1]) \cap \operatorname{dom}\delta$ and the following **Clark–Ocône type formula** holds (see also [80, 29]):

$$F = \mathbb{E}F + \delta(h) \qquad \mathbb{P}-\text{a.s.} \tag{2.19}$$

This formula will be essential in the proof of Corollary 4.3.

Generic sets. Let $\mu \subset \mathbb{R}^d$ be a finite set. We say that μ is **generic** if all pairwise distances between points are distinct. We say that a set $\mu \subset \mathbb{R}^d$ is **generic with respect to points** $x_1, \ldots, x_n \in \mathbb{R}^d$, for $n \in \mathbb{N}$, if $x_1, \ldots, x_n \notin \mu$ and $\mu \cup \{x_1, \ldots, x_n\}$ is generic. Note that for compact sets $H \subset \mathbb{R}^d$, any (H, dx)-Poisson measure χ can a.s. be identified with its support and this support is a.s. generic. To simplify the presentation, we will at times adopt the notation

$$F(\mu) := F(\xi_\mu), \qquad \text{where } \xi_\mu = \sum_{x\in\mu} \delta_x \tag{2.20}$$

for a finite set $\mu \in \mathbb{R}^d$ and a measurable functional $F : \mathbf{N}_{\mathbb{R}^d} \to \mathbb{R}$. Similar notation will be used for $\mathrm{D} F(\mu)$, $\mathrm{D}^{(2)} F(\mu)$ etc.
Stein's Method. For an integrable random variable F and a standard Gaussian N, the **Wasserstein distance** between the distributions of F and N is given by

$$d_W(F, N) = \sup_{h\in\mathcal{H}} |\mathbb{E}h(F) - \mathbb{E}h(N)| \tag{2.21}$$

where $\mathcal{H}$ is the set of Lipschitz-continuous functions $h : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant $\|h\|_L \leqslant 1$. On the other hand, the **Kolmogorov distance** between the distributions of F and N is defined as

$$d_K(F,N) = \sup_{z\in\mathbb{R}} |\mathbb{P}(F \leqslant z) - \mathbb{P}(N \leqslant z)|. \tag{2.22}$$

See e.g. [50, Appendix C], and the references therein, for a discussion of the basic properties of d_W and d_K. Given $h \in \mathcal{H}$ and $\Phi(u) = \mathbb{P}(N \leqslant u)$, Stein's equation

$$f'(x) = xf(x) + h(x) - \int_{\mathbb{R}} h(u)\Phi'(u)du, \qquad x \in \mathbb{R}, \tag{2.23}$$

admits a **canonical solution** f_h defined as

$$f_h(x) := e^{x^2/2} \int_{-\infty}^{x} (h(y) - \mathbb{E}[h(N)])e^{-y^2/2}dy, \tag{2.24}$$

satisfying the two properties: (a) $f_h \in \mathcal{C}^1(\mathbb{R})$ and $\|f_h'\|_\infty \leqslant \sqrt{\frac{2}{\pi}}$ and (b) f_h' is Lipschitz-continuous with $\|f_h'\|_L \leqslant 2$, see e.g. [54, Theorem 3, Chapter 6] and the references therein. In particular, one has that

$$d_W(F,N) = \sup_{h\in\mathcal{H}} |f_h'(F) - Ff_h(F)|. \tag{2.25}$$

Similarly, for fixed $z \in \mathbb{R}$, the canonical solution f_z to Stein's equation

$$f_z'(x) = xf_z(x) + \mathbb{1}_{(-\infty,z]}(x) - \Phi(z), \qquad x \in \mathbb{R}, \tag{2.26}$$

is differentiable everywhere except at z, where it is customary to define $f_z'(z) = zf_z(z) + 1 - \Phi(z)$. One has that

- $\|f_z\|_\infty \leqslant \frac{\sqrt{2\pi}}{4}$ and $\|f_z'\|_\infty \leqslant 1$
- $|xf_z(x)| \leqslant 1$ for all $x \in \mathbb{R}$ and the function $x \mapsto xf_z(x)$ is non-decreasing for all $z \in \mathbb{R}$,

(see e.g. [17, Lemma 2.3]). We consequently have that

$$d_K(F,N) = \sup_{z\in\mathbb{R}} |f_z'(F) - Ff_z(F)|. \tag{2.27}$$

2.1. *Notation.* For $x \in \mathbb{R}^d$ and $r > 0$, we write $B^d(x,r)$ to indicate the (open) ball of centre x and radius r. For a measurable set $A \subset \mathbb{R}^d$, we denote by $|A|$ the Lebesgue measure of A, unless A is finite, in which case $|A|$ denotes the number of elements in A. We use $\overline{A}$ to denote the closure of A. Unless otherwise indicated, $\eta_{|A}$ will denote the restriction of the measure η on some space $(\mathbb{W}, \mathcal{W})$ to a mesurable set $A \subset \mathbb{W}$. Throughout this paper, $\kappa_d = |B^d(0,1)|$. We use the symbols $\wedge$ (resp. $\vee$) to denote a minimum (resp. maximum) of two elements. We shall use LHS and RHS to denote 'left hand side' and 'right hand side' and use $|x|$ to denote the Euclidean norm of $x \in \mathbb{R}^d$. The supremum norm of a function $f : \mathbb{R} \to \mathbb{R}$ is denoted by $\|f\|_\infty$. By $\stackrel{d}{=}$ and $\xrightarrow{d}$ we mean equality and convergence in distribution respectively. We use the symbol $\simeq$ (resp. $\lesssim$) if there is equality (resp. inequality) up to multiplication by a positive constant.

3. Second order p-Poincaré inequalities in Wasserstein and Kolmogorov distances. In this section we state our new bounds on the distance between the distribution of a Poisson functional and the Normal law. These bounds are called second-order p-Poincaré inequalities, following a nomenclature coined in [16], where bounds of this type were given for the first time in a Gaussian context. We make use of the well-established Malliavin–Stein method, which was pioneered in [49] in the Wiener case, used for the first time in the Poisson case in [55] and subsequently extended and developed in a wide range of articles — see the references given in [40], the survey [6], the monograph [54] and the website [48]. Related bounds in the Kolmogorov distance have been studied in various places [21, 68, 40, 34].

Recall from (2.21) and (2.22) the definitions of the Wasserstein distance $d_W(F,N)$ and the Kolmogorov distance $d_K(F,N)$ respectively, for an integrable random variable F and a standard Gaussian N.

REMARK 3.1. For the rest of this section, we fix a σ-finite measure space $(\mathbb{X},\mathcal{X},\lambda)$. Before we state our main theorems, we introduce some simplified notation to improve legibility of the following results. Write $\mathbb{Y} := \mathbb{X}\times[0,1]$ and $\bar\lambda = \lambda\otimes dt$. We introduce a partial order on $\mathbb{Y}$ by saying that $x<y$ if $x=(z,s)$, $y=(w,u)$ and $s<u$. In the following, η will be a $(\mathbb{Y},\bar\lambda)$-Poisson measure and χ will be a $(\mathbb{X},\lambda)$-Poisson measure. We will write η_x for $\eta_{|\mathbb{X}\times[0,s)}$. Integrals with respect to $\bar\lambda$ are taken over $\mathbb{Y}$ and integrals with respect to λ are taken over $\mathbb{X}$.

The next statement contains the general abstract bounds on which our analysis will rely.

THEOREM 3.2. *Let $F\in L^2(\mathbb{P}_\eta)\cap\operatorname{dom}\mathrm{D}$ such that $\mathbb{E}F=0$. Then for any $q\in[1,2]$,*

$$d_W(F,N)\leqslant\sqrt{\frac{2}{\pi}}\,\mathbb{E}\left|1-\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar\lambda(dy)\right| + 2\,\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot|\,\mathrm{D}_y F|^q\bar\lambda(dy), \tag{3.1}$$

and

$$d_K(F,N)\leqslant\mathbb{E}\left|1-\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar\lambda(dy)\right| + \sup_{z\in\mathbb{R}}\mathbb{E}\int_{\mathbb{Y}}\left|\mathbb{E}\big[\,\mathrm{D}_y F\big|\eta_y\big]\right|\mathrm{D}_y F\cdot\mathrm{D}_y(Ff_z(F)+\mathbb{1}_{\{F>z\}})\bar\lambda(dy), \tag{3.2}$$

where f_z is the canonical solution to Stein's equation in the Kolmogorov case, as defined in (2.26).

As a next step, we derive the upper bounds we use in applications. Define the following quantities for $p,q\in[1,2]$:

$$\beta_1^{(p)} := \frac{2^{2/p}\sqrt{2}}{\sqrt{\pi}}\sigma^{-2}\Bigg(\int_{\mathbb{Y}}\Bigg(\int_{\mathbb{Y}}\mathbb{E}\left[\mathbb{E}\big[|\,\mathrm{D}_y F|\big|\eta_y\big]^{2p}\right]^{\frac{1}{2p}} \cdot\mathbb{E}\left[\mathbb{E}\big[|\,\mathrm{D}^{(2)}_{x,y}F|\big|\eta_{x\vee y}\big]^{2p}\right]^{\frac{1}{2p}}\bar\lambda(dy)\Bigg)^p\bar\lambda(dx)\Bigg)^{1/p},$$

$$\beta_2^{(p)} := \frac{2^{2/p}}{\sqrt{2\pi}}\sigma^{-2}\left(\int_{\mathbb{Y}}\left(\int_{\mathbb{Y}}\mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[|\,\mathrm{D}^{(2)}_{x,y}F|\big|\eta_y\big]^{2p}\right]^{1/p}\bar\lambda(dy)\right)^p\bar\lambda(dx)\right)^{1/p},$$

$$\beta_3^{(q)} := 2\sigma^{-(q+1)}\mathbb{E}\int_{\mathbb{Y}}\left|\mathbb{E}\left[\mathrm{D}_y F|\,\eta_y\right]\right|^{q+1}\bar\lambda(dy),$$

$$\beta_4^{(q)} := 2^{3-q}\sigma^{-(q+1)}\int_{\mathbb{Y}}\int_{\mathbb{Y}}\mathbb{1}_{\{y\leqslant x\}}\mathbb{E}\left[\mathbb{E}[\mathrm{D}_y F|\eta_y]^2\right]^{1/2}\cdot\mathbb{E}\left[\left|\mathbb{E}[\mathrm{D}^{(2)}_{x,y}F|\eta_x]\right|^{2q}\right]^{1/2}\bar\lambda(dx)\bar\lambda(dy).$$

The following statement is our first bound on Wasserstein distances, expressed in terms of moments of the first and second-order add-one costs conditional on past behaviour.

THEOREM 3.3. *Let η be a $(\mathbb{Y},\bar{\lambda})$-Poisson measure and let $F \in L^2(\mathbb{P}_\eta) \cap \operatorname{dom} \mathrm{D}$. Define $\sigma := \sqrt{\operatorname{Var}(F)}$ and $\hat{F} := \sigma^{-1}(F - \mathbb{E}F)$. Let $p, q \in [1,2]$. Then*

$$d_W(\hat{F}, N) \leqslant \beta_1^{(p)} + \beta_2^{(p)} + \beta_3^{(q)} + \beta_4^{(q)}. \tag{3.3}$$

The proof can be found in Appendix 8. Now define for $p, q \in [1,2]$,

$$\gamma_1^{(p)} := \frac{2^{2/p}\sqrt{2}}{\sqrt{\pi}}\sigma^{-2}\left(\int_{\mathbb{X}}\left(\int_{\mathbb{X}} \mathbb{E}\left[|\operatorname{D}_y F|^{2p}\right]^{\frac{1}{2p}} \cdot \mathbb{E}\left[|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right]^{\frac{1}{2p}} \lambda(dy)\right)^p \lambda(dx)\right)^{1/p},$$

$$\gamma_2^{(p)} := \frac{2^{2/p}}{\sqrt{2\pi}}\sigma^{-2}\left(\int_{\mathbb{X}}\left(\int_{\mathbb{X}} \mathbb{E}\left[|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right]^{1/p} \lambda(dy)\right)^p \lambda(dx)\right)^{1/p},$$

$$\gamma_3^{(q)} := 2\sigma^{-(q+1)}\int_{\mathbb{X}} \mathbb{E}\,|\operatorname{D}_y F|^{q+1}\,\lambda(dy),$$

and

$$\gamma_4^{(p)} := \sigma^{-2}\left(4\int_{\mathbb{X}} \mathbb{E}\left[|\operatorname{D}_y F|^{2p}\right]\lambda(dy)\right)^{1/p},$$

$$\gamma_5^{(p)} := \sigma^{-2}\left(4p\int_{\mathbb{X}}\int_{\mathbb{X}} \mathbb{E}\left[|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right]\lambda(dy)\lambda(dx)\right)^{1/p},$$

$$\gamma_6^{(p)} := \sigma^{-2}\left(2^{2+p}p\int_{\mathbb{X}}\int_{\mathbb{X}} \mathbb{E}\left[|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right]^{1/2} \cdot \mathbb{E}\left[|\operatorname{D}_x F|^{2p}\right]^{1/2}\lambda(dy)\lambda(dx)\right)^{1/p},$$

$$\gamma_7^{(p)} := \sigma^{-2}\Bigg(8p\int_{\mathbb{X}}\int_{\mathbb{X}}\left(\mathbb{E}|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right)^{\frac{1}{2p}} \cdot \left(\mathbb{E}|\operatorname{D}_x F|^{2p}\right)^{\frac{1}{2p}} \cdot \left(\mathbb{E}\,|\operatorname{D}_y F|^{2p}\right)^{1-1/p}\lambda(dy)\lambda(dx)\Bigg)^{1/p}.$$

Note that the quantities $\beta_1^{(p)}, \ldots, \beta_4^{(q)}$ and $\gamma_1^{(p)}, \ldots, \gamma_7^{(p)}$ contain only expressions related to $\operatorname{D} F$ and $\operatorname{D}^{(2)} F$.

The next statement contains our main estimates on Wasserstein and Kolmogorov distances, given in terms of moments of first and second-order add-one costs (without conditioning).

THEOREM 3.4. *Let χ be a $(\mathbb{X},\lambda)$-Poisson measure and let $F \in L^2(\mathbb{P}_\chi) \cap \operatorname{dom} \mathrm{D}$. Define $\sigma := \sqrt{\operatorname{Var} F}$ and $\hat{F} := \sigma^{-1}(F - \mathbb{E}F)$. Let $p, q \in [1,2]$. Then*

$$d_W(\hat{F}, N) \leqslant \gamma_1^{(p)} + \gamma_2^{(p)} + \gamma_3^{(q)} \tag{3.4}$$

and

$$d_K(\hat{F}, N) \leqslant \sqrt{\tfrac{\pi}{2}}\gamma_1^{(p)} + \sqrt{\tfrac{\pi}{2}}\gamma_2^{(p)} + \gamma_4^{(p)} + \gamma_5^{(p)} + \gamma_6^{(p)} + \gamma_7^{(p)}. \tag{3.5}$$

REMARK 3.5. Using Hölder's inequality, one can replace the term $\gamma_7^{(p)}$ by the slightly larger but simpler bound

$$\sigma^{-2}\left(8p\int_{\mathbb{X}}\int_{\mathbb{X}}\left(\mathbb{E}|\operatorname{D}_{x,y}^{(2)} F|^{2p}\right)^{\frac{1}{2p}} \cdot \left(\mathbb{E}|\operatorname{D}_x F|^{2p}\right)^{1-\frac{1}{2p}}\lambda(dx)\lambda(dy)\right)^{1/p}, \tag{3.6}$$

where $p \in [1,2]$. We will use this bound in the proof of Theorem 5.8 in the context of the Radial Spanning Tree.

REMARK 3.6. The inequalities (3.3), (3.4) and (3.5) hold for any choice of $p, q \in [1,2]$, even though the LHSs do not depend on the parameters p and q at all. In general, it is preferable to choose p and q as large as possible to achieve faster rates of convergence. However, when chosen too large, the RHSs of the inequalities might not converge at all due to integrability and moment constraints.

REMARK 3.7 (Discussion of literature). Our results in this section are a substantial extension of [40]. The bounds given in [40, Theorems 1.1 and 1.2] contain moments of first and second-order add-one costs with exponent 4 (or even $4+\epsilon$, see [40, Proposition 1.4]). While this is a very powerful tool for showing asymptotically Gaussian behaviour, a finite 4th moment is too strong a condition for some applications, most notably for the ONNG discussed in Section 5.1. Our Theorem 3.4 reduces this condition to finite $2p$ moments, where $p \in (1,2]$, while retaining similar bounds in the case $p = 2$. In particular, [40, Theorem 6.1 and Proposition 1.4] follow from our Theorem 3.4. (See also [75] for qualitative results requiring bounds on moments of order $2p$ under weak stabilisation assumptions).

