CONCENTRATION BOUNDS FOR GEOMETRIC POISSON FUNCTIONALS: LOGARITHMIC SOBOLEV INEQUALITIES REVISITED

SASCHA BACHMANN AND GIOVANNI PECCATI

ABSTRACT. We prove new concentration estimates for random variables that are functionals of a Poisson measure defined on a general measure space. Our results are specifically adapted to geometric applications, and are based on a pervasive use of a powerful logarithmic Sobolev inequality proved by L. Wu [43], as well as on several variations of the so-called Herbst argument. We provide several applications, in particular to edge counting and more general length power functionals in random geometric graphs, as well as to the convex distance for random point measures recently introduced by M. Reitzner [30].

Keywords: Concentration of Measure; Convex Distance; Herbst Argument; Logarithmic Sobolev Inequalities; Poisson Measure; Random Graphs; Stochastic Geometry.

2010 AMS Classification: 60D05; 60G57; 60C05

1. INTRODUCTION

1.1. **Overview.** Let η be a Poisson random measure over some measurable space $(\mathbb{X}, \mathcal{X})$ such that $\mathcal{X}$ is countably generated, and assume that η has a σ-finite intensity μ. Let $F = F(\eta)$ be a real-valued functional of η having finite expectation. In this paper, we are interested in proving several novel estimates for the upper and lower tails

$$\mathbb{P}(F \geq \mathbb{E}F + r) \quad \text{and} \quad \mathbb{P}(F \leq \mathbb{E}F - r), \quad r > 0,$$

that are well adapted for geometric applications, with particular emphasis on quantities appearing in the modern theory of random geometric graphs – see e.g. [29].

Our techniques are based on many variations of the so-called *Herbst argument* (see e.g. [3, 4, 24, 25]), basically consisting in using a logarithmic Sobolev inequality (or, alternatively, an integration by parts formula) in order to deduce a differential inequality involving the moment generating function $u \mapsto K(u) := \mathbb{E}[e^{uF}]$; solving the inequality then yields an upper bound on $K(u)$, implying in turn a tail estimate for F by means of Markov's inequality. The main insight developed in the present paper is that, by carefully combining the *Mecke formula* for Poisson point processes (see Section 2) with logarithmic Sobolev inequalities such as the one in Theorem 1.1 below, one can deduce bounds on $K(u)$ involving quantities of a fundamental geometric nature. As discussed below, other approaches to concentration via the Herbst argument on the Poisson space (see e.g. [7, 15, 43]) do not yield conditions that are amenable to geometric analysis.

1.2. **Logarithmic Sobolev inequalities and a motivating example.** Our starting point is the following powerful Theorem 1.1, proved by Wu in [43] (see also [8]), and extending previous breakthrough findings contained in [1, 2]. Such a result involves two objects: (i) the *entropy* of a random variable $Z > 0$ with $\mathbb{E}Z < \infty$, that is defined as

$$\mathrm{Ent}(Z) := \mathbb{E}(Z \log(Z)) - \mathbb{E}(Z) \log(\mathbb{E}Z),$$

and (ii) the *difference* (or *add-one cost*) operator DF, that is defined for any $x \in \mathbb{X}$ as

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

where δ_x denotes the Dirac mass at $x \in \mathbb{X}$.

Theorem 1.1 (See Corollary 2.3 in [43]). *For all $\lambda \in \mathbb{R}$ satisfying $\mathbb{E}(e^{\lambda F}) < \infty$ we have*

(1.1) $$\mathrm{Ent}(e^{\lambda F}) \leq \mathbb{E}\left[e^{\lambda F}\left(\int_{\mathbb{X}} \psi(\lambda D_x F(\eta))\ d\mu(x)\right)\right],$$

where $\psi(z) = ze^z - e^z + 1$.

A typical way of applying (1.1) to concentration estimates (via the Herbst argument) is demonstrated e.g. in [43, Proposition 3.1], where it is proved that, if DF and $\int_{\mathbb{X}}(DF)^2 d\mu$ are almost surely bounded by positive constants β and c, respectively, then $\mathbb{P}(F \geq \mathbb{E}F + r)$ is bounded by the function

$$r \mapsto \exp\left[-\frac{r}{2\beta} \log\left(1 + \frac{\beta r}{c}\right)\right], \quad r > 0.$$

Letting $\beta \to 0$, one deduces from this estimate that, if $DF \leq 0$ and $\int_{\mathbb{X}}(DF)^2 d\mu \leq c$, then,

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\frac{r^2}{2c}\right), \quad r > 0,$$

that is: such a concentration result only captures a Gaussian behaviour for the upper tail in the case of a non-increasing functional F such that $\int_{\mathbb{X}}(DF)^2 d\mu$ is deterministically bounded. Apart from the monotonicity requirement on F, a crucial limitation of a result of this kind is that, in most examples where F is a quantity arising in stochastic geometry (for instance, F is an *edge-counting statistic* such as the ones considered in Section 6 below), the quantity $\int_{\mathbb{X}}(DF)^2 d\mu$ *does not admit any meaningful geometric interpretation* – roughly because averaging DF over the deterministic measure μ completely cancels the special role played by those points in $\mathbb{X}$ that belong to the support of η. One should contrast such a situation with the following statement, that will be proved later on as a special case of Corollary 3.3 (such a result also implies the already quoted Proposition 3.1 in [43]):

Proposition 1.2. *Assume that there exists a finite constant $c > 0$ such that, almost surely,*

(1.2) $$V^+ := \int_{\mathbb{X}}(D_x F)_-^2 d\mu(x) + \int_{\mathbb{X}}(F(\eta) - F(\eta - \delta_x))_+^2 d\eta(x) \leq c,$$

where $(u)_-$ and $(u)_+$ stand for the negative and positive part of $u \in \mathbb{R}$, respectively. Then,

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\frac{r^2}{2c}\right).$$

Proposition 1.2 is particularly interesting when F is *non-decreasing*, that is, when $DF \geq 0$. Indeed, in this case one has that $V^+ = \int_{\mathbb{X}}(F(\eta) - F(\eta - \delta_x))_+^2 d\eta(x)$ and, since the role of μ is now immaterial, the relation $V_+ \leq c$ can be in principle verified by means of arguments of a purely geometric or combinatorial nature. For instance, we implement this strategy in Proposition 8.3 below, where we use Proposition 1.2 in order to deduce a novel intrinsic proof of the Gaussian upper tail behaviour of the convex distance for point processes – as recently introduced by Reitzner in [30].