The proofs of Theorems 3.2, 3.3 and 3.4 follow in spirit the ideas from [40] and [34, Theorem 1.12] (for the Kolmogorov distance). However, we work on a space $\mathbb{X} \times [0,1]$ extended by a time component and systematically replace the operator L^{-1} by the conditional expectation $\mathbb{E}[\mathrm{D}_{(x,t)} \cdot | \eta_{|\mathbb{X}\times[0,t)}]$ (see [65] for a similar approach for Poisson measures on the real line). Moreover, we apply the inequalities established in Section 4 to achieve the improvement in the exponent. For the Wasserstein distance, we also use an improvement due to [11] to obtain the terms $\beta_3^{(q)}, \beta_4^{(q)}, \gamma_3^{(q)}$. For the Kolmogorov distance, our bound in Theorem 3.4 makes use of an improvement implemented in [34], but we remove a strong condition on F. The resulting bound is close in spirit to the one given in [40, Theorem 1.2], but with an improvement from 4th moments to $2p$th moments. Moreover, our bound does not need the term corresponding to [40, term $\gamma_3^{(q)}$, p. 670] and replaces the term corresponding to [40, term $\gamma_4^{(p)}$, p. 671] by a term depending only on the add-one cost operators of F instead of $\mathbb{E}F^4$.

In Theorem 3.3, we do not take moments of the first and second-order add-one costs of our functionals, but of their expectation conditional on ‘past behaviour’. A bound of this type is new and the distinction is crucial to solve the critical case of the ONNG (see Theorem 5.3). As of now, such a bound is only available in the Wasserstein distance.

4. Ancillary results: new estimates for Skorohod integrals.

4.1. *A version of Itô’s formula.* We start this section by giving a version of Itô’s formula for Poisson integrals with anticipative integrands. This is a crucial ingredient for the proof of the new estimates given in Theorem 4.2. In the following, we will take η to be a $(\mathbb{X} \times [0,1], \mathcal{X} \otimes \mathcal{B}([0,1]), \lambda(dx) \otimes ds)$-Poisson measure, where $(\mathbb{X}, \mathcal{X}, \lambda)$ is a σ-finite measure space.

THEOREM 4.1 (Itô’s formula for non-adapted integrands). *Let $h, g \in L^1(\mathbf{N} \times \mathbb{X} \times [0,1])$ and assume h to be bounded. Let $X_0 \in \mathbb{R}$. For $t \in [0,1]$, define*

$$X_t(\eta) := X_0 + \int_{\mathbb{X}\times[0,t]} h(\eta - \delta_{(y,s)}, y, s)\eta(d(y,s)) - \int_{\mathbb{X}} \int_0^t g(\eta, y, s) ds \lambda(dy). \tag{4.1}$$

Then the process $(X_t)_{t\in[0,1]}$ *is well-defined and* $\mathbb{P}$*-a.s. càdlàg. Let* $\phi \in \mathcal{C}^1(\mathbb{R})$*. Then, for all* $t \in [0,1]$*,*

$$\phi(X_t) = \phi(X_0) + \int_{\mathbb{X}\times[0,t]} \left(\phi\left(X_{s-} + h(\eta - \delta_{(y,s)}, y, s)\right) - \phi(X_{s-})\right)\eta(d(y,s)) \\ - \int_{\mathbb{X}} \int_0^t \phi'(X_s) g(\eta, y, s) ds \lambda(dy) \qquad \mathbb{P}\text{-a.s.} \tag{4.2}$$

and the quantities in (4.2) *are well-defined.*

In the next three items we compare our version of Itô's formula (in the case $h = g$) with the classical one given in [30, Theorem II.5.1].

1. The main difference between (4.2) and [30, Thm. II.5.1] consists in the fact that we do not assume the integrand h to be predictable. There exist Itô-type formulae for anticipative integrands in various settings, e.g. in the Wiener case in [3] and [53]) and for pure jump and general Lévy processes in [19] and [2] respectively. To the best of our knowledge, our setting of a general Poisson point process is new.
2. Assume h to be predictable in the sense of (2.17). It follows that $h(\eta - \delta_{(y,s)}, y, s) = h(\eta, y, s)$ for all $(y,s) \in \mathbb{X} \times [0,1]$. For $\phi \in \mathcal{C}^2(\mathbb{R})$, formula (4.2) is now roughly equivalent to the Itô formula given in [30, Theorem II.5.1] in the special case where the semimartingale in the statement of [30, Theorem II.5.1] has the following properties:
   - the point process in question is a Poisson point process;
   - the only non-zero part is the one with respect to the compensated Poisson measure;
   - the integrand h is both in L^2 and in L^1.
3. Our setting is thus both more general ($\phi \in \mathcal{C}^1(\mathbb{R})$ and h anticipative) and more restrictive ($h \in L^1 \cap L^2$ instead of $h \in L^{2,\text{loc}}$ and the Gaussian, finite variation and non-compensated Poisson terms are zero) than the one given by Ikeda and Watanabe. The proof of our result relies however on the same ideas as the proof of [30, Theorem II.5.1].

4.2. *Moment Inequalities.* In this section, we present a number of functional inequalities that are of independent interest and also crucial to the improved bounds on Wasserstein and Kolmogorov distances presented in earlier sections.

To the best of our knowledge, Theorem 4.2 is the first bound of its kind on functionals of general Poisson–Skorohod integrals. Partial results are known in the particular case where h is predictable, see Corollary 4.3 and the discussion thereafter. In particular, Theorem 4.2 below contains the first general estimate in terms of add-one costs for p-moments of the Skorohod integral, where $p \in [1,2]$, the cases $p = 1$ and $p = 2$ being the only ones known. See also [39].

In the special case $\phi(x) = x^2$, the theorem below follows immediately from the isometry relation reported in formula (6.4) of Appendix 6.

THEOREM 4.2. *Let* $h \in L^2(\mathbf{N} \times \mathbb{X} \times [0,1])$ *satisfy* (2.13)*. Let* $\phi : \mathbb{R} \to \mathbb{R}$ *be a differentiable function with* $(p-1)$*-Hölder continuous derivative, for some* $p \in (1,2]$ *and assume that* $\phi(0) = 0$*. Then*

$$|\mathbb{E}\phi(\delta(h))| \leqslant \frac{c_\phi}{p} \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds\, |h(\eta, y, s)|^p \\ + c_\phi \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds \int_{\mathbb{X}} \lambda(dx) \int_0^s dt\, \left|\mathrm{D}_{(y,s)} h(\eta, x, t)\right| \cdot \left|\mathrm{D}_{(x,t)} h(\eta, y, s)\right|^{p-1} \\ + 2c_\phi \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds \int_{\mathbb{X}} \lambda(dx) \int_0^s dt\, \left|\mathrm{D}_{(y,s)} h(\eta, x, t)\right| \cdot |h(\eta, y, s)|^{p-1}, \tag{4.3}$$

where c_ϕ is the Hölder coefficient of ϕ'. In particular, this inequality holds with $\phi(x) = |x|^p$ and $c_\phi = p2^{2-p}$, for $p \in [1,2]$.

Our first corollary is a version of the above inequality for predictable functions h and contains a generalisation of the classical Poincaré inequality.

COROLLARY 4.3. *Let $h \in L^2(\mathbf{N} \times \mathbb{X} \times [0,1])$ be predictable in the sense of (2.17). Let $\phi : \mathbb{R} \to \mathbb{R}$ be a differentiable function with $(p-1)$-Hölder continuous derivative, for some $p \in (1,2]$. Assume $\phi(0) = 0$. Then*

$$|\mathbb{E}\phi(\delta(h))| \leqslant \frac{c_\phi}{p}\mathbb{E}\int_{\mathbb{X}}\int_0^1 |h(\eta,y,s)|^p ds\lambda(dy). \tag{4.4}$$

Moreover, for $F \in L^2(\mathbb{P}_\eta)$ and $p \in [1,2]$,

$$\mathbb{E}|F|^p - |\mathbb{E}F|^p \leqslant 2^{2-p}\mathbb{E}\int_{\mathbb{X}}\int_0^1 \left|\mathbb{E}[\mathrm{D}_{(x,t)} F|\eta_{|\mathbb{X}\times[0,t)}]\right|^p dt\lambda(dx). \tag{4.5}$$

REMARK 4.4. 1. We can extend inequality (4.5) to $F \in L^1(\mathbb{P}_\eta)$ at the cost of introducing an additional absolute value on the RHS:

$$\mathbb{E}|F|^p - |\mathbb{E}F|^p \leqslant 2^{2-p}\mathbb{E}\int_{\mathbb{X}}\int_0^1 \mathbb{E}[|\,\mathrm{D}_{(x,t)} F||\eta_{|\mathbb{X}\times[0,t)}]^p dt\lambda(dx). \tag{4.6}$$

This can be seen easily by approximating F by $F_n := (F \wedge n) \vee (-n)$ and using monotone and dominated convergence.

2. When removing the conditional expectation in (4.5) using Jensen's inequality, the inequality can be extended to functionals $G \in L^1(\mathbb{P}_\chi)$, where χ is a $(\mathbb{X},\lambda)$-Poisson measure without time component. Indeed, as discussed in Section 2, the marginal $\eta(.\times[0,1])$ has the same law as χ, which means one can see G as a functional on $\mathbf{N}_{\mathbb{X}\times[0,1]}$. We have then

$$\mathbb{E}|G|^p - |\mathbb{E}G|^p \leqslant 2^{2-p}\mathbb{E}\int_{\mathbb{X}} |\mathrm{D}_x G|^p \,\lambda(dx). \tag{4.7}$$

REMARK 4.5 (Literature review). The proof of Theorem 4.2 relies on a combination of the Clark–Ocône type representation result (2.19) and the version of Itô's formula given in Theorem 4.1. This method of combining a Clark–Ocône result with Itô formulae to deduce functional inequalities has been applied before in various settings, e.g. in [80] and [15], where it was used to deduce a modified log-Sobolev inequality and Φ-Sobolev inequalities respectively.

Inequalities (4.4), (4.5) and (4.7) can be seen as part of a larger family of functional inequalities on the Poisson space. The first to mention is the classical Poincaré inequality, given e.g. in [38, Theorem 10] (see also [28, Cor. 4.4] for a very early appearance of this inequality). Our inequality extends the classical one, which is (4.7) in the case $p = 2$. Another well-known inequality is the modified log-Sobolev inequality shown in [80] (see also [4]). It is extended in [15, equation (5.10)] to the so-called Φ-Sobolev inequalities, which in the case $\Phi(x) = x^p$, imply (4.7) when $F \geqslant 0$ or $\mathbb{E}F = 0$. Similarly, the Beckner type p inequalities discussed in [1, Section 4.6] imply (4.7) when $F \geqslant 0$ or $\mathbb{E}F = 0$, albeit with a worse constant. [81, Theorem 3.3.2] gives a version of (4.4) for p-norms in martingale type p Banach spaces. Although we did not check the details, it is reasonable to assume that one can deduce (4.4) in the case $\phi(x) = |x|^p$ from such a result when applied to $\mathbb{R}^d$.

REMARK 4.6 (Comparison with the Gaussian case and extensions when $p \geqslant 2$).

1. Inequality (4.7) for $p < 2$ does not hold for functionals of Gaussian random measures, as can be seen by taking $G = W_t$ and letting $t \to 0$, with W a standard Brownian motion. This is in contrast with the classical Poincaré inequality ($p = 2$) which holds in both Gaussian and Poisson settings.
2. Inequality (4.7) (and hence also (4.5)) is false in general for $p > 2$. Indeed, consider $(\mathbb{X}, \lambda) = (\mathbb{R}^d, dx)$ and $G = \chi(A) - \lambda(A)$ for some measurable $A \subset \mathbb{R}^d$. Then $\mathbb{E}G = 0$, $\mathbb{E}G^2 = \lambda(A)$ and $\mathrm{D}_x G = \mathbb{1}_A(x)$. On the LHS we have $\mathbb{E}|G|^p \geqslant (\mathbb{E}G^2)^{p/2} = \lambda(A)^{p/2}$ by Jensen's inequality and on the RHS
$$\int_{\mathbb{X}} \mathbb{E}|\mathrm{D}_x G|^p \lambda(dx) = \lambda(A). \tag{4.8}$$
However, since $p > 2$, we have $\lambda(A)^{p/2} \gg \lambda(A)$ for $\lambda(A)$ large enough. Hence the inequality fails for any multiplying constant.
3. Moment estimates for $p \geqslant 2$ are given in [26, Theorem 4.1] and [1, Proposition 4.20]. The RHSs of these inequalities involve related, but different quantities.

The versatility of Theorem 4.2 can be appreciated when considering the following corollary, which will be crucial in finding a bound on the Kolmogorov distance.

COROLLARY 4.7. *Let $h \in L^1(\mathbf{N} \times \mathbb{X} \times [0,1])$ and $G \in L^0(\mathbb{P}_\eta)$ bounded by a constant $c_G > 0$. Then for any $p \in [1, 2]$,*

$$\begin{aligned}
&\left|\mathbb{E}\int_{\mathbb{X}}\int_0^1 h(\eta, x, s)\,\mathrm{D}_{(x,s)} G ds\lambda(dx)\right| \\
&\leqslant c_G\Bigg(2^{2-p}\mathbb{E}\int_{\mathbb{X}}\int_0^1 |h(\eta,x,s)|^p ds\lambda(dx) \\
&\qquad + p2^{2-p}\mathbb{E}\int_{\mathbb{X}}\int_0^1\int_{\mathbb{X}}\int_0^1 |\mathrm{D}_{(x,s)} h(\eta,y,u)|^p ds\lambda(dx)du\lambda(dy) \\
&\qquad + p2^{3-p}\int_{\mathbb{X}}\int_0^1\int_{\mathbb{X}}\int_0^1 \mathbb{1}_{\{s<u\}}\mathbb{E}[|\mathrm{D}_{(y,u)} h(\eta,x,s)|^p]^{1/p} \\
&\qquad\qquad \mathbb{E}[|h(\eta,y,u)|^p]^{1-1/p} ds\lambda(dx)du\lambda(dy)\Bigg)^{1/p}.
\end{aligned} \tag{4.9}$$

REMARK 4.8. Provided that we upper bound the indicator in the third term on the RHS of (4.9) by 1, this inequality can be extended to a space $\mathbb{X}$ without time component.

5. Applications. In this section, we look at four types of graphs built on Poisson measures and assess the speeds of convergence to Normality of α-power-weighted edge-lengths such as (1.1). As was found in previous work [40, 69], we find for certain ranges of exponents α that the speed is given by $t^{-d/2}$, which corresponds to the order of the square root of the variance. This is the presumably optimal speed corresponding to the one in the classical Berry–Esseen theorem (see e.g. [64, Theorem 4, p. 111]). Beyond a certain threshold, we find a slower speed of convergence that depends on α. Generally speaking, a $2p$th moment integrability of the first and second-order add-one costs of the functionals leads to a speed of convergence of $t^{-d(1-1/p)}$. Whether this speed is optimal or not is an open question.

5.1. *Online Nearest Neighbour Graph.* Let $\mu \subset \mathbb{R}^d \times [0,1]$ be a finite set such that the projection of μ onto $\mathbb{R}^d$ is generic and does not contain any multiplicities and the projections onto $[0,1]$ are distinct. The ONNG on μ is an (undirected) graph in $\mathbb{R}^d$ constructed as follows:

- Vertices are given by $\{x \in \mathbb{R}^d : (x,s) \in \mu\}$
- Let $(x,s) \in \mu$. If $\mu \cap \big(\mathbb{R}^d \times [0,s)\big)$ is non-empty, then the online nearest neighbour of (x,s) is given by the point $(z,u) \in \mu \cap \big(\mathbb{R}^d \times [0,s)\big)$ which minimises $|x-z|$. In this case there is an edge from x to z and we denote this event by $\{(x,s) \to (z,u) \text{ in } \mu\}$.

For a point $(x,s) \in \mu$, the coordinate s can be seen as the arrival time of the point $x \in \mathbb{R}^d$, or its mark. Any point $(x,s) \in \mu$ has exactly one online nearest neighbour, except for the point in μ whose mark is minimal, which has none. Even though the graph is undirected, we think of arrows going from a point to its nearest neighbour, as this simplifies the discussion.

The ONNG is a relatively simple model for networks growing in time. Already mentioned in [73], the Online Nearest Neighbour graph came to general attention in [12], where it was presented as a simplified version of the FKP model developed in [22], used to model the internet graph. The name of the graph was coined in [57], where the martingale method is used to show central limit theorems for stabilising random systems satisfying a 4th moment condition.

To define our functional of interest, let

$$e(x,s,\mu) := \begin{cases} \inf\{|x-z| : (z,u) \in \mu \cap \big(\mathbb{R}^d \times [0,s)\big)\}, & \text{if } \mu \cap \big(\mathbb{R}^d \times [0,s)\big) \neq \emptyset \\ 0, & \text{otherwise,} \end{cases} \tag{5.1}$$

for $(x,s) \in \mu$. This is the length of the edge from x to its online nearest neighbour if there is one, and zero otherwise. Note that one can find a unique online nearest neighbour in μ for any point $(x,s) \in \mathbb{R}^d \times [0,1]$ such that the time coordinate s and the position x do not occur in μ. For convenience, we shall extend the above definitions to any such $(x,s) \in \mathbb{R}^d \times [0,1]$ and tacitly adopt the corresponding notation.