As anticipated, the principal aim of this paper is to prove a large collection of statements with the same flavour as Proposition 1.2 (see Section 3), and then to apply them to random variables arising in the theory of random geometric graphs.

1.3. **Plan.** Our work is organised as follows. After some preliminary facts discussed in Section 2, the subsequent Section 3 contains the statements of our main concentration estimates. Our results involve random variables having a form similar to the quantity V^+ introduced above, and largely generalise Proposition 1.2. Proofs are detailed in Section 4.

Section 5 presents several applications of the results of Section 3 to Poisson U-statistics of arbitrary order – as defined in the seminal reference [31] (see also [6, 9, 12, 19, 20, 28, 37, 21]).

The results of Section 5 are specialised in Section 6 to the case of edge-counting statistics associated with general random geometric graphs. Several careful comparisons with the existing literature (in particular [11, 32]) are presented.

Section 7 contains further estimates on U-statistics of order two, that are proved by adapting some techniques introduced in [16, 33]. Geometric applications to edge-length functionals are discussed in detail.

As anticipated, in Section 8 we apply the estimates of Section 3, in order to deduce a novel intrinsic proof of the concentration estimates for the convex distance for random point measures established in [30]. Such a fundamental object generalises to the framework of random point processes the celebrated *convex distance* introduced by Talagrand in [41]; see [32, 21] for several applications. We stress that the problem of finding an intrinsic proof of the striking concentration results from [30] has been one of the main motivations for elaborating the theory developed in the present paper.

1.4. **Further remarks on the literature.** The inequalities obtained in this paper (as well as some techniques exploited in the proofs) are very close in spirit to those appearing in the seminal references [3, 4, 25], where the so-called *entropy method* (roughly corresponding to a combination of the Herbst argument and of logarithmic Sobolev inequalities – see e.g. [24]) is developed in the framework of functions of

finite vectors of independent random elements. We recall that the results from [3, 4, 25] typically apply to random variables with the form $F = f(X_1, ..., X_n)$, where $X = (X_1, ..., X_n)$ is a vector of independent random elements and f is some deterministic measurable function, and are based on a pervasive use of *random difference operators* of the type

$$\Delta_i f(X) = f(X) - f(X_1, ..., X_{i-1}, X_i', X_{i+1}, ..., X_n), \quad i = 1, ..., n,$$

where X' is an independent copy of X. By inspection of the results presented below, it is not difficult to show that, in the case where the intensity μ of the Poisson measure η is *finite* and *non-atomic*, some versions of the main results of the present paper could be obtained by implementing the following rough strategy:

(i) Select a sequence of measurable partitions $\{B_1^n, ..., B_{k(n)}^n : n \geq 1\}$ of $\mathbb{X}$, in such a way that $k(n) \to \infty$ and $\max_{i=1,...,k(n)} \mu(B_i^n) \to 0$, as $n \to \infty$.

(ii) Consider a random variable $F = F(\eta)$ and represent it in the form $F = f_n(X_{n,1}, .., X_{n,k(n)})$, where $X_{n,i}$ is defined as the restriction of η to the set B_i^n, and f_n is some appropriate measurable mapping.

(iii) For a fixed n, prove a concentration estimate for $F(\eta)$ by applying the results from [3, 4, 25] to $f_n(X_{n,1}, .., X_{n,k(n)})$, in particular by considering an independent copy of $X_{n,1}, .., X_{n,k(n)}$ defined in terms of an independent Poisson measure η' on $\mathbb{X}$ with intensity μ.

(iv) Let $n \to \infty$, and recover a bound involving quantities related to the add-one cost operator DF described above, by exploiting the fact that independent Poisson measures with non-atomic intensities have almost surely disjoint supports.

Apart from the fact that this approach only works with finite intensity measures without atoms, some investigations in this direction have convincingly shown us that (to the best of our expertise), in order for the step described at Point (iv) to take place in a meaningful way, one should systematically add to our statements some additional technical assumptions, that are indeed *not required* if one implements the direct approach based on the Mecke formula that is systematically adopted in this paper. An analogous phenomenon can be observed for instance in [14, Theorem 4], where a weaker version of the Poincaré inequality on the Poisson space is deduced by means of a discretisation procedure similar to the one outlined above, and of the use of the classical Efron-Stein inequality. In view of these remarks, we decided not to directly exploit the connection with the entropy method on product spaces in the proofs of our main results.

Another collection of results that is relevant for our paper is contained in references [7, 15], where the authors obtain concentration estimates by applying integration by parts techniques, in particular by using the properties of the so-called *Ornstein-Uhlenbeck semigroup* associated with a given Poisson measure – see also [40]. As in the already discussed examples from [43], the estimates contained in these references have an equally problematic geometric interpretation, since they involve integrals of add-one cost operators with respect to the underlying intensity measure

μ. Moreover, in order to exploit some probabilistic representation of the Ornstein-Uhlenbeck semigroup, one has also to work on *extended* probability spaces. It is a natural question to ask whether the Mecke formula could be combined with some of the estimates from [7, 15] in order to obtain concentration inequalities that are adapted to a geometric framework. We prefer to think of this issue as a separate problem, and leave it open for further research.

2. Framework

For the rest of the paper, we shall denote by $(\mathbb{X}, \mathcal{X}, \mu)$ a σ-finite measure space, such that the σ-field $\mathcal{X}$ is countably generated and $\mu(\mathbb{X}) > 0$. We write η to indicate a Poisson point process on $(\mathbb{X}, \mathcal{X})$. This means that $\eta = \{\eta(A) : A \in \mathcal{X}_0\}$ is a collection of random variables, defined on some probability space $(B, \mathcal{B}, \mathbb{P})$ and indexed by the elements of $\mathcal{X}_0 = \{A \in \mathcal{X} : \mu(A) < \infty\}$, such that the following properties are satisfied: **(i)** for every fixed $A \in \mathcal{X}_0$, $\eta(A)$ is a Poisson random variable with parameter $\mu(A)$, and **(ii)** for every collection of pairwise disjoint $A_1, ..., A_n \in \mathcal{X}_0$, one has that the random variables $\eta(A_1), ..., \eta(A_n)$ are stochastically independent.

As usual, we interpret η as a random element in the space $\mathbf{N} = \mathbf{N}(\mathbb{X})$ of integer-valued σ-finite measures ξ on $\mathbb{X}$ equipped with the smallest σ-field $\mathcal{N}$ making the mappings $\xi \mapsto \xi(B)$ measurable for all $B \in \mathcal{X}$; see e.g. [35] or [22]. The standard notation $x \in \eta$ is shorthand in order to indicate that the point x is an element of the support of η. We write $\hat{\eta}$ for the compensated random (signed) measure $\eta - \mu$. We shall write $F = F(\eta)$ to indicate that a given random variable F can be written in the form $F = \mathfrak{f}(\eta)$, $\mathbb{P}$-a.s., for some measurable function $\mathfrak{f} : \mathbf{N} \to \mathbb{R}$; such a function $\mathfrak{f}$ (which is uniquely determined by F up to sets of $\mathbb{P}$-measure zero) is customarily called a *representative* of F, and F is called a *Poisson functional*.

Remark 2.1 (Some conventions). In what follows, we will indifferently use the notation F and $F(\eta)$ if there is no ambiguity. Also, for any $\xi \in \mathbf{N}$, the notation $F(\xi)$ refers to $\mathfrak{f}(\xi)$, where $\mathfrak{f}$ is a fixed representative of F. Finally, we observe that, in the statements of some of our main results, we will often work under the assumption that the add-one cost operator $D_x F(\xi)$ verifies a given property $\mathcal{P}$ (for instance, $D_x F(\xi) \leq 0$) for every $x \in \mathbb{X}$ and every $\xi \in \mathbf{N}$: this requirement means of course that there exists a representative $\mathfrak{f}$ of F such that the quantity $\mathfrak{f}(\xi + \delta_x) - \mathfrak{f}(\xi)$ verifies $\mathcal{P}$ for every $x \in \mathbb{X}$ and every $\xi \in \mathbf{N}$.

We will systematically use the following standard notation: for every $\xi \in \mathbf{N}$,

$$(2.3) \qquad \xi_{\neq}^k := \{\mathbf{x} = (x_1, ..., x_k) : x_i \in \xi, \forall i = 1, ..., k, \text{ and } x_i \neq x_j, i \neq j\}.$$

A result that we shall use in several occasions (and that in some sense represents the backbone of our approach) is the following well-known *Slivnyak-Mecke formula*:

for every $k \geq 1$ and every non-negative measurable function H on $\mathbf{N} \times \mathbb{X}^k$, one has that

$$(2.4) \qquad \mathbb{E}\left[\int_{\mathbb{X}^k} H(\eta, x_1, \ldots, x_k) d\eta^{(k)}(x_1, \ldots, x_k)\right]$$

$$= \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} \mathbb{E}\left[H\left(\eta + \sum_{i=1}^{k} \delta_{x_i}, x_1, \ldots, x_k\right)\right] d\mu(x_1) \cdots d\mu(x_k),$$

where δ_x stands for the Dirac mass at x and $\eta^{(k)}$ is the point process on $\mathbb{X}^k$ with support $\eta_{\neq}^k$ and $\eta^{(k)}(\{\mathbf{x}\}) = \eta(\{x_1\}) \cdots \eta(\{x_k\})$ for any $\mathbf{x} = (x_1, \ldots, x_k) \in \eta_{\neq}^k$. A standard proof of the fundamental relation (2.4) can be found e.g. in [35, Theorem 3.2.5 and Corollary 3.2.3], in the case of a non-atomic intensity μ. The result extends straightforwardly to the case of a general σ-finite measure μ – see e.g. [22]. When specialised to the case $k = 1$, relation (2.4) is known as *Mecke formula*, and boils down to the following identity: for every non-negative measurable function H on $\mathbf{N} \times \mathbb{X}$, one has that

$$(2.5) \qquad \mathbb{E}\left[\int_{\mathbb{X}} H(\eta, x) d\eta(x)\right] = \int_{\mathbb{X}} \mathbb{E}[H(\eta + \delta_x, x)] d\mu(x).$$

For the rest of the paper, for every integer $k \geq 1$ and every real $p > 0$, we will write $L^p(\mu^k) := L^p(\mathbb{X}^k, \mathcal{X}^{\otimes k}, \mu^k)$, and also use the shorthand notation $L^p(\mu^1) = L^p(\mu)$. In Section 5.4, the symbol $L^2(\mu^0)$ is used to denote the real line $\mathbb{R}$, endowed with the usual Euclidean inner product.

3. Deviation Inequalities for Poisson Functionals

In the following, we are going to develop new tools for proving deviation inequalities of Poisson functionals. Our approach is an adaptation of the entropy method for product space functionals that was particularly investigated in [4].