We will be studying the sums of power-weighted edge-lengths defined as follows: for $\alpha > 0$, let

$$F^{(\alpha)}(\mu) := \sum_{(x,s)\in\mu} e(x,s,\mu)^{\alpha}. \tag{5.2}$$

Note that here we make use of the convention explained in Section 2 to identify a set of points μ with the point measure whose support is given by μ.

Let η be a Poisson measure on $\mathbb{R}^d \times [0,1]$ with Lebesgue intensity. Let $H \subset \mathbb{R}^d$ be a convex body. For $t \geqslant 1$, define

$$F_t^{(\alpha)} := F^{(\alpha)}(\eta_{|tH\times[0,1]}). \tag{5.3}$$

For this functional, a CLT is shown in [57]:

THEOREM 5.1 ([57, Theorem 3.6]). *For $0 \leqslant \alpha < \frac{d}{4}$, there is a constant $\sigma_{\alpha,d} > 0$ such that as $t \to \infty$,*

$$t^{-d}\operatorname{Var}\left(F_t^{(\alpha)}\right) \longrightarrow \sigma_{\alpha,d} \quad \textit{and} \quad \frac{F_t^{(\alpha)} - \mathbb{E}F_t^{(\alpha)}}{t^{d/2}} \xrightarrow{d} \mathcal{N}(0,\sigma_{\alpha,d}). \tag{5.4}$$

A quantitative counterpart to this result is shown in [34]. Our Theorem 5.3 provides a speed of convergence that is faster than the one given in [34].

In [57, 60, 78], results similar to Theorem 5.1 were conjectured to hold for $\alpha \in \left[\frac{d}{4}, \frac{d}{2}\right]$. In particular, part of the Conjectures 2.1 and 2.2. in [78] states

CONJECTURE 5.2 ([78]). *For $\alpha \in \left[\frac{d}{4}, \frac{d}{2}\right)$, there is a constant $\sigma_{\alpha,d} > 0$ such that* (5.4) *holds.*

For $\alpha = \frac{d}{2}$, there is a constant $\sigma_d > 0$ such that

$$\log(t)^{-1} t^{-d} \operatorname{Var}\left(F_t^{(\alpha)}\right) \longrightarrow \sigma_d \quad \text{and} \quad \frac{F_t^{(\alpha)} - \mathbb{E}F_t^{(\alpha)}}{\log(t)^{1/2} t^{d/2}} \xrightarrow{d} \mathcal{N}(0, \sigma_d). \tag{5.5}$$

Our forthcoming Theorem 5.3 confirms this conjecture by giving quantitative central limit theorems for $\alpha \in \left(0, \frac{d}{2}\right]$ and upper and lower bounds for the variances that match the conjectured orders. Upper bounds of the conjectured orders were already given in [78, Theorem 2.1] for the variances involved. They are shown for an ONNG built on n uniformly distributed random variables and the corresponding result for the Poisson version follows by Poissonisation. For the sake of completeness, we give purely Poissonian proofs of the upper bounds in the Supplementary Material [74] (see [74, Proposition 7.14]), following however a similar strategy as in [78]. A law of large numbers is shown in [77] and the case $\alpha > \frac{d}{2}$ is discussed in [60] (especially for $d = 1$) and in [78], where it is shown that a limit exists in this case, but is non-Gaussian for $\alpha > d$. For more related results we refer to the survey [61] (for results up to 2010) and to the paper [44].

THEOREM 5.3. *For $0 < \alpha < \frac{d}{2}$, and for every $1 < p < \frac{d}{2\alpha}$ such that $p \leqslant 2$, there is a constant $c_1 > 0$ such that for all $t \geqslant 1$ large enough*

$$\max\left\{ d_W\left(\frac{F_t^{(\alpha)} - \mathbb{E}F_t^{(\alpha)}}{\sqrt{\operatorname{Var}\left(F_t^{(\alpha)}\right)}}, N\right), d_K\left(\frac{F_t^{(\alpha)} - \mathbb{E}F_t^{(\alpha)}}{\sqrt{\operatorname{Var}\left(F_t^{(\alpha)}\right)}}, N\right)\right\} \leqslant c_1 t^{-d\left(1-\frac{1}{p}\right)}, \tag{5.6}$$

where N denotes a standard normal random variable. Moreover, there are constants $c_2, C_2 > 0$ such that for all $t \geqslant 1$ large enough

$$c_2 t^d < \operatorname{Var}(F_t^{(\alpha)}) < C_2 t^d. \tag{5.7}$$

For $\alpha = \frac{d}{2}$, there is a constant $c_3 > 0$ such that for all $t \geqslant 1$ large enough

$$d_W\left(\frac{F_t^{(d/2)} - \mathbb{E}F_t^{(d/2)}}{\sqrt{\operatorname{Var}\left(F_t^{(d/2)}\right)}}, N\right) \leqslant c_3 \log(t)^{-1}. \tag{5.8}$$

Moreover, there are constants $c_4, C_4 > 0$ such that for all $t \geqslant 1$ large enough

$$c_4 t^d \log(t^d) < \operatorname{Var}(F_t^{(d/2)}) < C_4 t^d \log(t^d). \tag{5.9}$$

The constants $c_1, c_2, C_2, c_3, c_4, C_4$ may depend on H, α, d and p.

Note that, if one chooses $\alpha = 0$, then $F_t^{(\alpha)} = \eta(tH \times [0,1]) - 1$, and hence

$$\frac{F_t^{(\alpha)} - \mathbb{E}F_t^{(\alpha)}}{\operatorname{Var}(F_t^{(\alpha)})} \stackrel{d}{=} \frac{X_t - t^d|H|}{\sqrt{t^d|H|}}, \tag{5.10}$$

where X_t is a Poisson distributed random variable with parameter $t^d|H|$. One can derive a CLT for this quantity from e.g. (1.9) and (1.12). The speed of convergence is of the order $t^{-d/2}$, which corresponds to the square root of the order of the variance. This speed is in fact optimal, as follows from e.g. [71, inequality (2.27)].

In the case $0 < \alpha < \frac{d}{4}$, Theorem 5.3 also gives a speed of convergence of $t^{-d/2}$. As such, our bound is a natural extension of the optimal speed achieved when $\alpha = 0$.

In the case $\frac{d}{4} \leqslant \alpha < \frac{d}{2}$, the choice of p in Theorem 5.3 depends on α and we can achieve any speed strictly slower than $t^{-(d-2\alpha)}$. Note however that the constant c_1 implicitly contains the factor $(d - 2p\alpha)^{-1}$, and hence it will explode when $p \to \frac{d}{2\alpha}$.

5.2. *Gilbert Graph.* For a finite set $\mu \subset \mathbb{R}^d$ and a real number $\epsilon > 0$, the Gilbert graph $G(\mu, \epsilon)$ has vertex set μ and an edge between $x, y \in \mu$, $x \neq y$ if and only if $|x-y| \leqslant \epsilon$. To construct our functional of interest, we consider

- $W \subset \mathbb{R}^d$ a convex body;
- for every $t > 0$, we take η^t a $(W, t\,dx)$-Poisson measure;
- $(\epsilon_t)_{t>0}$ a sequence of positive real numbers s.t. $\epsilon_t \to 0$ as $t \to \infty$.

Then for $\alpha \in \mathbb{R}$, define

$$L_t^{(\alpha)} := \sum_{e \in G(\eta^t, \epsilon_t)} |e|^\alpha = \frac{1}{2} \sum_{x,y \in \eta^t, x \neq y} \mathbb{1}_{\{|x-y| \leqslant \epsilon_t\}} |x-y|^\alpha, \tag{5.11}$$

where e denote the edges of the graph and $|e|$ their length.

Define

$$\hat{L}_t^{(\alpha)} := \frac{L_t^{(\alpha)} - \mathbb{E} L_t^{(\alpha)}}{\sqrt{\operatorname{Var}\left(L_t^{(\alpha)}\right)}}. \tag{5.12}$$

REMARK 5.4. For the sake of continuity with the article [67], we use the convention that the intensity $t\,dx$ of η^t grows and the observation window W stays constant. In the other applications presented in this article, we keep the intensity constant and instead let the observation window grow as tW. Note that one can pass from one setting to the other by a simple rescaling. Indeed, consider $\tilde{\eta}$ a Poisson measure on $\mathbb{R}^d$ with Lebesgue intensity and for $s \geqslant 1$, construct a Gilbert graph on $\tilde{\eta}_{|sH}$ by connecting two points $x \neq y \in \tilde{\eta}_{|sH}$ if and only if $|x-y| \leqslant \tilde{\epsilon}_s$. For $\alpha \in \mathbb{R}$, let

$$F_s^{(\alpha)} := \frac{1}{2} \sum_{x,y \in \tilde{\eta}_{|sW}, x \neq y} \mathbb{1}_{\{|x-y| \leqslant \tilde{\epsilon}_s\}} |x-y|^\alpha. \tag{5.13}$$

Then $F_t^{(\alpha)}$ is equal in law to $s^\alpha L_{s^d}^{(\alpha)}$ with $\epsilon_{s^d} = s^{-1} \tilde{\epsilon}_s$. The central limit theorem for $F_t^{(\alpha)}$ can be deduced from the one for $L_t^{(\alpha)}$.

The first mention of the Gilbert graph was by Gilbert in [23], in dimension $d = 2$. It has been treated in many works under various names: geometric or proximity graph, interval graph (when $d = 1$) or disk graph (when $d = 2$). The book [56] provides a vast background and literature review and we also refer to [33, 35] for central limit theorems of generalisations of the Gilbert graph and [66] for a quantitative CLT on a sum of weighted edge-lengths. For a comprehensive overview of the Gilbert graph in the context of U-statistics, see [36], especially Section 4.3. See also [45, 47, 27, 14, 18, 25]. In [67], the authors give a complete picture of the asymptotic behaviour of $\hat{L}_t^{(\alpha)}$ for $\alpha \in \mathbb{R}$. In particular, they show that for $\alpha > -\frac{d}{2}$, the quantity $\hat{L}_t^{(\alpha)}$ converges in distribution to a standard Gaussian as $t \to \infty$, provided that $t^2 \epsilon_t^2 \to \infty$. They also give a quantitative bound on the speed of convergence in Kolmogorov distance in the case $\alpha > -\frac{d}{4}$. As an application of our estimates, we recover this speed of convergence below and extend to the case $-\frac{d}{2} < \alpha \leqslant -\frac{d}{4}$. The authors of [67] show that CLTs hold also for $-d < \alpha \leqslant -\frac{d}{2}$ with different rescalings; however, establishing corresponding speeds of convergence in this range is still an open problem.

THEOREM 5.5. *Let $\alpha > -\frac{d}{2}$ and assume that $t^2 \epsilon_t^d \to \infty$ as $t \to \infty$. Then for $t \geqslant 1$ large enough*

- *if* $\alpha > -\frac{d}{4}$*, there is a constant* $c_1 > 0$ *such that*

$$\max\left\{d_W\left(\hat{L}_t^{(\alpha)}, N\right), d_K\left(\hat{L}_t^{(\alpha)}, N\right)\right\} \leqslant c_1\left(t^{-1/2} \vee (t^2\epsilon_t^d)^{-1/2}\right). \tag{5.14}$$

- *if* $-\frac{d}{2} < \alpha \leqslant -\frac{d}{4}$*, then for any* $1 < p < -\frac{d}{2\alpha}$*, there is a constant* $c_2 > 0$ *such that*

$$\max\left\{d_W\left(\hat{L}_t^{(\alpha)}, N\right), d_K\left(\hat{L}_t^{(\alpha)}, N\right)\right\} \leqslant c_2\left(t^{-1+1/p} \vee (t^2\epsilon_t^d)^{-1+1/p}\right). \tag{5.15}$$

REMARK 5.6. A careful inspection of the bounds applied to $\gamma_3^{(q)}$ in the proof of Theorem 5.5 reveals that in the sparse regime ($t\epsilon_t^d \to 0$) when $-\frac{d}{2} < \alpha \leqslant -\frac{d}{4}$, a slightly improved rate can be found for the Wasserstein distance. Indeed, for any $1 < p < -\frac{d}{2\alpha}$ and any $0 < r < -\frac{d}{\alpha} - 2$, there is a constant $c > 0$ such that

$$d_W(\hat{L}_t^{(\alpha)}, N) \leqslant c\left(t^{-1+1/p} \vee (t^2\epsilon_t^d)^{-r/2}\right). \tag{5.16}$$

Since $\frac{r}{2} \in \left(0, -\frac{d}{2\alpha} - 1\right)$ and $1 - \frac{1}{p} \in \left(0, 1 + \frac{2\alpha}{d}\right)$ and $1 + \frac{2\alpha}{d} < -\frac{d}{2\alpha} - 1$, one can choose $\frac{r}{2} > 1 - \frac{1}{p}$. This then gives a slightly faster convergence rate. As an illustrating example, consider the case where $\epsilon_t^d = t^{-\theta}$ with $1 < \theta < 2$. Then $t\epsilon_t^d \to 0$ and $t^2\epsilon_t^d \to \infty$. Theorem 5.5 provides the rate of convergence $t^{(-1+\frac{1}{p})(2-\theta)}$, and by following this strategy it can be improved to $t^{-1+\frac{1}{p}} \vee t^{-\frac{r}{2}(2-\theta)}$ with $\frac{r}{2} > 1 - \frac{1}{p}$.

5.3. *k-Nearest Neighbour graphs.* For a finite generic set $\mu \subset \mathbb{R}^d$ and a positive integer $k \in \mathbb{N}$, the **k-Nearest Neighbour graph** has vertex set μ and an edge between $x, y \in \mu$ if and only if y is one of the k nearest points (in Euclidean distance) to x or vice-versa.

For our functional of interest, consider the following framework:

- $H \subset \mathbb{R}^d$ is a convex body;
- η is an $(\mathbb{R}^d, dx)$-Poisson measure;
- $\phi : (0, \infty) \to (0, \infty)$, a decreasing function, and $r > 2$ are such that

$$\int_0^1 \phi(s)^r s^{d-1} ds < \infty. \tag{5.17}$$

For any finite generic set $\mu \subset \mathbb{R}^d$, define

$$F(\mu) := \frac{1}{2} \sum_{x,y\in\mu, x\neq y} \mathbb{1}_{\{x\in N(y,\mu) \text{ or } y\in N(x,\mu)\}} \phi(|x-y|), \tag{5.18}$$

where $N(x, \mu)$ is the set consisting of the k nearest neighbours of x in μ. For $t \geqslant 1$, define $F_t := F(\eta_{|tH})$ and set $\hat{F}_t := (F_t - \mathbb{E}F_t)\operatorname{Var}(F_t)^{-1/2}$.

The k-nearest neighbour graph is a model frequently used in e.g. social sciences or geography, see [77] for a discussion of applications. Quantitative central limit theorems for edge-related quantities were shown in [5, 63] and subsequently improved in [40]. For a discussion of the literature, we refer to [40].

In [40], the authors give a quantitative central limit theorem for the sum of power-weighted edge-lengths with powers $\alpha \geqslant 0$, at a speed of convergence of $t^{-d/2}$. We complement this result by dealing with the case $\alpha \in \left(-\frac{d}{2}, 0\right)$. Note that in this case, the CLT is new even in its qualitative version. Power-weighted edge-length statistics with negative powers are used in spatial statistics, where spatial proximity implies a high correlation between quantities (see e.g. [46, Chapter 13.1.4], [79, Chapter 7.4] and [31, pages 65–69]). In the regime $\alpha \in \left(-\frac{d}{4}, 0\right)$, we recover the same, presumably optimal, speed of convergence of $t^{-d/2}$ as in [40], whereas in the case $\alpha \in \left(-\frac{d}{2}, -\frac{d}{4}\right]$, we find a speed of convergence that decreases as α approaches $-\frac{d}{2}$. It is natural to ask what happens when $\alpha \leqslant -\frac{d}{2}$. We consider this a separate issue and leave it open for further research.

THEOREM 5.7. *Under the conditions stated above, for any $p \in (1,2]$ such that $p < \frac{r}{2}$, there is a constant $c > 0$ such that, for $t \geqslant 1$ large enough,*

$$\max\left\{d_W\left(\hat{F}_t, N\right), d_K\left(\hat{F}_t, N\right)\right\} \leqslant ct^{d(1/p-1)}. \tag{5.19}$$

This inequality holds in particular for the function $\phi(x) = x^{-\alpha}$ with $0 < \alpha < \frac{d}{2}$, for any $p \in (1,2]$ such that $p < \frac{d}{2\alpha}$.

5.4. *Radial Spanning Tree.* Let $\mu \subset \mathbb{R}^d \setminus \{0\}$ be a finite set, generic with respect to the point 0. The radial spanning tree on μ, in short $RST(\mu)$, is constructed as follows:

- The set of vertices is given by $\mu \cup \{0\}$;
- for every $x \in \mu$, we add exactly one edge to the point $z \in \left(\mu \cap B^d(0,|x|)\right) \cup \{0\}$ which minimises $|x - z|$. We call z the radial nearest neighbour of x and say 'x connects to z', denoted by '$x \to z$ in μ'. We denote the length $|x - z|$ by $g(x, \mu)$.

In order to define our functional of interest, consider the following setting:

- $H \subset \mathbb{R}^d$ a convex body such that $B^d(0,\epsilon) \subset H$ for some $\epsilon > 0$;
- η is an $(\mathbb{R}^d, dx)$-Poisson measure;
- $\phi : (0,\infty) \to (0,\infty)$, a decreasing function, and $r > 2$ are such that

$$\int_0^1 \phi(s)^r s^{d-1} ds < \infty. \tag{5.20}$$

For any finite set $\mu \subset \mathbb{R}^d$ generic with respect to 0, define

$$F(\mu) := \sum_{x \in \mu} \phi(g(x,\mu)) \tag{5.21}$$

and for $t \geqslant 1$, define $F_t := F(\eta_{|tH})$. Set $\hat{F}_t := (F_t - \mathbb{E}F_t)\operatorname{Var}(F_t)^{-1/2}$.

Aside from an earlier reference in [24] under the name of 'exodic graph', the radial spanning tree was developed in [7] as a model related to the minimal directed spanning tree and to Poisson forests. The paper also discusses various applications, most notably in communication networks. Further work on the radial spanning tree has been done in [61, 8, 69]. In [69], the authors give a quantitative central limit theorem for sums of power-weighted edge-lengths of the radial spanning tree for powers $\alpha \geqslant 0$. The framework is one where the intensity of the Poisson measure increases while the observation window stays constant. After rescaling to our framework of a constant intensity and a growing window, one obtains by [69, Theorem 1.2] a speed of convergence of $t^{-d/2}$. We add quantitative central limit theorems for $\alpha \in \left(-\frac{d}{2}, 0\right)$, recovering the same speed of $t^{-d/2}$ for $\alpha \in \left(-\frac{d}{4}, 0\right)$. Note that this CLT is new even in its qualitative version. As for the k-Nearest Neighbour graph, the case $\alpha \leqslant -\frac{d}{2}$ will be the object of further research.

THEOREM 5.8. *Under the conditions stated above, for any $p \in (1,2]$ such that $p < \frac{r}{2}$, there is a constant $c > 0$ such that for $t \geqslant 1$ large enough,*

$$\max\left\{d_W\left(\hat{F}_t, N\right), d_K\left(\hat{F}_t, N\right)\right\} \leqslant ct^{d(1/p-1)}. \tag{5.22}$$

This inequality holds in particular for the function $\phi(x) = x^{-\alpha}$ with $0 < \alpha < \frac{d}{2}$, for any $p \in (1,2]$ such that $p < \frac{d}{2\alpha}$.

The rest of the paper is devoted to providing the proofs of the results in Sections 3 to 5.

6. Background on Malliavin Calculus. In this section, we present several useful notions related to Malliavin calculus. Unless otherwise indicated, these results are explained in [38]. We work in the setting of Section 2: in particular, χ indicates a $(\mathbb{W}, \nu)$-Poisson measure.

We start with three useful **isometry** relations.

Let $f \in L^2\left(\mathbb{W}^n, \nu^{(n)}\right)$ and $g \in L^2\left(\mathbb{W}^m, \nu^{(m)}\right)$. Then

$$\mathbb{E}\,\mathrm{I}_n(f)\,\mathrm{I}_m(g) = \mathbb{1}_{\{m=n\}} n! \int_{\mathbb{W}^n} \tilde{f}(x)\tilde{g}(x)\,\nu^{(n)}(dx), \tag{6.1}$$

where $\tilde{f}$ and $\tilde{g}$ are the **symmetrisations** of f and g defined by

$$\tilde{f}(x_1, \dots, x_n) = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} f(x_{\sigma(1)}, \dots, x_{\sigma(n)}), \tag{6.2}$$

the set Σ_n being the set of all permutations of $\{1, \dots, n\}$. See [38, Lemma 4] and the remark thereafter on page 10 for a proof.

Relation (6.1) implies that for $F, G \in L^2(\mathbb{P}_\chi)$ having an expansion (2.5) with kernels f_n and g_n respectively,

$$\mathbb{E} FG = \mathbb{E} F \mathbb{E} G + \sum_{n=1}^{\infty} n! \int_{\mathbb{W}^n} f_n(x) g_n(x)\,\nu^{(n)}(dx). \tag{6.3}$$

By [38, Theorem 5], if $h \in L^2(\mathbf{N} \times \mathbb{W})$ satisfies (2.13), then

$$\mathbb{E}\delta(h)^2 = \mathbb{E} \int_{\mathbb{W}} h(\chi, w)^2 \nu(dw) + \mathbb{E} \int_{\mathbb{W}} \int_{\mathbb{W}} \mathrm{D}_z\, h(\chi, w)\, \mathrm{D}_w\, h(\chi, z)\, \nu(dz)\nu(dw). \tag{6.4}$$

A well-known relation in Malliavin calculus is the so-called **integration by parts** formula: for $F \in \operatorname{dom} \mathrm{D}$ and $h \in \operatorname{dom} \delta$, we have $\mathbb{E} \int_{\mathbb{W}} h(\chi, w)\, \mathrm{D}_w\, F \nu(dw) = \mathbb{E}[F\delta(h)]$ (cf. [38, Theorem 4]). The condition on F is however suboptimal in our context, which is why we need a version of integration by parts under slightly different assumptions.

LEMMA 6.1. *Let $h \in \operatorname{dom} \delta \cap L^1(\mathbf{N} \times \mathbb{W})$ and $F \in L^0(\mathbb{P}_\chi)$ bounded. Then*

$$\mathbb{E} \int_{\mathbb{W}} h(\chi, w)\, \mathrm{D}_w\, F \nu(dw) = \mathbb{E}[F\delta(h)]. \tag{6.5}$$

PROOF. Since $h \in \operatorname{dom} \delta \cap L^1(\mathbf{N} \times \mathbb{W})$, it is easy to check that the expectations appearing in the statement are well-defined and finite. Note that

$$\begin{aligned} \mathbb{E} \int_{\mathbb{X}} \mathrm{D}_w\, F\, h(\chi, w)\nu(dw) &= \mathbb{E} \int_{\mathbb{X}} (F(\chi + \delta_w) - F(\chi)) h(\chi, w)\nu(dw) \\ &= \mathbb{E} \int_{\mathbb{X}} F(\chi + \delta_w) h(\chi, w)\nu(dw) - \mathbb{E} \int_{\mathbb{X}} F(\chi) h(\chi, w)\nu(dw), \end{aligned} \tag{6.6}$$

where the last line is justified by the fact that F is bounded and $h \in L^1(\mathbf{N} \times \mathbb{W})$, so both integrals are well-defined. We now apply Mecke's formula (2.7) to deduce that (6.6) equals

$$\begin{aligned} &\mathbb{E} \int_{\mathbb{X}} F(\chi) h(\chi - \delta_w, w)\chi(dw) - \mathbb{E} \int_{\mathbb{X}} F(\chi) h(\chi, w)\nu(dw) \\ &\qquad = \mathbb{E} F(\chi) \left(\int_{\mathbb{X}} h(\chi - \delta_w, w)\chi(dw) - \int_{\mathbb{X}} h(\chi, w)\nu(dw) \right). \end{aligned} \tag{6.7}$$

Since $h \in \operatorname{dom} \delta \cap L^1(\mathbf{N} \times \mathbb{X})$, the result follows by (2.14). ∎

Next, we introduce the **Ornstein–Uhlenbeck** operators P_τ, for $\tau \in [0,1]$. For $F \in L^1(\mathbb{P}_\chi)$ and $\tau \in [0,1]$, we define

$$P_\tau F(\chi) = \int \mathbb{E}\big[F(\chi^\tau + \xi)\big|\chi\big]\Pi_\tau(d\xi), \tag{6.8}$$

where χ^τ is a τ-thinning of χ (see [38, p. 24] and the reference given therein) and Π_τ is the law of an independent Poisson measure with intensity measure $(1-\tau)\nu$. It follows by Jensen's inequality that for all $p \geqslant 1$, one has

$$\mathbb{E}|P_\tau F|^p \leqslant \mathbb{E}|F|^p. \tag{6.9}$$

By [38, Lemma 6], for all $F \in L^2(\mathbb{P}_\chi)$ and all $\tau \in [0,1]$, for $\nu^{(n)}$-a.e. $w_1, \dots, w_n \in \mathbb{W}$ it holds $\mathbb{P}$-a.s. that

$$\mathrm{D}^{(n)}_{w_1,\dots,w_n}(P_\tau F) = \tau^n P_\tau \, \mathrm{D}^{(n)}_{w_1,\dots,w_n} F. \tag{6.10}$$

This implies that for all $F \in L^2(\mathbb{P}_\chi)$, the following expansion holds (see also [38, equation (79)]):

$$P_\tau F = \mathbb{E}F + \sum_{n=1}^{\infty} \tau^n \, \mathrm{I}_n(f_n). \tag{6.11}$$

The following lemma summarises some useful approximation properties of the Ornstein–Uhlenbeck operator.

LEMMA 6.2. *Let $h \in L^2(\mathbf{N}_\mathbb{W} \times \mathbb{W})$ and let $\tau \in (0,1)$. Then $P_\tau h$ satisfies condition (2.13) and $P_\tau h \to h$ in $L^2(\mathbf{N}_\mathbb{W} \times \mathbb{W})$ as $\tau \to 1$. Moreover, for $w, z \in \mathbb{W}$, and all $p \geqslant 1$,*

$$\mathbb{E}|P_\tau h(\chi, w)|^p \leqslant \mathbb{E}|h(\chi, w)|^p \tag{6.12}$$

and

$$\mathbb{E}|\,\mathrm{D}_z \, P_\tau h(\chi, w)|^p \leqslant \mathbb{E}|\,\mathrm{D}_z \, h(\chi, w)|^p. \tag{6.13}$$

Under the additional assumption that $h \in \operatorname{dom}\delta$, it holds that $\delta(P_\tau h) \to \delta(h)$ in $L^2(\mathbb{P}_\chi)$ as $\tau \to 1$.

PROOF. Combining the expansion formulae (6.11), (2.9) and (2.6), we derive that

$$\mathrm{D}_w \, P_\tau h(\chi, z) = \sum_{n=1}^{\infty} n\tau^n \, \mathrm{I}_{n-1}(h_n(z, w, .)). \tag{6.14}$$

Applying now the isometry property (6.3), it follows that

$$\mathbb{E}\int_\mathbb{W}\int_\mathbb{W} (\mathrm{D}_w \, P_\tau h(\chi, z))^2 \, \nu(dw)\nu(dz) = \sum_{n=0}^{\infty} n \cdot n! \, \tau^{2n} \|h_n\|^2_{n+1}, \tag{6.15}$$

where $\|.\|_n$ is the norm in $L^2(\mathbb{W}^n, \nu^{(n)})$. Now note that $\sup_{n\geqslant 1} n\tau^{2n} < \infty$ and

$$\sum_{n=0}^{\infty} n! \, \|h_n\|^2_{n+1} = \int_\mathbb{W} \mathbb{E}h(w)^2 \nu(dw) < \infty, \tag{6.16}$$

hence $P_\tau h$ satisfies (2.13). Similarly using the expansions, we deduce that

$$\mathbb{E}\int_\mathbb{W} (P_\tau h(\chi, w) - h(\chi, w))^2 \, \nu(dw) = \sum_{n=1}^{\infty} n! \, (\tau^n - 1)^2 \|h_n\|^2_{n+1}. \tag{6.17}$$

By dominated convergence, this expression tends to 0 as $\tau \to 1$. Properties (6.12) and (6.13) follow immediately from (6.9) and (6.10). For the last point, note that

$$\delta(P_\tau h) - \delta(h) = \sum_{n=0}^{\infty} \mathrm{I}_{n+1}((\tau^n - 1)h_n) \tag{6.18}$$

and

$$\mathbb{E}\left(\delta(P_\tau h) - \delta(h)\right)^2 = \sum_{n=0}^{\infty} (n+1)!(1-\tau^n)^2 \|\tilde{h}_n\|_{n+1}^2, \tag{6.19}$$

which converges to 0 as $\tau \to 1$ by dominated convergence since $h \in \operatorname{dom}\delta$. ■

The following lemma is used on several occasions:

LEMMA 6.3 ([41, Theorem 1.5]). *Let η be a $(\mathbb{W} \times [0,1], \nu \otimes ds)$-Poisson measure and let $F, G \in L^2(\mathbb{P}_\eta)$. Then*

$$\mathbb{E} \int_{\mathbb{W}} \int_0^1 \mathbb{E}[\mathrm{D}_{(y,s)} F | \eta_{|\mathbb{W}\times[0,1]}]^2 ds \lambda(dy) < \infty, \tag{6.20}$$

and an analogous estimate holds for G. Moreover,

$$\operatorname{Cov}(F, G) = \mathbb{E} \int_{\mathbb{W}} \int_0^1 \mathbb{E}[\mathrm{D}_{(y,s)} F | \eta_{|\mathbb{W}\times[0,s)}] \mathbb{E}[\mathrm{D}_{(y,s)} G | \eta_{|\mathbb{W}\times[0,s)}] ds \lambda(dy). \tag{6.21}$$

7. Proofs of Theorem 4.2 and Corollaries 4.3 and 4.7.

LEMMA 7.1. *For $p \in (1, 2]$, the function $\phi : \mathbb{R} \to \mathbb{R} : x \mapsto |x|^p$ is continuously differentiable and its derivative ϕ' is $(p-1)$-Hölder continuous with Hölder coefficient $c_\phi = p2^{2-p}$.*

PROOF. We want to show that

$$c_\phi := \sup_{\substack{a \neq b \\ a,b \in \mathbb{R}}} \frac{|\phi'(a) - \phi'(b)|}{|a-b|^{p-1}} = p2^{2-p}. \tag{7.1}$$

First we observe that $\phi'(x) = p\operatorname{sgn}(x)|x|^{p-1}$, with the convention that $\operatorname{sgn}(0) = 1$. Let $a \neq b$ and assume without loss of generality that $|a| \geqslant |b| \geqslant 0$. Then $a \neq 0$ and

$$\frac{\left|\operatorname{sgn}(a)|a|^{p-1} - \operatorname{sgn}(b)|b|^{p-1}\right|}{|a-b|^{p-1}} = \frac{\left|1 - \operatorname{sgn}\left(\frac{b}{a}\right)\left|\frac{b}{a}\right|^{p-1}\right|}{\left|1 - \frac{b}{a}\right|^{p-1}}. \tag{7.2}$$

It follows that

$$c_\phi = p \sup_{x \in [-1,1)} \frac{1 - \operatorname{sgn}(x)\,|x|^{p-1}}{(1-x)^{p-1}} =: p \sup_{x \in [-1,1)} f(x) \tag{7.3}$$

The function f is differentiable on $[-1, 1)$ with derivative $f'(x) = (p-1)(1-x)^{-p}(1 - |x|^{p-2}) \leqslant 0$, so f is decreasing and therefore

$$f(x) \leqslant f(-1) = 2^{2-p}. \tag{7.4}$$

We conclude that $c_\phi = p2^{2-p}$. ■

PROOF OF THEOREM 4.2. We start by showing the result for an approximation of h. Let $(U_m)_{m\in\mathbb{N}}\subset\mathbb{X}$ s.t. $\bigcup_m U_m=\mathbb{X}$ and for all $m\in\mathbb{N}$, $\lambda(U_m)<\infty$ and $U_m\subset U_{m+1}$. Define for $\mu\in\mathbf{N}$, $(y,s)\in\mathbb{X}\times[0,1]$:

$$h_m(\mu,y,s):=[(h(\mu,y,s)\wedge m)\vee(-m)]\,\mathbb{1}_{\{y\in U_m\}}. \tag{7.5}$$

Now $h_m\in L^1(\mathbf{N}\times\mathbb{X}\times[0,1])\cap L^2(\mathbf{N}\times\mathbb{X}\times[0,1])$ and $h_m\to h$ in L^2 as $m\to\infty$. Moreover, $|h_m|\leqslant|h|$ and $|\,\mathrm{D}_{(x,t)}\,h_m(\eta,y,s)|\leqslant|\,\mathrm{D}_{(x,t)}\,h(\eta,y,s)|$ for all $(x,t),(y,s)\in\mathbb{X}\times[0,1]$. This implies that h_m satisfies (2.13) and hence $h_m\in\operatorname{dom}\delta$.
Finally, h_m is also bounded. In particular, for any $\mu\in\mathbf{N}$, $(y,s)\in\mathbb{X}\times[0,1]$,

$$|h_m(\mu,y,s)|\leqslant m\mathbb{1}_{\{y\in U_m\}}. \tag{7.6}$$

We conclude that $\delta(h_m)$ is well-defined and has by (2.14) the pathwise expression

$$\delta(h_m)=\int_{\mathbb{X}\times[0,1]}h_m(\eta-\delta_{(y,s)},y,s)\eta(d(y,s))-\int_{\mathbb{X}}\int_0^1 h_m(\eta,y,s)ds\lambda(dy). \tag{7.7}$$