The heart of the method we are about to present is the modified logarithmic Sobolev inequality stated below. For the remainder of the section, we consider a Poisson functional F. As above, we will use the difference (or add-one cost) operator DF, that is defined for any $(x, \xi) \in \mathbb{X} \times \mathbf{N}$ by

$$D_x F(\xi) = F(\xi + \delta_x) - F(\xi),$$

where δ_x denotes the Dirac mass at $x \in \mathbb{X}$. To shorten notations, we write

$$D_x^I F(\xi) = D_x F(\xi) \mathbb{1}\{(x, \xi) \in I\}$$

whenever $I \subseteq \mathbb{X} \times \mathbf{N}$ is measurable, $x \in \mathbb{X}$ and $\xi \in \mathbf{N}$. In the same spirit we will also use the notations

$$D_x^{\geq \beta} F(\xi) = D_x F(\xi) \mathbb{1}\{D_x F(\xi) \geq \beta\},$$
$$D_x^+ F(\xi) = D_x F(\xi) \mathbb{1}\{D_x F(\xi) \geq 0\}.$$

The quantities $D_x^{\leq\beta}F(\xi)$ and $D_x^-F(\xi)$ are defined analogously, and so are the operators $D_x^{>\beta}F$ and $D_x^{<\beta}F$ (with strict inequalities). The following observation is derived by combining Wu's modified logarithmic Sobolev inequality for Poisson point processes (1.1) with the Mecke formula (2.5).

Proposition 3.1. *Let $I \subseteq \mathbb{X} \times \mathbf{N}$ be a measurable set. Then for all $\lambda \in \mathbb{R}$ satisfying $\mathbb{E}(e^{\lambda F}) < \infty$ we have*

$$\mathrm{Ent}(e^{\lambda F}) \leq \mathbb{E}\left[e^{\lambda F}\left(\int_{\mathbb{X}} \psi(\lambda D_x^I F(\eta))\, d\mu(x) + \int_{\mathbb{X}} \phi(-\lambda D_x^{I^c} F(\eta - \delta_x))\, d\eta(x)\right)\right],$$

where $\phi(z) = e^z - z - 1$ and $\psi(z) = ze^z - e^z + 1$.

For any $\beta \in \mathbb{R}$ we define the random variables $V_\beta^+ = V_\beta^+(F)$ and $V_\beta^- = V_\beta^-(F)$ by

$$V_\beta^+ = \int_{\mathbb{X}} (D_x^{\leq\beta}F(\eta))^2\, d\mu(x) + \int_{\mathbb{X}} (D_x^{>\beta}F(\eta - \delta_x))^2 d\eta(x),$$

$$V_\beta^- = \int_{\mathbb{X}} (D_x^{\geq\beta}F(\eta))^2\, d\mu(x) + \int_{\mathbb{X}} (D_x^{<\beta}F(\eta - \delta_x))^2 d\eta(x).$$

Note that we will write $V^+ = V_0^+$ and $V^- = V_0^-$. The notation is in correspondence with [4] where the entropy method for product spaces was investigated. The upcoming result can be regarded as a generalized analogue of [4, Theorem 2] for the Poisson space. The proof is similar to the product space version where Proposition 3.1 takes now the role of the log Sobolev inequality. The generalization is achieved using arguments similar to those in the proof of [43, Proposition 3.1].

To get prepared for the presentation of the theorem, for $\beta \in \mathbb{R}$ and $z > 0$, let

$$\Phi_\beta(z) = \begin{cases} \psi(z\beta)/(z\beta^2) & \beta > 0 \\ z/2 & \beta = 0 \\ \phi(-z\beta)/(z\beta^2) & \beta < 0, \end{cases}$$

and

$$\Psi_\beta(z) = \begin{cases} \phi(z\beta)/(z\beta^2) & \beta > 0 \\ z/2 & \beta = 0 \\ \phi(-z\beta)/(z\beta^2) & \beta < 0, \end{cases}$$

where ϕ and ψ are as in Proposition 3.1. Note that we will frequently use the fact that these functions are non-decreasing.

Theorem 3.2. *Consider the Poisson functional F and assume that one of the following conditions is satisfied:*

(i) *$\beta > 0$ and $D_x F(\xi) < \beta$ holds for all $(x, \xi) \in \mathbb{X} \times \mathbf{N}$;*
(ii) *$\beta < 0$ and $D_x F(\xi) > \beta$ holds for all $(x, \xi) \in \mathbb{X} \times \mathbf{N}$;*
(iii) *$\beta = 0$.*

Let $\lambda > 0$ be such that $\mathbb{E}\exp(\lambda F) < \infty$ and assume that F is integrable. Then for any $\theta > 0$ satisfying $\Phi_\beta(\lambda)\theta < 1$, we have

$$(3.6) \qquad \log\mathbb{E}[\exp(\lambda(F - \mathbb{E}F))] \le \frac{\Psi_\beta(\lambda)\theta}{1 - \Phi_\beta(\lambda)\theta} \log\mathbb{E}\left[\exp\left(\frac{\lambda V_\beta^+}{\theta}\right)\right].$$

Let $\lambda > 0$ be such that $\mathbb{E}\exp(-\lambda F) < \infty$ and assume that F is integrable. Then for any $\theta > 0$ with $\Phi_{-\beta}(\lambda)\theta < 1$, we have

$$(3.7) \qquad \log\mathbb{E}[\exp(-\lambda(F - \mathbb{E}F))] \le \frac{\Psi_{-\beta}(\lambda)\theta}{1 - \Phi_{-\beta}(\lambda)\theta} \log\mathbb{E}\left[\exp\left(\frac{\lambda V_\beta^-}{\theta}\right)\right].$$

In addition to this, the relation $\mathbb{E}\exp(\lambda V_\beta^+/\theta) < \infty$ implies that $\mathbb{E}|F| < \infty$ and $\mathbb{E}\exp(\lambda F) < \infty$. Also, the relation $\mathbb{E}\exp(\lambda V_\beta^-/\theta) < \infty$ implies that $\mathbb{E}|F| < \infty$ and $\mathbb{E}\exp(-\lambda F) < \infty$.

With the above results the methods for deriving deviation inequalities presented in [4] naturally carry over to the Poisson space. In the following we present some variations of these techniques that will be used for the applications later on.