Define X_t as in Theorem 4.1 with $g=h=h_m$ and $X_0=0$. Then $\delta(h_m)=X_1$ and we infer from (4.2) that

$$\begin{aligned}\phi(\delta(h_m))-\phi(0)=&\int_{\mathbb{X}\times[0,1]}\big(\phi(X_{s-}+h_m(\eta-\delta_{(y,s)},y,s))-\phi(X_{s-})\big)\,\eta(d(y,s))\\&-\int_{\mathbb{X}}\int_0^1\phi'(X_s)h_m(\eta,y,s)ds\lambda(dy)\qquad\mathbb{P}\text{-a.s.}\end{aligned} \tag{7.8}$$

We would now like to take expectations on both sides of (7.8), but in order to do so, we must first show that the terms in question are in $L^1(\mathbb{P}_\eta)$. Start with a simple estimate for ϕ that uses Hölder continuity of ϕ'. Let $a,b\in\mathbb{R}$. Then

$$\begin{aligned}|\phi(a+b)-\phi(a)|&=\left|\int_0^1\phi'(a+ub)b\,du\right|\\&\leqslant|\phi'(a)b|+\int_0^1|b|\cdot\big|\phi'(a+ub)-\phi'(a)\big|\,du\\&\leqslant|\phi'(0)b|+|\phi'(a)-\phi'(0)|\cdot|b|+|b|\int_0^1 c_\phi|ub|^{p-1}du\\&\leqslant|\phi'(0)|\cdot|b|+c_\phi|a|^{p-1}\cdot|b|+\frac{c_\phi}{p}|b|^p.\end{aligned} \tag{7.9}$$

We are also going to need a bound on $\sup_{s\in[0,1]}\big\{\max\big\{|X_s|,|X_{s-}(\eta+\delta_{(y,s)})|\big\}\big\}$. Using (7.6), we find the following:

$$\begin{aligned}\max\big\{|X_s|,|X_{s-}(\eta+\delta_{(y,s)})|\big\}&\leqslant\int_{\mathbb{X}\times[0,1]}m\mathbb{1}_{\{x\in U_m\}}\eta(d(x,t))+\int_{\mathbb{X}}\int_0^1 m\mathbb{1}_{\{x\in U_m\}}dt\lambda(dx)\\&\leqslant m\left(\eta(U_m\times[0,1])+\lambda(U_m)\right).\end{aligned} \tag{7.10}$$

Using that $\phi(0)=0$ and (7.9) with $a=0$ and $b=\delta(h_m)$, we get for the LHS in (7.8)

$$\mathbb{E}\,|\phi(\delta(h_m))|\leqslant|\phi'(0)|\cdot\mathbb{E}|\delta(h_m)|+\frac{c_\phi}{p}\mathbb{E}|\delta(h_m)|^p, \tag{7.11}$$

which is finite since $\delta(h_m) \in L^2(\mathbb{P}_\eta)$. To show that the first term on the RHS in (7.8) is integrable, we first apply Mecke's formula (2.7) to get

$$\begin{aligned}(7.12)\quad & \mathbb{E}\left|\int_{\mathbb{X}\times[0,1]} \big(\phi(X_{s-}+h_m(\eta-\delta_{(y,s)},y,s))-\phi(X_{s-})\big)\,\eta(d(y,s))\right| \\ &\leqslant \mathbb{E}\int_{\mathbb{X}}\int_0^1 \left|\phi(X_{s-}(\eta+\delta_{(y,s)})+h_m(\eta,y,s))-\phi(X_{s-}(\eta+\delta_{(y,s)}))\right| ds\lambda(dy).\end{aligned}$$

Now we combine (7.9), (7.10) and (7.6) to get that the RHS of (7.12) is bounded by

$$\begin{aligned}(7.13)\quad & \mathbb{E}\int_{\mathbb{X}}\int_0^1 \Big(|\phi'(0)|\cdot m\mathbb{1}_{\{y\in U_m\}}+c_\phi|m\left(\eta(U_m\times[0,1])+\lambda(U_m)\right)|^{p-1}\cdot m\mathbb{1}_{\{y\in U_m\}} \\ &\qquad + \frac{c_\phi}{p}m^p\mathbb{1}_{\{y\in U_m\}}\Big)ds\lambda(dy) \\ &\leqslant m\lambda(U_m)\left(|\phi'(0)|+c_\phi m^{p-1}\mathbb{E}\left[(\eta(U_m\times[0,1])+\lambda(U_m))^{p-1}\right]+\frac{c_\phi m^{p-1}}{p}\right),\end{aligned}$$

which is finite since $\eta(U_m\times[0,1])$ is a Poisson random variable with parameter $\lambda(U_m)$ and thus all its moments are finite. This also shows that $\phi(X_{s-}(\eta+\delta_{(y,s)})+h_m(\eta,y,s)) - \phi(X_{s-}(\eta+\delta_{(y,s)}))\in L^1(\mathbf{N}\times\mathbb{X}\times[0,1])$. The second term on the RHS can be treated by the same method.

We can now take expectations on both sides of (7.8) and apply Mecke's formula (2.7) to get:

(7.14)

$$\begin{aligned}\mathbb{E}\phi(\delta(h_m)) &= \mathbb{E}\int_{\mathbb{X}}\int_0^1 \phi(X_{s-}(\eta+\delta_{(y,s)})+h_m(\eta,y,s))-\phi(X_{s-}(\eta+\delta_{(y,s)}))ds\lambda(dy) \\ &\quad -\mathbb{E}\int_{\mathbb{X}}\int_0^1 \phi'(X_s(\eta))h_m(\eta,y,s)ds\lambda(dy).\end{aligned}$$

Now add and subtract the integral of $\phi(X_{s-}(\eta)+h_m(\eta,y,s))-\phi(X_{s-}(\eta))$ (which can be shown to be in $L^1(\mathbf{N}\times\mathbb{X}\times[0,1])$ by the same methods as above). We obtain:

$$\begin{aligned}\mathbb{E}\phi(\delta(h_m)) &= \mathbb{E}\int_{\mathbb{X}}\int_0^1 \phi(X_{s-}(\eta+\delta_{(y,s)})+h_m(\eta,y,s))-\phi(X_{s-}(\eta+\delta_{(y,s)})) \\ &\qquad -\phi(X_{s-}(\eta)+h_m(\eta,y,s))+\phi(X_{s-}(\eta))ds\lambda(dy) \\ &\quad +\mathbb{E}\int_{\mathbb{X}}\int_0^1 \phi(X_{s-}(\eta)+h_m(\eta,y,s))-\phi(X_{s-}(\eta)) \\ &\qquad -\phi'(X_s(\eta))h_m(\eta,y,s)ds\lambda(dy) \\ (7.15)\qquad &=: I_1+I_2.\end{aligned}$$

To deal with I_1, note that for $a,b,c\in\mathbb{R}$

$$\begin{aligned}|\phi(a+c)-\phi(a)-\phi(b+c)+\phi(b)| &= \left|\int_a^b \big(\phi'(u+c)-\phi'(u)\big)\,du\right| \\ &\leqslant \int_{a\wedge b}^{a\vee b} c_\phi|c|^{p-1}du \\ (7.16)\qquad &= c_\phi|a-b|\cdot|c|^{p-1}.\end{aligned}$$

Hence

$$|I_1| \leqslant c_\phi \mathbb{E} \int_{\mathbb{X}} \int_0^1 \left| X_{s-}(\eta + \delta_{(y,s)}) - X_{s-}(\eta) \right| \cdot |h_m(\eta, y, s)|^{p-1} ds \lambda(dy). \tag{7.17}$$

Now we bound and rewrite part of the integrand on the RHS of (7.17) to find

$$\begin{aligned}
&\left| X_{s-}(\eta + \delta_{(y,s)}) - X_{s-}(\eta) \right| \\
&= \bigg| \int_{\mathbb{X} \times [0,s)} h_m(\eta + \delta_{(y,s)} - \delta_{(x,t)}, x, t) \eta(d(x,t)) - \int_{\mathbb{X}} \int_0^s h_m(\eta + \delta_{(y,s)}, x, t) dt \lambda(dx) \\
&\quad - \int_{\mathbb{X} \times [0,s)} h_m(\eta - \delta_{(x,t)}, x, t) \eta(d(x,t)) + \int_{\mathbb{X}} \int_0^s h_m(\eta, x, t) dt \lambda(dx) \bigg| \\
&= \left| \int_{\mathbb{X} \times [0,s)} \mathrm{D}_{(y,s)} h_m(\eta - \delta_{(x,t)}, x, t) \eta(d(x,t)) - \int_{\mathbb{X}} \int_0^s \mathrm{D}_{(y,s)} h_m(\eta, x, t) dt \lambda(dx) \right| \\
&\leqslant \int_{\mathbb{X} \times [0,s)} \left| \mathrm{D}_{(y,s)} h_m(\eta - \delta_{(x,t)}, x, t) \right| \eta(d(x,t)) + \int_{\mathbb{X}} \int_0^s \left| \mathrm{D}_{(y,s)} h_m(\eta, x, t) \right| dt \lambda(dx).
\end{aligned} \tag{7.18}$$

Multiplying this by $|h_m(\eta, x, t)|^{p-1}$ and taking expectations, after an application of Mecke's formula (2.7) we deduce that

$$\begin{aligned}
|I_1| \leqslant c_\phi \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds \int_{\mathbb{X}} \lambda(dx) \int_0^s dt \left| \mathrm{D}_{(y,s)} h_m(\eta, x, t) \right| \\
\left(|h_m(\eta + \delta_{(x,t)}, y, s)|^{p-1} + |h_m(\eta, y, s)|^{p-1} \right).
\end{aligned} \tag{7.19}$$

Since $|a - b|^{p-1} \leqslant |a|^{p-1} + |b|^{p-1}$ for all $a, b \in \mathbb{R}$, one has that

$$|h_m(\eta + \delta_{(x,t)}, y, s)|^{p-1} \leqslant |\, \mathrm{D}_{(x,t)} h_m(\eta, y, s)|^{p-1} + |h_m(\eta, y, s)|^{p-1}. \tag{7.20}$$

This implies that

$$\begin{aligned}
|I_1| &\leqslant c_\phi \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds \int_{\mathbb{X}} \lambda(dx) \int_0^s dt \left| \mathrm{D}_{(y,s)} h_m(\eta, x, t) \right| \cdot |\, \mathrm{D}_{(x,t)} h_m(\eta, y, s)|^{p-1} \\
&\quad + 2c_\phi \mathbb{E} \int_{\mathbb{X}} \lambda(dy) \int_0^1 ds \int_{\mathbb{X}} \lambda(dx) \int_0^s dt \left| \mathrm{D}_{(y,s)} h_m(\eta, x, t) \right| \cdot |h_m(\eta, y, s)|^{p-1}.
\end{aligned} \tag{7.21}$$

Now we consider $|I_2|$. For this, note that for $a, b \in \mathbb{R}$,

$$\begin{aligned}
|\phi(a + b) - \phi(a) - \phi'(a) b| &= \left| \int_0^b \left(\phi'(a + u) - \phi'(a) \right) du \right| \\
&\leqslant c_\phi \int_0^{|b|} u^{p-1} du \\
&= \frac{c_\phi}{p} |b|^p.
\end{aligned} \tag{7.22}$$

Applying this to $a = X_s(\eta)$ and $h = h_m(\eta, y, s)$ yields

$$|I_2| \leqslant \frac{c_\phi}{p} \mathbb{E} \int_{\mathbb{X}} \int_0^1 |h_m(\eta, y, s)|^p ds \lambda(dy). \tag{7.23}$$

Observe that in the previous computation we implicitly used the fact that, $\mathbb{P}$-a.s., the set $\{s\in[0,1]: X_s(\eta)\neq X_{s-}(\eta)\}$ has zero Lebesgue measure. We have therefore shown the following inequality:

$$
\begin{aligned}
&|\mathbb{E}\phi(\delta(h_m))| \\
&\quad\leqslant \frac{c_\phi}{p}\mathbb{E}\int_{\mathbb{X}}\lambda(dy)\int_0^1 ds\,|h_m(\eta,y,s)|^p \\
&\qquad + c_\phi\mathbb{E}\int_{\mathbb{X}}\lambda(dy)\int_0^1 ds\int_{\mathbb{X}}\lambda(dx)\int_0^s dt\,\left|\mathrm{D}_{(y,s)}\,h_m(\eta,x,t)\right|\cdot|\,\mathrm{D}_{(x,t)}\,h_m(\eta,y,s)|^{p-1} \\
&\qquad + 2c_\phi\mathbb{E}\int_{\mathbb{X}}\lambda(dy)\int_0^1 ds\int_{\mathbb{X}}\lambda(dx)\int_0^s dt\,\left|\mathrm{D}_{(y,s)}\,h_m(\eta,x,t)\right|\cdot|h_m(\eta,y,s)|^{p-1}.
\end{aligned}
\tag{7.24}
$$

By the construction of h_m, the RHS of this inequality is upper bounded by

$$
\begin{aligned}
&\frac{c_\phi}{p}\mathbb{E}\int_{\mathbb{X}}\lambda(dy)\int_0^1 ds\,|h(\eta,y,s)|^p \\
&\quad + c_\phi\mathbb{E}\int_{\mathbb{X}}\lambda(dx)\int_0^1 ds\int_{\mathbb{X}}\lambda(dx)\int_0^s dt\,\left|\mathrm{D}_{(y,s)}\,h(\eta,x,t)\right|\cdot|\,\mathrm{D}_{(x,t)}\,h(\eta,y,s)|^{p-1} \\
&\qquad + 2c_\phi\mathbb{E}\int_{\mathbb{X}}\lambda(dy)\int_0^1 ds\int_{\mathbb{X}}\lambda(dx)\int_0^s dt\,\left|\mathrm{D}_{(y,s)}\,h(\eta,x,t)\right|\cdot|h(\eta,y,s)|^{p-1}.
\end{aligned}
\tag{7.25}
$$

In order to conclude the proof, it remains to show that $\mathbb{E}\phi(\delta(h_m))\to\mathbb{E}\phi(\delta(h))$, as $m\to\infty$. For this, we use (7.9) with $a=\delta(h)$ and $b=\delta(h_m)-\delta(h)$ to get

$$
\begin{aligned}
&\mathbb{E}|\phi(\delta(h_m))-\phi(\delta(h))| \\
&\qquad\leqslant \mathbb{E}|\delta(h_m)-\delta(h)|\left(|\phi'(0)|+c_\phi|\delta(h)|^{p-1}+\frac{c_\phi}{p}|\delta(h_m)-\delta(h)|^{p-1}\right).
\end{aligned}
\tag{7.26}
$$

Using, in order, the Cauchy–Schwarz, Minkowski and Jensen inequalities, we obtain

$$
\begin{aligned}
&\mathbb{E}|\phi(\delta(h_m))-\phi(\delta(h))| \\
&\leqslant \mathbb{E}\left[(\delta(h_m)-\delta(h))^2\right]^{1/2}\mathbb{E}\left[\left(|\phi'(0)|+c_\phi|\delta(h)|^{p-1}+\frac{c_\phi}{p}|\delta(h_m)-\delta(h)|^{p-1}\right)^2\right]^{1/2} \\
&\leqslant \mathbb{E}\left[\delta(h_m-h)^2\right]^{1/2}\left(|\phi'(0)|+c_\phi\mathbb{E}\left[|\delta(h)|^2\right]^{(p-1)/2}+\frac{c_\phi}{p}\mathbb{E}\left[|\delta(h_m-h)|^2\right]^{(p-1)/2}\right).
\end{aligned}
\tag{7.27}
$$

By the isometry property (6.4) and the Cauchy–Schwarz inequality, we infer that

$$
\begin{aligned}
&\mathbb{E}\left[|\delta(h)|^2\right] \\
&\quad\leqslant \mathbb{E}\int_{\mathbb{X}}\int_0^1 h(\eta,y,s)^2 ds\lambda(dy)+\mathbb{E}\int_{\mathbb{X}}\int_0^1\int_{\mathbb{X}}\int_0^1\left(\mathrm{D}_{(x,t)}\,h(\eta,y,s)\right)^2 ds\lambda(dy)dt\lambda(dx)
\end{aligned}
\tag{7.28}
$$

and

(7.29)
$$\mathbb{E}\left[|\delta(h_m-h)|^2\right] \leqslant \mathbb{E}\int_{\mathbb{X}}\int_0^1 (h_m(\eta,y,s)-h(\eta,y,s))^2 ds\lambda(dy)$$
$$+\,\mathbb{E}\int_{\mathbb{X}}\int_0^1\int_{\mathbb{X}}\int_0^1 \left(\mathrm{D}_{(x,t)}\, h_m(\eta,y,s)-\mathrm{D}_{(x,t)}\, h(\eta,y,s)\right)^2 ds\lambda(dy)dt\lambda(dx).$$

The RHS of (7.28) is finite because h satisfies (2.13), and the RHS of (7.29) tends to 0 as $m\to\infty$ by dominated convergence and the assumption (2.13). This implies convergence to 0 on the RHS of (7.27).