In the case where V_β^+ and V_β^- for $\beta \ne 0$ are almost surely bounded by a constant, the above entropy inequalities yield exponential tails for the random variable $F(\eta)$. If V^+ and V^- are almost surely bounded (i.e. $\beta = 0$), we even obtain Gaussian tails. Note that Wu's deviation inequality [43, Proposition 3.1] is implied by the following more general results. Indeed, in the case where $DF < \beta$ for some $\beta > 0$, one has $V_\beta^+ = \int_{\mathbb{X}}(D_x F)^2 d\mu(x)$, hence the upcoming Corollary 3.3 specializes to [43, Proposition 3.1] in this case.

Corollary 3.3. *Assume that F satisfies $V_\beta^+ \le c$ almost surely. Then F is integrable and the following statements hold:*

(i) *If either condition (i) or (ii) of Theorem 3.2 is satisfied, then for all $r \ge 0$,*

$$\mathbb{P}(F \ge \mathbb{E}F + r) \le \exp\left(-\left(\frac{c}{\beta^2} + \frac{r}{|\beta|}\right)\log\left(1 + \frac{|\beta|r}{c}\right) + \frac{r}{|\beta|}\right)$$
$$\le \exp\left(-\frac{r}{2|\beta|}\log\left(1 + \frac{|\beta|r}{c}\right)\right).$$

(ii) *If $\beta = 0$, that is if $V^+ \le c$ holds almost surely, then for all $r \ge 0$,*

$$\mathbb{P}(F \ge \mathbb{E}F + r) \le \exp\left(-\frac{r^2}{2c}\right).$$

We continue with the corresponding version for the lower tail. This corollary is obtained in the same way as the above one where inequality (3.7) is used instead of (3.6). The proof is therefore omitted.

Corollary 3.4. *Assume that F satisfies $V_\beta^- \le c$ almost surely. Then F is integrable and the following statements hold:*

(i) *If either condition (i) or (ii) of Theorem 3.2 is satisfied, then for all $r \ge 0$,*

$$\mathbb{P}(F \le \mathbb{E}F - r) \le \exp\left(-\left(\frac{c}{\beta^2} + \frac{r}{|\beta|}\right)\log\left(1 + \frac{|\beta|r}{c}\right) + \frac{r}{|\beta|}\right)$$

$$\le \exp\left(-\frac{r}{2|\beta|}\log\left(1 + \frac{|\beta|r}{c}\right)\right).$$

(ii) *If $\beta = 0$, that is if $V^- \le c$ holds almost surely, then for all $r \ge 0$,*

$$\mathbb{P}(F \le \mathbb{E}F - r) \le \exp\left(-\frac{r^2}{2c}\right).$$

The following result is useful to obtain deviation inequalities under less restrictive boundedness conditions on V^+. This is an adaptation for Poisson processes that corresponds to a combination of the results [4, Theorem 8 and Theorem 9].

Corollary 3.5. *Assume that $F \ge 0$ and that there is a random variable $G \ge 0$ and an $\alpha \in [0, 2)$ such that almost surely*

$$V^+ \le GF^\alpha.$$

Let $\theta > 0$ and $\lambda \in (0, 2/\theta)$ be such that $\mathbb{E}\exp(\lambda G/\theta) < \infty$. Then $\mathbb{E}F^{1-\alpha/2} < \infty$ and

$$\log\mathbb{E}(\exp(\lambda(F^{1-\alpha/2} - \mathbb{E}F^{1-\alpha/2}))) \le \frac{\lambda\theta}{2 - \lambda\theta}\log\mathbb{E}\left(\exp\left(\frac{\lambda G}{\theta}\right)\right).$$

In the case where the random variable G in the above corollary is just a constant, we obtain the following deviation inequality for the upper tail.

Corollary 3.6. *Assume that $F \ge 0$ and that for some $\alpha \in [0, 2)$ and $c > 0$ we have almost surely*

$$V^+ \le cF^\alpha.$$

Then F is integrable and for all $r \ge 0$,

$$\mathbb{P}(F \ge \mathbb{E}F + r) \le \exp\left(-\frac{((r + \mathbb{E}F)^{1-\alpha/2} - (\mathbb{E}F)^{1-\alpha/2})^2}{2c}\right).$$

The next result is a variation of Corollary 3.5 for Poisson functionals that are not necessarily non-negative. This is the Poisson space analogue of [4, Theorem 5].

Theorem 3.7. *Assume that the Poisson functional F is integrable and that for some $a > 0$ and $b \ge 0$ we have almost surely*

$$V^+ \le aF + b.$$

Then for any $\lambda \in (0, 2/a)$ we have $\mathbb{E}(e^{\lambda F}) < \infty$ and

$$\log\mathbb{E}(\exp(\lambda(F - \mathbb{E}F))) \le \frac{\lambda^2}{2 - a\lambda}(a\mathbb{E}F + b).$$

Moreover, for any $r \geq 0$,

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\frac{r^2}{2a\mathbb{E}F + 2b + ar}\right).$$

We continue with a result that applies whenever F is non-decreasing and V^- is non-decreasing and integrable. In this case, the random variable F has a Gaussian lower tail:

Theorem 3.8. *Assume that the Poisson functional F satisfies*

$$D_x F(\xi), \ D_x V^-(\xi) \geq 0 \ \ \text{for all} \ \ (x, \xi) \in \mathbb{X} \times \mathbf{N} \ \ \text{and} \ \ \mathbb{E}V^-(\eta) < \infty.$$

Then $F = F(\eta)$ is integrable and for all $r \geq 0$ we have

$$\mathbb{P}(F \leq \mathbb{E}F - r) \leq \exp\left(-\frac{r^2}{2\mathbb{E}V^-}\right).$$

Remark 3.9. Assume that the Poisson functional F is non-decreasing. A sufficient condition for the assumption $DV^- \geq 0$ in the above theorem is that the second iteration of the difference operator of F is non-negative. Indeed, assume that for every $(x, \xi) \in \mathbb{X} \times \mathbf{N}$,

$$D_z D_x F(\xi) \geq 0 \ \ \text{for } \mu\text{-almost every } z \in \mathbb{X}.$$