To show that the inequality holds in particular for $\phi(x)=|x|^p$ with $p\in(1,2]$, it suffices to note that by Lemma 7.1, this function satisfies the conditions of the theorem. For $\phi(x)=|x|$, we define h_m as in (7.5). Then $h_m\in L^1(\mathbf{N}\times\mathbb{X}\times[0,1])\cap L^2(\mathbf{N}\times\mathbb{X}\times[0,1])\cap\operatorname{dom}\delta$ and the pathwise representation (2.14) holds. By the triangle inequality, we have

(7.30)
$$\mathbb{E}|\delta(h_m)|\leqslant\mathbb{E}\int_{\mathbb{X}\times[0,1]}|h_m(\eta-\delta_{y,s},y,s)|\eta(d(y,s))+\mathbb{E}\int_{\mathbb{X}}\int_0^1|h_m(\eta,y,s)|ds\lambda(dy).$$

By Mecke's formula (2.7) and the fact that $|h_m|\leqslant|h|$, inequality (4.3) follows with h_m on the LHS instead of h, and the second and third term on the RHS being zero. As we have shown before, $\delta(h_m)\to\delta(h)$ in $L^2(\mathbb{P}_\eta)$ as $m\to\infty$, and hence the inequality follows for h. ■

8. Proofs of Theorems 3.2, 3.3 and 3.4. Throughout this section, we work with the simplified notation adopted in Remark 3.1; recall also the definitions of the distances d_W and d_K given in (2.21) and (2.22), as well as the definitions and properties of the canonical solutions f_h and f_z as given in (2.21), (2.22) and thereafter. Let $N\sim\mathcal{N}(0,1)$.

PROOF OF THEOREM 3.2. We will show the upper bounds in (3.1) and (3.2) by exploiting the representations (2.25) and (2.27). The proof is divided into three steps. First, we are going to derive the first terms on the RHSs of (3.1) and (3.2) for both Wasserstein and Kolmogorov distances at the same time. Second, we deduce the second term on the RHS of (3.1) and, as a last step, we find the second term on the RHS of (3.2). Throughout the proof, we fix $h\in\mathcal{H}$ and $z\in\mathbb{R}$ and consider the corresponding canonical solutions f_h and f_z.

Step 1. Write f for either f_h or f_z. Then, f is Lipschitz and there is a version of f' which is bounded. Since $F\in\operatorname{dom}\mathrm{D}$ and f' is bounded, the expression

$$\mathbb{E}f'(F)\int_{\mathbb{Y}}\mathrm{D}_y\,F\,\mathbb{E}[\mathrm{D}_y\,F|\eta_y]\bar{\lambda}(dy) \tag{8.1}$$

is well-defined. Add and subtract this term to $\mathbb{E}f'(F)-\mathbb{E}Ff(F)$ and bound the resulting first term as follows:

$$\left|\mathbb{E}f'(F)-\mathbb{E}f'(F)\int_{\mathbb{Y}}\mathrm{D}_y\,F\,\mathbb{E}[\mathrm{D}_y\,F|\eta_y]\bar{\lambda}(dy)\right|$$
$$\leqslant\mathbb{E}|f'(F)|\cdot\left|1-\int_{\mathbb{Y}}\mathrm{D}_y\,F\,\mathbb{E}[\mathrm{D}_y\,F|\eta_y]\bar{\lambda}(dy)\right|$$
$$\leqslant\|f'\|_\infty\,\mathbb{E}\left|1-\int_{\mathbb{Y}}\mathrm{D}_y\,F\,\mathbb{E}[\mathrm{D}_y\,F|\eta_y]\bar{\lambda}(dy)\right|. \tag{8.2}$$

As $\|f_h'\|_\infty\leqslant\sqrt{\frac{2}{\pi}}$ and $\|f_z'\|_\infty\leqslant 1$, the bounds follow.

We are left to deal with

$$\left|\mathbb{E}Ff(F)-\mathbb{E}f'(F)\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy)\right|. \tag{8.3}$$

Since f is Lipschitz and $F\in L^2(\mathbb{P}_\eta)$, it follows that $f(F)\in L^2(\mathbb{P}_\eta)$. Hence by Lemma 6.3

$$\mathbb{E}Ff(F)=\mathrm{Cov}(F,f(F))=\mathbb{E}\int_{\mathbb{Y}}\mathbb{E}[\mathrm{D}_y F|\eta_y]\,\mathbb{E}[\mathrm{D}_y f(F)|\eta_y]\bar{\lambda}(dy). \tag{8.4}$$

Again by Lipschitzianity of f, it follows that $|\mathrm{D}_y f(F)|\leqslant|\mathrm{D}_y F|$ and hence an application of the Cauchy–Schwarz inequality, together with the fact that $F\in\operatorname{dom}\mathrm{D}$, justifies that

$$\mathbb{E}Ff(F)=\mathbb{E}\int_{\mathbb{Y}}\mathbb{E}[\mathrm{D}_y F|\eta_y]\,\mathrm{D}_y f(F)\bar{\lambda}(dy). \tag{8.5}$$

Therefore we are left to bound

$$\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot\left|f'(F)\,\mathrm{D}_y F-\mathrm{D}_y f(F)\right|\bar{\lambda}(dy). \tag{8.6}$$

Step 2. To bound (8.6) for the Wasserstein distance, we use an argument borrowed from the proof of [11, Theorem 3.1] to upper bound $|f_h(b)-f_h(a)-f_h'(a)(b-a)|$. Let $a,b\in\mathbb{R}$. Then by the properties stated above, f_h is Lipschitz and hence

$$|f_h(b)-f_h(a)-f_h'(a)(b-a)|\leqslant|f_h(b)-f_h(a)|+|f_h'(a)(b-a)|\leqslant 2\sqrt{\tfrac{2}{\pi}}|b-a|. \tag{8.7}$$

But at the same time by Lipschitzianity of f_h',

$$|f_h(b)-f_h(a)-f_h'(a)(b-a)|=\left|\int_a^b f_h'(x)-f_h'(a)dx\right|\leqslant 2(b-a)^2. \tag{8.8}$$

We deduce that for any $q\in[1,2]$

$$|f_h(b)-f_h(a)-f_h'(a)(b-a)|\leqslant 2\min\{|b-a|,(b-a)^2\}\leqslant 2|b-a|^q. \tag{8.9}$$

It follows that

$$\begin{aligned}\left|f_h'(F)\,\mathrm{D}_y F-\mathrm{D}_y f_h(F)\right|&=|f_h(F(\eta+\delta_y))-f_h(F(\eta))-f_h'(F(\eta))(F(\eta+\delta_y)-F(\eta))|\\&\leqslant 2|\mathrm{D}_y F|^q,\end{aligned} \tag{8.10}$$

and therefore (8.6) is bounded by

$$2\,\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\,|\mathrm{D}_y F|^q\bar{\lambda}(dy). \tag{8.11}$$

The required bound in the Wasserstein distance follows suit.
Step 3. We reason as in the proof of [34, Theorem 1.12] to conclude

$$\left|f_z'(F)\,\mathrm{D}_y F-\mathrm{D}_y f_z(F)\right|\leqslant\mathrm{D}_y F\cdot\mathrm{D}_y(Ff_z(F)+\mathbb{1}_{\{F>z\}}). \tag{8.12}$$

Thus (8.6) is upper bounded by

$$\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\,\mathrm{D}_y F\cdot\mathrm{D}_y\left(Ff_z(F)+\mathbb{1}_{\{F>z\}}\right)\bar{\lambda}(dy). \tag{8.13}$$

The desired bound now follows by taking the supremum over all $z\in\mathbb{R}$. ■

PROOF OF THEOREM 3.3. We apply Theorem 3.2 and bound the ensuing RHS of (3.1). The proof will be split into two steps.

Step 1. We start by showing that under the conditions of the theorem,

$$\mathbb{E}\left|1-\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy)\right| \leqslant \sqrt{\frac{\pi}{2}}\beta_1^{(p)}+\sqrt{\frac{\pi}{2}}\beta_2^{(p)}. \tag{8.14}$$

We can assume that $\mathbb{E}F=0$ and $\sigma=1$ (indeed, the result then follows since $\mathrm{D}\,\hat{F}=\sigma^{-1}\,\mathrm{D}\,F$). For ease of notation, define

$$G:=\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy)-1. \tag{8.15}$$

As a first step, we show that $\mathbb{E}G=0$. As $F\in\operatorname{dom}\mathrm{D}$, we can use Fubini's theorem and Lemma 6.3 to deduce

$$\mathbb{E}\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy)=\mathbb{E}\int_{\mathbb{Y}}\mathbb{E}[\mathrm{D}_y F|\eta_y]^2\bar{\lambda}(dy)=\operatorname{Var}(F)=1. \tag{8.16}$$

Now by the modification (4.6) of Corollary 4.3 given in Remark 4.4 and since $G\in L^1(\mathbb{P}_\eta)$, we have for $p\in[1,2]$,

$$\mathbb{E}|G|\leqslant(\mathbb{E}|G|^p)^{1/p}\leqslant\left(2^{2-p}\mathbb{E}\int_{\mathbb{Y}}\mathbb{E}[|\,\mathrm{D}_x G||\eta_x]^p\bar{\lambda}(dx)\right)^{1/p}. \tag{8.17}$$

Let us now study the term

$$\mathrm{D}_x G=\mathrm{D}_x\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy) \tag{8.18}$$

and define

$$h(\eta,y):=\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]. \tag{8.19}$$

We want to show that since $h\in L^1(\mathbf{N}\times\mathbb{Y})$,

$$\left|\mathrm{D}_x\int_{\mathbb{Y}}h(\eta,y)\bar{\lambda}(dy)\right|\leqslant\int_{\mathbb{Y}}|\mathrm{D}_x h(\eta,y)|\,\bar{\lambda}(dy). \tag{8.20}$$

and to do so we need an argument taken from the proof of [40, Proposition 4.1]: Assume the RHS of (8.20) to be finite (if it is not, then the inequality (8.20) is trivially true). Then

$$\int_{\mathbb{Y}}|h(\eta+\delta_x,y)|\bar{\lambda}(dy)\leqslant\int_{\mathbb{Y}}|\mathrm{D}_x h(\eta,y)|+|h(\eta,y)|\bar{\lambda}(dy)<\infty \tag{8.21}$$

and hence

$$\mathrm{D}_x\int_{\mathbb{Y}}h(\eta,y)\bar{\lambda}(dy)=\int_{\mathbb{Y}}\mathrm{D}_x h(\eta,y)\bar{\lambda}(dy). \tag{8.22}$$

The inequality (8.20) follows. We have therefore shown

$$\begin{aligned}\mathbb{E}|G|&\leqslant\left(2^{2-p}\mathbb{E}\int_{\mathbb{Y}}\mathbb{E}\left[\int_{\mathbb{Y}}|\,\mathrm{D}_x h(\eta,y)|\bar{\lambda}(dy)\middle|\,\eta_x\right]^p\bar{\lambda}(dx)\right)^{1/p}\\&=\left(2^{2-p}\int_{\mathbb{Y}}\mathbb{E}\left(\int_{\mathbb{Y}}\mathbb{E}\big[|\,\mathrm{D}_x h(\eta,y)||\eta_x\big]\bar{\lambda}(dy)\right)^p\bar{\lambda}(dx)\right)^{1/p},\end{aligned} \tag{8.23}$$

where the second line follows from Tonelli's theorem. By Minkowski's integral inequality,

$$\mathbb{E}\left(\int_{\mathbb{Y}} \mathbb{E}\big[|\operatorname{D}_x h(\eta,y)|\big|\eta_x\big]\,\bar\lambda(dy)\right)^p \leqslant \left(\int_{\mathbb{Y}} \mathbb{E}\big[\mathbb{E}\big[|\operatorname{D}_x h(\eta,y)|\big|\eta_x\big]^p\big]^{1/p}\bar\lambda(dy)\right)^p. \tag{8.24}$$

By the formula (2.4) for products,

$$\begin{aligned}\operatorname{D}_x h(\eta,y) &= \operatorname{D}_x\left(\operatorname{D}_y F\,\mathbb{E}[\operatorname{D}_y F|\eta_y]\right)\\ &= \operatorname{D}^{(2)}_{x,y} F\cdot \mathbb{E}[\operatorname{D}_y F|\eta_y] + \operatorname{D}_y F\cdot \operatorname{D}_x \mathbb{E}[\operatorname{D}_y F|\eta_y] + \operatorname{D}^{(2)}_{x,y} F\cdot \operatorname{D}_x \mathbb{E}[\operatorname{D}_y F|\eta_y].\end{aligned} \tag{8.25}$$

Since $\mathbb{E}[\operatorname{D}_y F|\eta_y]$ depends on η only through η_y, it follows that $\operatorname{D}_x \mathbb{E}[\operatorname{D}_y F|\eta_y] = 0$ whenever $x \geqslant y$. Moreover, since $F \in L^2(\mathbb{P}_\eta)$ implies that $\operatorname{D}_y F, \operatorname{D}^{(2)}_{x,y} F \in L^1(\mathbb{P}_\eta)$ by [41, Theorem 1.1], if $x < y$, we can put the add-one cost operator inside the conditional expectation. Therefore

$$\operatorname{D}_x \mathbb{E}[\operatorname{D}_y F|\eta_y] = \mathbb{1}_{\{x<y\}}\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]. \tag{8.26}$$

By Minkowski's norm inequality, it follows that

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}_x h(\eta,y)|\big|\eta_x\big]^p\right]^{1/p} &\leqslant \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F\cdot \mathbb{E}[\operatorname{D}_y F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}\\ &\quad + \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}_y F\cdot \mathbb{1}_{\{x<y\}}\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}\\ &\quad + \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F\cdot \mathbb{1}_{\{x<y\}}\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}.\end{aligned} \tag{8.27}$$

By the properties of conditional expectations and splitting the first term into two parts, (8.27) is bounded by

$$\begin{aligned}&\mathbb{1}_{\{y\leqslant x\}}\mathbb{E}\left[|\mathbb{E}[|\operatorname{D}^{(2)}_{x,y} F||\eta_x]^p\cdot \mathbb{E}[\operatorname{D}_y F|\eta_y]|^p\right]^{1/p}\\ &+ \mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[\mathbb{E}[|\operatorname{D}^{(2)}_{x,y} F||\eta_y]\cdot|\mathbb{E}[\operatorname{D}_y F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}\\ &+ \mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[\mathbb{E}[|\operatorname{D}_y F||\eta_y]\cdot|\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}\\ &+ \mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[\mathbb{E}[|\operatorname{D}^{(2)}_{x,y} F||\eta_y]\cdot|\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]|\big|\eta_x\big]^p\right]^{1/p}.\end{aligned} \tag{8.28}$$

By an application of Jensen's inequality (8.28) is now bounded by

$$\begin{aligned}&\mathbb{E}\left[|\mathbb{E}[\operatorname{D}_y F|\eta_y]|^p\cdot \mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_{x\vee y}\big]^p\right]^{1/p}\\ &+ \mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}_y F|\big|\eta_y\big]^p\cdot\big|\mathbb{E}\big[\operatorname{D}^{(2)}_{x,y} F\big|\eta_y\big]\big|^p\right]^{1/p}\\ &+ \mathbb{1}_{\{x<y\}}\mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_y\big]^{2p}\right]^{1/p}.\end{aligned} \tag{8.29}$$

Plugging the conclusion from (8.24) – (8.29) into (8.23) and applying Minkowski's norm inequality again yields

$$\begin{aligned}&\mathbb{E}|G|\\ &\leqslant 2^{2/p-1}\left(\int_{\mathbb{Y}}\left(\int_{\mathbb{Y}}\mathbb{E}\left[\big|\mathbb{E}[\operatorname{D}_y F|\eta_y]\big|^p\cdot \mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_{x\vee y}\big]^p\right]^{1/p}\bar\lambda(dy)\right)^p\bar\lambda(dx)\right)^{1/p}\end{aligned}$$

$$+ 2^{2/p-1} \left(\int_{\mathbb{Y}} \left(\int_{\mathbb{Y}} \mathbb{1}_{\{x<y\}} \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}_y F|\big|\eta_y\big]^p \cdot \big|\mathbb{E}[\operatorname{D}^{(2)}_{x,y} F|\eta_y]\big|^p \right]^{1/p} \bar{\lambda}(dy) \right)^p \bar{\lambda}(dx) \right)^{1/p}$$