Then, since obviously $D_z D_x F(\xi) = D_x D_z F(\xi)$, one has $D_z F(\xi + \delta_x) \geq D_z F(\xi) \geq 0$ and hence $D_z F(\xi + \delta_x)^2 \geq D_z F(\xi)^2$ for all $(x, \xi) \in \mathbb{X} \times \mathbf{N}$ and μ-a.e. $z \in \mathbb{X}$. So we see that also $D(DF)^2$ is non-negative, thus yielding that for every $(x, \xi) \in \mathbb{X} \times \mathbf{N}$,

$$D_x V^-(\xi) = D_x \int_{\mathbb{X}} (D_z F(\xi))^2 d\mu(z) = \int_{\mathbb{X}} D_x (D_z F(\xi))^2 d\mu(z) \geq 0.$$

We conclude this section with a result that deals with the situation when F is non-decreasing and the difference operator DF is bounded. In this case, we obtain a deviation inequality for the lower tail by controlling the random variable V^+. This is a Poisson space analogue of [26, Theorem 13].

Theorem 3.10. *Assume that $F \geq 0$ and that for some $a > 0$ we have*

$$0 \leq D_x F(\xi) \leq 1 \ \ \text{for any} \ \ (x, \xi) \in \mathbb{X} \times \mathbf{N} \ \ \text{and almost surely} \ \ V^+(\eta) \leq aF(\eta).$$

Then F is integrable and for any $r \geq 0$ we have

$$\mathbb{P}(F \leq \mathbb{E}F - r) \leq \exp\left(-\frac{r^2}{2\max(a, 1)\mathbb{E}F}\right).$$

4. Proofs

We begin with the proof of the crucial logarithmic Sobolev type inequality, namely Proposition 3.1, that is the foundation of our techniques.

Proof of Proposition 3.1. By (1.1), the inequality

$$\text{Ent}(e^{\lambda F}) \leq \mathbb{E}\left[e^{\lambda F}\int_{\mathbb{X}}\psi(\lambda D_x F)\,d\mu(x)\right]$$

holds. Now, since $\psi(0) = 0$, $\phi(0) = 0$ and $\psi(z) = e^z\phi(-z)$ for any $z \in \mathbb{R}$, we have

$$\psi(\lambda D_x F(\eta)) = \psi(\lambda D_x^I F(\eta)) + e^{\lambda D_x F(\eta)}\phi(-\lambda D_x^{I^c} F(\eta)).$$

Hence, we compute

$$\begin{aligned}\text{Ent}(e^{\lambda F}) &\leq \mathbb{E}\int_{\mathbb{X}} e^{\lambda F}\psi(\lambda D_x^I F)\,d\mu(x) + \mathbb{E}\int_{\mathbb{X}} e^{\lambda D_x F+\lambda F}\phi(-\lambda D_x^{I^c} F)\,d\mu(x)\\ &= \mathbb{E}\int_{\mathbb{X}} e^{\lambda F}\psi(\lambda D_x^I F)\,d\mu(x) + \mathbb{E}\int_{\mathbb{X}} e^{\lambda F(\eta+\delta_x)}\phi(-\lambda D_x^{I^c} F)\,d\mu(x)\\ &= \mathbb{E}\int_{\mathbb{X}} e^{\lambda F}\psi(\lambda D_x^I F)\,d\mu(x) + \mathbb{E}\int_{\mathbb{X}} e^{\lambda F}\phi(-\lambda D_x^{I^c} F(\eta-\delta_x))d\eta(x).\end{aligned}$$

At this, the last equality holds by the Mecke formula (2.5). □

The following lemma will be used occasionally in the upcoming proofs.

Lemma 4.1. *Assume that for some $\beta \in \mathbb{R}$ we have $\mathbb{E}V_\beta^+ < \infty$ or $\mathbb{E}V_\beta^- < \infty$. Then F is integrable.*

Proof. The proof uses a truncation argument that is standard in this context, see e.g. the proof of [43, Proposition 3.1]. The statement for V_β^- is proved in the same way than for V_β^+. Consider for any $n \in \mathbb{N}$ the truncation

$$F_n = \min(\max(F, -n), n).$$

Then $\mathbb{E}(F_n)^2 < \infty$, hence the Poincaré inequality for Poisson point processes (see e.g. [43, Remark 1.4]) together with the Mecke formula (2.5) yield

$$\mathbb{V}F_n \leq \mathbb{E}\int_{\mathbb{X}}(D_x F_n(\eta))^2 d\mu(x) \leq \mathbb{E}\int_{\mathbb{X}}(D_x F(\eta))^2 d\mu(x) = \mathbb{E}V_\beta^+.$$

Therefore, $\sup_{n\in\mathbb{N}}\mathbb{V}F_n \leq \mathbb{E}V_\beta^+ < \infty$. Now, if there is a subsequence $\{F_{n_k}\}_{k\in\mathbb{N}}$ satisfying $\lim_{k\to\infty}\mathbb{E}F_{n_k} = \pm\infty$, then $F_{n_k} \to \pm\infty$ in probability. This would be a contradiction to $F_n \to F$ in probability. We see that the family $\{\mathbb{E}F_n\}_{n\in\mathbb{N}}$ is bounded. Together with $\sup_{n\in\mathbb{N}}\mathbb{V}F_n < \infty$ this implies also $\sup_{n\in\mathbb{N}}\mathbb{E}(F_n^2) < \infty$. Thus, the family $\{F_n\}_{n\in\mathbb{N}}$ is uniformly integrable. In particular, as desired, we have $\mathbb{E}|F| < \infty$. □