(8.30)

$$+ 2^{2/p-1} \left(\int_{\mathbb{Y}} \left(\int_{\mathbb{Y}} \mathbb{1}_{\{x<y\}} \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_y\big]^{2p} \right]^{1/p} \bar{\lambda}(dy) \right)^p \bar{\lambda}(dx) \right)^{1/p}$$

which is in turn bounded by

$$2^{2/p} \left(\int_{\mathbb{Y}} \left(\int_{\mathbb{Y}} \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}_y F|\big|\eta_y\big]^p \cdot \mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_{x\vee y}\big]^p \right]^{1/p} \bar{\lambda}(dy) \right)^p \bar{\lambda}(dx) \right)^{1/p}$$

$$+ 2^{2/p-1} \left(\int_{\mathbb{Y}} \left(\int_{\mathbb{Y}} \mathbb{1}_{\{x<y\}} \mathbb{E}\left[\mathbb{E}\big[|\operatorname{D}^{(2)}_{x,y} F|\big|\eta_y\big]^{2p} \right]^{1/p} \bar{\lambda}(dy) \right)^p \bar{\lambda}(dx) \right)^{1/p}. \tag{8.31}$$

The result follows by an application of the Cauchy–Schwarz inequality.
Step 2. We want to show that

$$\mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}[\operatorname{D}_y F|\eta_y]| \cdot |\operatorname{D}_y F|^q \bar{\lambda}(dy) \leqslant \tfrac{1}{2}\beta_3^{(q)} + \tfrac{1}{2}\beta_4^{(q)}. \tag{8.32}$$

Again it suffices to show (8.32) for F with $\mathbb{E}F = 0$ and $\mathbb{E}F^2 = 1$. Assume the $\beta_3^{(q)}, \beta_4^{(q)}$ to be finite (otherwise there is nothing to prove). Then, in particular

$$\mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}\,[\operatorname{D}_y F|\,\eta_y]|^{q+1}\,\bar{\lambda}(dy) < \infty \tag{8.33}$$

and we can add and subtract this term to the LHS of (8.32), yielding $\beta_3^{(q)}$ and the following rest term:

$$\begin{aligned}
&\left| \mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}[\operatorname{D}_y F|\eta_y]| \cdot |\operatorname{D}_y F|^q \bar{\lambda}(dy) - \mathbb{E}\int_{\mathbb{Y}} |\mathbb{E}\,[\operatorname{D}_y F|\,\eta_y]|^{q+1}\,\bar{\lambda}(dy) \right| \\
&= \left| \mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}[\operatorname{D}_y F|\eta_y]| \cdot \mathbb{E}[|\operatorname{D}_y F|^q|\eta_y] \bar{\lambda}(dy) - \mathbb{E}\int_{\mathbb{Y}} |\mathbb{E}\,[\operatorname{D}_y F|\,\eta_y]|^{q+1}\,\bar{\lambda}(dy) \right| \\
&\leqslant \mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}[\operatorname{D}_y F|\eta_y]| \cdot \big| \mathbb{E}[|\operatorname{D}_y F|^q|\eta_y] - |\mathbb{E}\,[\operatorname{D}_y F|\,\eta_y]|^q \big| \bar{\lambda}(dy),
\end{aligned} \tag{8.34}$$

where the equality is justified by the fact that $\mathbb{E}[G|\eta_y]$ is defined for all $G \in L^0(\mathbb{P}_\eta)$ which are integrable or non-negative (see the definition (2.18)).
The term $\mathbb{E}[|\operatorname{D}_y F|^q|\eta_y] - |\mathbb{E}\,[\operatorname{D}_y F|\,\eta_y]|^q$ can be rewritten as

$$\tilde{\mathbb{E}}|g_y(\chi)|^q - |\tilde{\mathbb{E}} g_y(\chi)|^q, \tag{8.35}$$

where $g_y(\chi) = \operatorname{D}_y F(\eta_y + \chi)$ and $\tilde{\mathbb{E}}$ is the expectation with respect to χ, a Poisson measure on the space $\{x \in \mathbb{Y} : x \geqslant y\}$ with intensity $\lambda \otimes ds$, independent of η. As $F \in \operatorname{dom} \operatorname{D}$, it can be shown that $g_y \in L^2(\mathbb{P}_\chi)$ for a.e. $y \in \mathbb{Y}$.
By (4.5) in Corollary 4.3,

$$\tilde{\mathbb{E}}|g_y(\chi)|^q - |\tilde{\mathbb{E}} g_y(\chi)|^q \leqslant 2^{2-q} \tilde{\mathbb{E}} \int_{\mathbb{Y}} \mathbb{1}_{\{y \leqslant x\}} \big| \tilde{\mathbb{E}}[\operatorname{D}_x g_y(\chi)|\chi_x] \big|^q \bar{\lambda}(dx). \tag{8.36}$$

Plugging (8.36) into (8.34), we deduce that the RHS of (8.34) is upper bounded by

$$2^{2-q} \mathbb{E} \int_{\mathbb{Y}} |\mathbb{E}[\operatorname{D}_y F|\eta_y]| \int_{\mathbb{Y}} \mathbb{1}_{\{y \leqslant x\}} \tilde{\mathbb{E}} \big| \tilde{\mathbb{E}}[\operatorname{D}_x g_y(\chi)|\chi_x] \big|^q \bar{\lambda}(dx) \bar{\lambda}(dy). \tag{8.37}$$

Since χ and η are independent and have the same intensity measure,

$$\mathbb{1}_{\{y\leqslant x\}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot\tilde{\mathbb{E}}\big|\tilde{\mathbb{E}}[\mathrm{D}_x g_y(\chi)|\chi_x]\big|^q \tag{8.38}$$

has the same law as

$$\mathbb{1}_{\{y\leqslant x\}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot\mathbb{E}\big[|\mathbb{E}[\mathrm{D}^{(2)}_{x,y} F|\eta_x]|^q\big|\eta_y\big]. \tag{8.39}$$

This implies finally that the RHS of (8.34) is upper bounded by

$$2^{2-q}\mathbb{E}\int_{\mathbb{Y}}\int_{\mathbb{Y}}\mathbb{1}_{\{y\leqslant x\}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot|\mathbb{E}[\mathrm{D}^{(2)}_{x,y} F|\eta_x]|^q\bar{\lambda}(dx)\bar{\lambda}(dy) \tag{8.40}$$

and the result follows by the Cauchy–Schwarz inequality. ■

PROOF OF THEOREM 3.4. It suffices to prove this result for F such that $\mathbb{E}F=0$ and $\sigma=1$. If Theorem 3.4 holds for all σ-finite spaces $\mathbb{X}$, it holds in particular for the space $\mathbb{X}\times[0,1]$. In fact, it suffices to prove Theorem 3.4 for functionals $F\in L^2(\mathbb{P}_\eta)\cap\operatorname{dom}\mathrm{D}$, where η is a $(\mathbb{X}\times[0,1],\lambda\otimes ds)$-Poisson measure. Indeed, as discussed in Section 2, we can regard any functional $F\in L^2(\mathbb{P}_\chi)$, with the $(\mathbb{X},\lambda)$-Poisson measure χ, as a functional of η without changing the law of F or its add-one costs. If we replace $\mathbb{X}$ by $\mathbb{X}\times[0,1]$ and χ by η in the terms $\gamma_1^{(p)},\dots,\gamma_7^{(p)}$, the expressions do not change either since for any $F\in L^2(\mathbb{P}_\chi)\cap\operatorname{dom}\mathrm{D}$, the integrands do not depend on time. For the rest of this proof, we let thus $F\in L^2(\mathbb{P}_\eta)\cap\operatorname{dom}\mathrm{D}$ and recall the notation from Remark 3.1 where $\mathbb{Y}:=\mathbb{X}\times[0,1]$ and $\bar{\lambda}:=\lambda\otimes ds$ are explicitly defined.

We divide this proof into two steps: first, we discuss the bound on the Wasserstein distance, second, we show the bound on the Kolmogorov distance.

Step 1. By a combination of Theorem 3.2 and the proof of Theorem 3.3, we see that

$$\begin{aligned} d_W(F,N) &\leqslant \sqrt{\frac{2}{\pi}}\,\mathbb{E}\left|1-\int_{\mathbb{Y}}\mathrm{D}_y F\,\mathbb{E}[\mathrm{D}_y F|\eta_y]\bar{\lambda}(dy)\right|+2\,\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot|\mathrm{D}_y F|^q\bar{\lambda}(dy)\\ &\leqslant \beta_1^{(p)}+\beta_2^{(p)}+2\,\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot|\mathrm{D}_y F|^q\bar{\lambda}(dy). \end{aligned} \tag{8.41}$$

To bound $\beta_1^{(p)}$ and $\beta_2^{(p)}$, we simply apply Jensen's inequality and bound $\mathbb{1}_{\{x<y\}}$ by 1. This gives

$$\beta_1^{(p)}+\beta_2^{(p)}\leqslant\gamma_1^{(p)}+\gamma_2^{(p)}. \tag{8.42}$$

For the second term, apply Hölder's inequality to deduce

$$\begin{aligned} &\mathbb{E}\int_{\mathbb{Y}}|\mathbb{E}[\mathrm{D}_y F|\eta_y]|\cdot|\mathrm{D}_y F|^q\bar{\lambda}(dy)\\ &\qquad\leqslant\int_{\mathbb{Y}}\left(\mathbb{E}\,|\mathbb{E}[\mathrm{D}_y F|\eta_y]|^{q+1}\right)^{1/(q+1)}\cdot\left(\mathbb{E}|\,\mathrm{D}_y F|^{q+1}\right)^{1-1/(q+1)}\bar{\lambda}(dy). \end{aligned} \tag{8.43}$$

A further application of Jensen's inequality yields the result.

Step 2. We start from inequality (3.2) in Theorem 3.2. The first term was dealt with in Step 1 of this proof. Let us deal with the second term. Define

$$h(\eta,y):=\mathrm{D}_y F\cdot|\mathbb{E}[\mathrm{D}_y F|\eta_y]| \tag{8.44}$$

and for $z\in\mathbb{R}$,

$$Z_z:=Ff_z(F)+\mathbb{1}_{\{F>z\}}. \tag{8.45}$$

The term we want to bound is thus given by

$$\sup_{z\in\mathbb{R}} \mathbb{E} \int_{\mathbb{Y}} h(\eta, y) \, \mathrm{D}_y \, Z_z \bar{\lambda}(dy). \tag{8.46}$$

Since $F \in \operatorname{dom} \mathrm{D}$, it follows by Cauchy–Schwarz inequality that $h \in L^1(\mathbf{N} \times \mathbb{Y})$. Moreover, by the properties of f_z, we have $|Z_z| \leqslant 2$ for all $z \in \mathbb{R}$. Hence we can apply Corollary 4.7 and get for any $p \in [1, 2]$

$$\begin{aligned}
&\mathbb{E} \int_{\mathbb{Y}} h(\eta, y) \, \mathrm{D}_y \, Z_z \bar{\lambda}(dy) \\
&\leqslant 2 \Bigg(2^{2-p} \mathbb{E} \int_{\mathbb{Y}} |h(\eta, y)|^p \bar{\lambda}(dy) \\
&\quad + p 2^{2-p} \mathbb{E} \int_{\mathbb{Y}} \int_{\mathbb{Y}} |\mathrm{D}_y \, h(\eta, x)|^p \, \bar{\lambda}(dy) \bar{\lambda}(dx) \\
&\quad + p 2^{3-p} \int_{\mathbb{Y}} \int_{\mathbb{Y}} \mathbb{1}_{\{x<y\}} \left(\mathbb{E} \, |\mathrm{D}_y \, h(\eta, x)|^p \right)^{1/p} \cdot \left(\mathbb{E} \, |h(\eta, y)|^p \right)^{1-1/p} \bar{\lambda}(dy) \bar{\lambda}(dx) \Bigg)^{1/p} \\
&= (I_1 + I_2 + I_3)^{1/p} \\
&\leqslant (I_1)^{1/p} + (I_2)^{1/p} + (I_3)^{1/p},
\end{aligned} \tag{8.47}$$

since $|a + b|^{1/p} \leqslant |a|^{1/p} + |b|^{1/p}$.

Let us first look at I_1. By applying the Cauchy–Schwarz and Jensen inequalities, it follows that

$$\begin{aligned}
I_1 &= 4 \, \mathbb{E} \int_{\mathbb{Y}} | \, \mathrm{D}_y \, F|^p \cdot |\mathbb{E}[\mathrm{D}_y \, F | \eta_y]|^p \bar{\lambda}(dy) \\
&\leqslant 4 \int_{\mathbb{Y}} \mathbb{E} \left[| \, \mathrm{D}_y \, F|^{2p} \right] \bar{\lambda}(dy).
\end{aligned} \tag{8.48}$$

Now we deal with I_2. By the formula (2.4),

$$\mathrm{D}_y \, h(\eta, x) = \mathrm{D}^{(2)}_{x,y} \, F \cdot \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| + \mathrm{D}^{(2)}_{x,y} \, F \cdot |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| + \mathrm{D}_x \, F \cdot \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]|. \tag{8.49}$$

By Minkowski's norm inequality,

$$\begin{aligned}
&\left(\mathbb{E} \int_{\mathbb{Y}} \int_{\mathbb{Y}} |\mathrm{D}_y \, h(\eta, x)|^p \, \bar{\lambda}(dy) \bar{\lambda}(dx) \right)^{1/p} \\
&\leqslant \left(\mathbb{E} \int_{\mathbb{Y}} \int_{\mathbb{Y}} \left| \mathrm{D}^{(2)}_{x,y} \, F \cdot \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| \right|^p \bar{\lambda}(dy) \bar{\lambda}(dx) \right)^{1/p} \\
&\quad + \left(\mathbb{E} \int_{\mathbb{Y}} \int_{\mathbb{Y}} \left| \mathrm{D}^{(2)}_{x,y} \, F \cdot |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| \right|^p \bar{\lambda}(dy) \bar{\lambda}(dx) \right)^{1/p} \\
&\quad + \left(\mathbb{E} \int_{\mathbb{Y}} \int_{\mathbb{Y}} \left| \, \mathrm{D}_x \, F \cdot \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| \right|^p \bar{\lambda}(dy) \bar{\lambda}(dx) \right)^{1/p}.
\end{aligned} \tag{8.50}$$

By the triangle inequality, and as in the proof of Theorem 3.3,

$$\big| \, \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x \, F | \eta_x]| \big| \leqslant \mathbb{1}_{\{y<x\}} | \, \mathrm{D}_y \, \mathbb{E}[\mathrm{D}_x \, F | \eta_x]| \leqslant |\mathbb{E}[\mathrm{D}^{(2)}_{x,y} \, F | \eta_x]|. \tag{8.51}$$

By the Cauchy–Schwarz and Jensen inequalities,

$$
\begin{aligned}
(I_2)^{1/p} &\leqslant (4p)^{1/p} \left(\int_{\mathbb{Y}} \int_{\mathbb{Y}} \mathbb{E}\left[|\, \mathrm{D}^{(2)}_{x,y} F|^{2p} \right] \bar{\lambda}(dy)\bar{\lambda}(dx) \right)^{1/p} \\
&\quad + 2^{2/p+1} p^{1/p} \left(\int_{\mathbb{Y}} \int_{\mathbb{Y}} \mathbb{E}\left[|\, \mathrm{D}^{(2)}_{x,y} F|^{2p} \right]^{1/2} \cdot \mathbb{E}\left[|\, \mathrm{D}_x F|^{2p} \right]^{1/2} \bar{\lambda}(dy)\bar{\lambda}(dx) \right)^{1/p}. \qquad (8.52)
\end{aligned}
$$

Lastly, we look at the term I_3. Since $\mathbb{E}[\mathrm{D}_x F|\eta_x]$ depends only on η_x, we get

$$\mathbb{1}_{\{x<y\}} \mathrm{D}_y \, |\mathbb{E}[\mathrm{D}_x F|\eta_x]| = 0. \qquad (8.53)$$

Hence

$$\mathbb{1}_{\{x<y\}} \left(\mathbb{E}\, |\mathrm{D}_y h(\eta, x)|^p \right)^{1/p} = \mathbb{1}_{\{x<y\}} \left(\mathbb{E} \left| \mathrm{D}^{(2)}_{x,y} F \cdot \mathbb{E}[\mathrm{D}_x F|\eta_x] \right|^p \right)^{1/p}. \qquad (8.54)$$

Applying Cauchy–Schwarz and Jensen again yields

$$I_3 \leqslant 8p \int_{\mathbb{Y}} \int_{\mathbb{Y}} \left(\mathbb{E}|\, \mathrm{D}^{(2)}_{x,y} F|^{2p} \right)^{\frac{1}{2p}} \cdot \left(\mathbb{E}|\, \mathrm{D}_x F|^{2p} \right)^{\frac{1}{2p}} \cdot \left(\mathbb{E}\, |\mathrm{D}_y F|^{2p} \right)^{1-1/p} \bar{\lambda}(dy)\bar{\lambda}(dx). \qquad (8.55)$$

This inequality concludes the proof. ∎

Acknowledgments. I would like to thank my supervisor Giovanni Peccati for extensive discussions and his invaluable help on this project. I would also like to thank Pierre Perruchaud for his contribution to the computation of the constant discussed in Lemmas 7.24 and 7.25 of the Supplementary Material [74]. I am grateful to Matthias Schulte and Raluca Balan for the careful reading of part or even all of this work and making many helpful suggestions. I would equally like to thank Günter Last, Joseph Yukich, Andrew Wade, Mark Podolskij and Gesine Reinert for useful discussions and comments, as well as the anonymous reviewers for their helpful comments.