We continue with the proof of Theorem 3.2. As in the proof of the product space version [4, Theorem 2], we also need [25, Lemma 11]. This result states that for any $\lambda > 0$ and any two random variables X and Y satisfying $\mathbb{E}(e^{\lambda X}), \mathbb{E}(e^{\lambda Y}) < \infty$, we have

$$(4.8)\qquad \frac{\lambda\mathbb{E}(Xe^{\lambda Y})}{\mathbb{E}(e^{\lambda Y})} \leq \frac{\lambda\mathbb{E}(Ye^{\lambda Y})}{\mathbb{E}(e^{\lambda Y})} + \log\mathbb{E}(e^{\lambda X}) - \log\mathbb{E}(e^{\lambda Y}).$$

Proof of Theorem 3.2. We prove (3.6). We only deal with the case $\beta \neq 0$, whereas the case $\beta = 0$ can be obtained with by similar arguments. To prove the desired inequality we adapt the proof of [4, Theorem 2] and combine this with arguments from the proof of [43, Proposition 3.1]. We first consider the case where F is bounded. Let ϕ and ψ be as in Proposition 3.1. Then $\psi(z)/z^2$ and $\phi(z)/z^2$ are non-decreasing. Hence, for any $u \in (0, \lambda]$ we have

$$\psi(uz) \leq (\psi(u\beta)/\beta^2)z^2 \leq u\Phi_\beta(u)z^2 \text{ for } z \leq \beta,$$
$$\phi(-uz) \leq (\phi(-u\beta)/\beta^2)z^2 \leq u\Phi_\beta(u)z^2 \text{ for } z \geq \beta.$$

Together with $\psi(0) = \phi(0) = 0$ this gives

$$\psi(uD_x^{\leq \beta}F(\eta)) \leq u\Phi_\beta(u)(D_x^{\leq \beta}F(\eta))^2,$$
$$\phi(-uD_x^{>\beta}F(\eta - \delta_x)) \leq u\Phi_\beta(u)(D_x^{>\beta}F(\eta - \delta_x))^2.$$

Hence, taking $I = \{(x, \xi) \in \mathbb{X} \times \mathbf{N} : D_xF(\xi) \leq \beta\}$, it follows from Proposition 3.1 that

$$\text{Ent}(e^{uF}) \leq \mathbb{E}\left[e^{uF}\left(\int_\mathbb{X} \psi(uD_x^{\leq \beta}F(\eta))\, d\mu(x) + \int_\mathbb{X} \phi(-uD_x^{>\beta}F(\eta - \delta_x))d\eta(x)\right)\right]$$
$$\leq u\Phi_\beta(u)\, \mathbb{E}\left[e^{uF}\left(\int_\mathbb{X} (D_x^{\leq \beta}F(\eta))^2\, d\mu(x) + \int_\mathbb{X}(D_x^{>\beta}F(\eta - \delta_x))^2 d\eta(x)\right)\right]$$
$$= u\Phi_\beta(u)\, \mathbb{E}(V_\beta^+ e^{uF}).$$

Moreover, taking $X = V_\beta^+/\theta$ and $Y = F$, it follows from (4.8) that

$$\frac{\mathbb{E}(V_\beta^+ e^{uF})}{\mathbb{E}(e^{uF})} \leq \frac{\theta\mathbb{E}(Fe^{uF})}{\mathbb{E}(e^{uF})} + \frac{\theta}{u}\log\mathbb{E}(e^{uV_\beta^+/\theta}) - \frac{\theta}{u}\log\mathbb{E}(e^{uF}).$$

Invoking the definition of the entropy, it follows from the last two displays that

$$\frac{\mathbb{E}(Fe^{uF})}{u\mathbb{E}(e^{uF})} - \frac{\log\mathbb{E}(e^{uF})}{u^2} \leq \Phi_\beta(u)\left(\frac{\theta\mathbb{E}(Fe^{uF})}{u\mathbb{E}(e^{uF})} + \frac{\theta}{u^2}\log\mathbb{E}(e^{uV_\beta^+/\theta}) - \frac{\theta}{u^2}\log\mathbb{E}(e^{uF})\right).$$

Since by assumption $\Phi_\beta(u)\theta \leq \Phi_\beta(\lambda)\theta < 1$, the latter inequality is equivalent to

$$\frac{\mathbb{E}(Fe^{uF})}{u\mathbb{E}(e^{uF})} - \frac{\log\mathbb{E}(e^{uF})}{u^2} \leq \frac{\Phi_\beta(u)\theta\log\mathbb{E}(e^{uV_\beta^+/\theta})}{u^2(1 - \Phi_\beta(u)\theta)}.$$

Defining $h(u) = \frac{1}{u}\log\mathbb{E}(e^{uF})$ and $g(u) = \log\mathbb{E}(e^{uV_\beta^+})$, the above estimate can be restated as follows:

$$h'(u) \leq \frac{\Phi_\beta(u)\theta g(u/\theta)}{u^2(1 - \Phi_\beta(u)\theta)}$$

for any $u \in (0, \lambda]$. Since $\lim_{u \to 0+} h(u) = \mathbb{E}F$, integration from 0 to λ gives

(4.9) $$h(\lambda) \leq \mathbb{E}F + \int_0^\lambda \frac{\Phi_\beta(u)\theta g(u/\theta)}{u^2(1 - \Phi_\beta(u)\theta)}\, du.$$

It's a well known fact that the logarithm of a moment generating function is convex, hence g is convex on the interval $[0, \lambda/\theta]$. In particular, we have for any $u \in (0, \lambda]$ that

$$\frac{g(u/\theta)}{u(1 - \Phi_\beta(u)\theta)} \leq \frac{\frac{g(\lambda/\theta)}{\lambda} u}{u(1 - \Phi_\beta(u)\theta)} \leq \frac{g(\lambda/\theta)}{\lambda(1 - \Phi_\beta(\lambda)\theta)}.$$

Hence,

$$\int_0^\lambda \frac{\Phi_\beta(u)\theta g(u/\theta)}{u^2(1 - \Phi_\beta(u)\theta)} \, du \leq \frac{\theta g(\lambda/\theta)}{\lambda(1 - \Phi_\beta(\lambda)\theta)} \int_0^\lambda \frac{\Phi_\beta(u)}{u} \, du.$$