Funding. The author was supported by the Luxembourg National Research Fund (PRIDE17/1224660/GPS) for most of the project. Revision were carried out while the author was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) grant (EP/T018445/1). The author would also like to thank the Hausdorff Research Institute for Mathematics, where she carried out part of the revisions during the dual trimester program *Synergies between modern probability, geometric analysis and stochastic geometry*.

SUPPLEMENTARY MATERIAL

Selected Proofs
This material contains the proofs of Theorem 4.1, Corollary 4.3, Corollary 4.7, as well as the proofs of Theorems 5.5, 5.7, 5.8 and 5.3 for the applications, including all necessary technical estimates.

REFERENCES

[1] ADAMCZAK, R., POLACZYK, B. and STRZELECKI, M. (2022). Modified log-Sobolev inequalities, Beckner inequalities and moment estimates. *Journal of Functional Analysis* **282** 109349. https://doi.org/10.1016/j.jfa.2021.109349

[2] ALÒS, E., LEÓN, J. A. and VIVES, J. (2008). An anticipating Itô formula for Lévy processes. *Alea* **4** 285–305.

[3] ALÒS, E. and NUALART, D. (1998). An extension of Itô's Formula for Anticipating Processes. *Journal of Theoretical Probability* **11** 493–514. https://doi.org/10.1023/A:1022692024364

[4] ANÉ, C. and LEDOUX, M. (2000). On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probability Theory and Related Fields* **116** 573–602. https://doi.org/10.1007/s004400050263

[5] AVRAM, F. and BERTSIMAS, D. (1993). On Central Limit Theorems in Geometrical Probability. *The Annals of Applied Probability* **3** 1033 – 1046. https://doi.org/10.1214/aoap/1177005271

[6] AZMOODEH, E., PECCATI, G. and YANG, X. (2018). Malliavin-Stein Method: a Survey of Recent Developments. *arXiv:1809.01912*.

[7] BACCELLI, F. and BORDENAVE, C. (2007). The radial spanning tree of a Poisson point process. *The Annals of Applied Probability* **17** 305–359.

[8] BACCELLI, F., COUPIER, D. and TRAN, V. C. (2013). Semi-infinite paths of the two-dimensional radial spanning tree. *Advances in Applied Probability* **45** 895–916.

[9] BALAN, R. M., XIA, P. and ZHENG, G. (2023). Almost sure central limit theorem for the hyperbolic Anderson model with Lévy white noise. *arXiv:2310.10784v1*.

[10] BALAN, R. M. and ZHENG, G. (2023). Hyperbolic Anderson model with Lévy white noise: spatial ergodicity and fluctuation. *arXiv:2302.14178*.

[11] BASSE-O'CONNOR, A., PODOLSKIJ, M. and THÄLE, C. (2020). A BerryEsseen theorem for partial sums of functionals of heavy-tailed moving averages. *Electronic Journal of Probability* **25** 1 – 31. https://doi.org/10.1214/20-EJP435

[12] BERGER, N., BOLLOBÁS, B., BORGS, C., CHAYES, J. and RIORDAN, O. (2007). Degree distribution of the FKP network model. *Theoretical Computer Science* **379** 306–316. https://doi.org/10.1016/j.tcs.2007.02.049

[13] BHATTACHARJEE, C. (2022). Gaussian approximation for rooted edges in a random minimal directed spanning tree. *Random Structures Algorithms* **61** 462–492. MR4474209

[14] BOURGUIN, S. and PECCATI, G. (2014). Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering. *Electronic Journal of Probability* **19** 1–42.

[15] CHAFAÏ, D. (2004). Entropies, convexity, and functional inequalities, On Φ-entropies and Φ-Sobolev inequalities. *Journal of Mathematics of Kyoto University* **44** 325 – 363. https://doi.org/10.1215/kjm/1250283556

[16] CHATTERJEE, S. (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probability Theory and Related Fields* **143** 1–40. https://doi.org/10.1007/s00440-007-0118-6

[17] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal approximation by Stein's method*. *Probability and its Applications (New York)*. Springer, Heidelberg. https://doi.org/10.1007/978-3-642-15007-4

[18] DECREUSEFOND, L., SCHULTE, M. and THÄLE, C. (2016). Functional Poisson approximation in KantorovichRubinstein distance with applications to U-statistics and stochastic geometry. *The Annals of Probability* **44** 2147 – 2197. https://doi.org/10.1214/15-AOP1020

[19] DI NUNNO, G., MEYER-BRANDIS, T., ØKSENDAL, B. and PROSKE, F. (2005). Malliavin Calculus And Anticipative Itô Formulae for Lévy Processes. *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **08** 235–258. https://doi.org/10.1142/S0219025705001950

[20] EICHELSBACHER, P., REDNOSS, B., THÄLE, C. and ZHENG, G. (2021). A simplified second-order Gaussian Poincaré inequality in discrete setting with applications. *arXiv:2108.05216*. https://doi.org/10.48550/arxiv.2108.05216

[21] EICHELSBACHER, P. and THÄLE, C. (2014). New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electronic Journal of Probability* **19** 1 – 25. https://doi.org/10.1214/EJP.v19-3061

[22] FABRIKANT, A., KOUTSOUPIAS, E. and PAPADIMITRIOU, C. H. (2002). Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet. In *Automata, Languages and Programming* (P. Widmayer, S. Eidenbenz, F. Triguero, R. Morales, R. Conejo and M. Hennessy, eds.) 110–122. Springer Berlin Heidelberg.

[23] GILBERT, E. N. (1961). Random Plane Networks. *Journal of the Society for Industrial and Applied Mathematics* **9** 533–543. https://doi.org/10.1137/0109045

[24] GILBERT, E. N. (1965). Random minimal trees. *J. Soc. Indust. Appl. Math.* **13** 376–387.

[25] GRYGIEREK, J. and THÄLE, C. (2020). Gaussian fluctuations for edge counts in high-dimensional random geometric graphs. *Statistics & Probability Letters* **158** 108674. https://doi.org/10.1016/j.spl.2019.108674

[26] GUSAKOVA, A., SAMBALE, H. and THÄLE, C. (2021). Concentration on Poisson spaces via modified Φ-Sobolev inequalities. *Stochastic Processes and their Applications* **140** 216–235. https://doi.org/10.1016/j.spa.2021.06.009

[27] HAN, G. and MAKOWSKI, A. M. (2009). One-Dimensional Geometric Random Graphs With Nonvanishing DensitiesPart I: A Strong Zero-One Law for Connectivity. *IEEE Transactions on Information Theory* **55** 5832–5839. https://doi.org/10.1109/TIT.2009.2032799

[28] HOUDRÉ, C. and PÉREZ-ABREU, V. (1995). Covariance identities and inequalities for functionals on Wiener and Poisson spaces. *Ann. Probab.* **23** 400–419.

[29] HOUDRÉ, C. and PRIVAULT, N. (2002). Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli* **8** 697–720.

[30] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. *North-Holland Mathematical Library* **24**. Elsevier.

[31] KALKHAN, M. A. *Spatial Statistics: Geospatial Information Modeling and Thematic Mapping*, 1 ed. CRC Press.

[32] KESTEN, H. and LEE, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **6** 495–527. https://doi.org/10.1214/aoap/1034968141

[33] LACHIÈZE-REY, R. and PECCATI, G. (2013). Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs. *Electronic Journal of Probability* **18** 1–32.

[34] LACHIÈZE-REY, R., PECCATI, G. and YANG, X. (2022). Quantitative two-scale stabilization on the Poisson space. *Ann. Appl. Probab.* **32** 3085–3145. https://doi.org/10.1214/21-aap1768 MR4474528

[35] LACHIÈZE-REY, R. and PECCATI, G. (2013). Fine Gaussian fluctuations on the Poisson space II: Rescaled kernels, marked processes and geometric U-statistics. *Stochastic Processes and their Applications* **123** 4186–4218. https://doi.org/10.1016/j.spa.2013.06.004

[36] LACHIÈZE-REY, R. and REITZNER, M. (2016). U-Statistics in Stochastic Geometry. In *Stochastic Analysis for Poisson Point Processes*, (G. Peccati and M. Reitzner, eds.). *Bocconi & Springer Series* **7** 229–253. Springer. https://doi.org/10.1007/978-3-319-05233-5

[37] LACHIÈZE-REY, R., SCHULTE, M. and YUKICH, J. E. (2019). Normal Approximation for Stabilizing Functionals. *The Annals of Applied Probability* **29** 931–993. https://doi.org/10.1214/18-AAP1405

[38] LAST, G. (2016). Stochastic Analysis for Poisson Processes. In *Stochastic Analysis for Poisson Point Processes*, (G. Peccati and M. Reitzner, eds.). *Bocconi & Springer Series* **7** 1–36. Springer. https://doi.org/10.1007/978-3-319-05233-5

[39] LAST, G., MOLCHANOV, I. and SCHULTE, M. (2022). Normal approximation of Kabanov-Skorohod integrals on Poisson spaces. *arXiv:2211.02009*.

[40] LAST, G., PECCATI, G. and SCHULTE, M. (2016). Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields* **165** 667–723. https://doi.org/10.1007/s00440-015-0643-7

[41] LAST, G. and PENROSE, M. D. (2011). Poisson process Fock space representation, chaos expansion and covariance inequalities. *Probability Theory and Related Fields* **150** 663–690. https://doi.org/10.1007/s00440-010-0288-5

[42] LAST, G. and PENROSE, M. D. (2011). Martingale representation for Poisson processes with applications to minimal variance hedging. *Stochastic Processes and their Applications* **121** 1588–1606. https://doi.org/10.1016/j.spa.2011.03.014

[43] LAST, G. and PENROSE, M. D. (2018). *Lectures on the Poisson process*. Cambridge University Press.

[44] LICHEV, L. and MITSCHE, D. (2021). New results for the random nearest neighbor tree. *arXiv:2108.13014*.

[45] MCDIARMID, C. (2003). Random channel assignment in the plane. *Random Structures & Algorithms* **22** 187–212. https://doi.org/10.1002/rsa.10077

[46] MORAGA, P. *Spatial Statistics for Data Science*, 1 ed. Chapman & Hall.

[47] MÜLLER, T. (2008). Two-point concentration in random geometric graphs. *Combinatorica* **28**. https://doi.org/10.1007/s00493-008-2283-3

[48] NOURDIN, I. (2022). Malliavin-Stein approach. https://sites.google.com/site/malliavinstein.

[49] NOURDIN, I. and PECCATI, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Relat. Fields* **145** 75–118. https://doi.org/10.1007/s00440-008-0162-x

[50] NOURDIN, I. and PECCATI, G. (2012). *Normal Approximations with Malliavin Calculus*. *Cambridge Tracts in Mathematics*. Cambridge University Press.

[51] NOURDIN, I., PECCATI, G. and REINERT, G. (2009). Second order Poincaré inequalities and CLTs on Wiener space. *Journal of Functional Analysis* **257** 593–609. https://doi.org/10.1016/j.jfa.2008.12.017

[52] NOURDIN, I., PECCATI, G. and YANG, X. (2020). Restricted hypercontractivity on the Poisson space. *Proc. Amer. Math. Soc.* **148** 3617–3632. https://doi.org/10.1090/proc/14964

[53] NUALART, D. and PARDOUX, É. (1988). Stochastic calculus with anticipating integrands. *Probability Theory and Related Fields* **78** 535–581. https://doi.org/10.1007/BF00353876

[54] PECCATI, G. and REITZNER, M., eds. (2016). *Stochastic Analysis for Poisson Point Processes*. *Bocconi & Springer Series*. Springer.

[55] PECCATI, G., SOLÉ, J. L., TAQQU, M. S. and UTZET, F. (2010). Stein's method and Normal approximation of Poisson functionals. *The Annals of Probability* **38** 443 – 478. https://doi.org/10.1214/09-AOP477

[56] PENROSE, M. D. (2003). *Random Geometric Graphs*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198506263.001.0001

[57] PENROSE, M. D. (2005). Multivariate Spatial Central Limit Theorems with Applications to Percolation and Spatial Graphs. *The Annals of Probability* **33** 1945–1991.

[58] PENROSE, M. D. and WADE, A. R. (2006). On the total length of the random minimal directed spanning tree. *Advances in Applied Probability* **38** 336372. https://doi.org/10.1239/aap/1151337075

[59] PENROSE, M. D. and WADE, A. R. (2008). Multivariate Normal Approximation in Geometric Probability. *Journal of Statistical Theory and Practice* **2** 293-326. https://doi.org/10.1080/15598608.2008.10411876

[60] PENROSE, M. D. and WADE, A. R. (2008). Limit theory for the random on-line nearest-neighbor graph. *Random Structures & Algorithms* **32** 125–156. https://doi.org/10.1002/rsa.20185

[61] PENROSE, M. D. and WADE, A. R. (2009). Random Directed and on-line Networks. In *New Perspectives in Stochastic Geometry* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199232574.003.0007

[62] PENROSE, M. D. and YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Annals of Applied probability* **11** 1005–1041.

[63] PENROSE, M. D. and YUKICH, J. E. (2005). Normal approximation in geometric probability. *Stein's method and applications* **5** 37–58.

[64] PETROV, V. V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag.

[65] PRIVAULT, N. and TORRISI, G. L. (2013). Probability approximation by Clark-Ocone covariance representation. *Electron. J. Probab.* **18** no. 91, 25. https://doi.org/10.1214/EJP.v18-2787

[66] REITZNER, M. and SCHULTE, M. (2013). Central limit theorems for U-statistics of Poisson point processes. *The Annals of Probability* **41** 3879 – 3909. https://doi.org/10.1214/12-AOP817

[67] REITZNER, M., SCHULTE, M. and THÄLE, C. (2017). Limit theory for the Gilbert graph. *Advances in Applied Mathematics* **88** 26–61. https://doi.org/10.1016/j.aam.2016.12.006

[68] SCHULTE, M. (2016). Normal Approximation of Poisson Functionals in Kolmogorov Distance. *Journal of Theoretical Probability* **29** 96–117. https://doi.org/10.1007/s10959-014-0576-6

[69] SCHULTE, M. and THÄLE, C. (2017). Central limit theorems for the radial spanning tree. *Random Structures Algorithms* **50** 262–286. https://doi.org/10.1002/rsa.20651

[70] SCHULTE, M. and YUKICH, J. E. (2019). Multivariate second order Poincaré inequalities for Poisson functionals. *Electron. J. Probab.* **24** 1–42. https://doi.org/10.1214/19-EJP386

[71] SCHULTE, M. and YUKICH, J. E. (2023). Rates of multivariate normal approximation for statistics in geometric probability. *The Annals of Applied Probability* **33** 507 – 548. https://doi.org/10.1214/22-AAP1822

[72] SHI, Z., BALASUBRAMANIAN, K. and POLONIK, W. (2022). A Flexible Approach for Normal Approximation of Geometric and Topological Statistics. *arXiv:2210.10744*.

[73] STEELE, J. M. (1989). Cost of sequential connection for points in space. *Operations Research Letters* **8** 137–142. https://doi.org/10.1016/0167-6377(89)90039-4

[74] TRAUTHWEIN, T. (2025). Supplement to "Quantitative CLTs on the Poisson space via Skorohod estimates and p-Poincaré inequalities".

[75] TRINH, K. D. (2019). On central limit theorems in stochastic geometry for add-one cost stabilizing functionals. *Electron. Commun. Probab.* **24** Paper No. 76, 15. https://doi.org/10.1214/19-ecp279

[76] VIDOTTO, A. (2020). An improved second-order Poincaré inequality for functionals of Gaussian fields. *J. Theoret. Probab.* **33** 396–427. https://doi.org/10.1007/s10959-019-00883-3 MR4064306

[77] WADE, A. R. (2007). Explicit Laws of Large Numbers for Random Nearest-Neighbour-Type Graphs. *Advances in Applied Probability* **39** 326–342.

[78] WADE, A. R. (2009). Asymptotic theory for the multidimensional random on-line nearest-neighbour graph. *Stochastic Processes and their Applications* **119** 1889–1911. https://doi.org/10.1016/j.spa.2008.09.006

[79] WALLER, L. A. and GOTWAY, C. A. *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons.

[80] WU, L. (2000). A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probability Theory and Related Fields* **118** 427–438. https://doi.org/10.1007/PL00008749

[81] ZHU, J. (2010). A study of SPDEs w.r.t. compensated Poisson random measures and related topics. *PhD Thesis, University of York*.