In the case $\beta < 0$, we bound the integral on the right-hand side by $\Phi_\beta(\lambda) = \Psi_\beta(\lambda)$. This works since $\Phi_\beta(u)/u$ is non-decreasing. In the case $\beta > 0$, the integral can be explicitly computed and one obtains $\int_0^\lambda (\Phi_\beta(u)/u)du = \phi(\lambda\beta)/(\lambda\beta^2) = \Psi_\beta(\lambda)$. Combining this with (4.9) gives

$$\log \mathbb{E}(e^{\lambda F}) \leq \lambda \mathbb{E}F + \frac{\Psi_\beta(\lambda)\theta}{1 - \Phi_\beta(\lambda)\theta} \, g(\lambda/\theta).$$

This proves inequality (3.6) for bounded F. Applying (3.6) to the Poisson functional $-F$ and observing that $V_\beta^+(-F) = V_{-\beta}^-(F)$ yields (3.7) for bounded F.

Now assume that F is not necessarily bounded and consider for $n \in \mathbb{N}$ the truncated random variables

$$F_n = \min(\max(F, -n), n).$$

We will now conclude that if $\mathbb{E}\exp(\lambda V_\beta^+/\theta) < \infty$, then F is integrable and the family of random variables

(4.10) $$\{\exp(\lambda(F_n - \mathbb{E}F_n))\}_{n \in \mathbb{N}}$$

converges in probability to $\exp(\lambda(F - \mathbb{E}F))$ and is uniformly integrable. Thus, it follows that $\lim_{n \to \infty} \mathbb{E}\exp(\lambda(F_n - \mathbb{E}F_n)) = \mathbb{E}\exp(\lambda(F - \mathbb{E}F)) < \infty$ and hence also $\mathbb{E}\exp(\lambda(F)) < \infty$. This together with the observation $V_\beta^+(F_n) \leq V_\beta^+$ and the fact that we already proved (3.6) for the bounded random variables F_n then yields the result for the unbounded F.

Integrability of F follows from Lemma 4.1 since the assumption $\mathbb{E}(\exp(\lambda V_\beta^+/\theta)) < \infty$ implies that $\mathbb{E}V_\beta^+ < \infty$. By dominated convergence, integrability of F now implies the convergence in probability of the sequence in (4.10).

To prove the uniform integrability, first observe, that if (i), (ii) or (iii) holds, then

$$V_\beta^+(F_n) \leq V_\beta^+(F) = V_\beta^+.$$

Also note that we can choose $\nu > 1$ such that $\Phi_\beta(\nu\lambda)\nu\theta < 1$. Then $\mathbb{E}(e^{\lambda\nu F_n}) < \infty$ for all $n \in \mathbb{N}$, so it follows from (3.6) that

$$\log \mathbb{E}(\exp(\lambda\nu(F_n - \mathbb{E}F_n))) \leq \frac{\Psi_\beta(\nu\lambda)\nu\theta}{1 - \Phi_\beta(\nu\lambda)\nu\theta} \log \mathbb{E}\left(\exp\left(\frac{\lambda V_\beta^+(F_n)}{\theta}\right)\right)$$

$$\leq \frac{\Psi_\beta(\nu\lambda)\nu\theta}{1 - \Phi_\beta(\nu\lambda)\nu\theta} \log \mathbb{E}\left(\exp\left(\frac{\lambda V_\beta^+}{\theta}\right)\right) < \infty.$$

Denoting the map $x \mapsto x^\nu$ by Λ, the above inequality yields

$$\sup_{n\in\mathbb{N}} \mathbb{E}[\Lambda(\exp(\lambda(F_n - \mathbb{E}F_n)))] < \infty.$$

By the Theorem of de la Vallée-Poussin this implies uniform integrability of the family in (4.10).

Repeating the above reasoning for $-F$ instead of F, we can also extend inequality (3.7) to the case where F is unbounded. $\qquad\square$

Proof of Corollary 3.3. It follows from Lemma 4.1 that F is integrable. Theorem 3.2 yields

$$\log \mathbb{E}(e^{\lambda(F-\mathbb{E}F)}) \leq \inf_{\theta\in(0,1/\Phi_\beta(\lambda))} \frac{\Psi_\beta(\lambda)\lambda c}{1 - \Phi_\beta(\lambda)\theta} = \Psi_\beta(\lambda)\lambda c.$$

Markov's inequality now gives for any $\lambda > 0$,

$$\mathbb{P}(F \geq \mathbb{E}F + r) = \mathbb{P}(e^{\lambda(F-\mathbb{E}F)} \geq e^{\lambda r}) \leq \frac{\mathbb{E}(e^{\lambda(F-\mathbb{E}F)})}{e^{\lambda r}} \leq \exp\left(\Psi_\beta(\lambda)\lambda c - \lambda r\right).$$

Optimizing in λ yields the desired deviation bounds. $\qquad\square$

Proof of Corollary 3.5. Here we adapt and combine the proofs of [4, Theorem 8 and Theorem 9]. For $\alpha = 0$, the statement follows directly from Theorem 3.2, so let $\alpha \in (0, 2)$. Let $\gamma = 1 - \alpha/2$. Then, on the event $\{F \neq 0\}$, we have

$$\int_{\mathbb{X}} (D_x^+ F^\gamma(\eta - \delta_x))^2 d\eta(x)$$

$$= \int_{\mathbb{X}} \mathbb{1}\{F(\eta)^\gamma \geq F(\eta - \delta_x)^\gamma > 0\}(F(\eta)^\gamma - F(\eta - \delta_x)^\gamma)^2 d\eta(x)$$

$$+ \int_{\mathbb{X}} \mathbb{1}\{F(\eta)^\gamma \geq F(\eta - \delta_x)^\gamma = 0\}F(\eta)^{2\gamma} d\eta(x)$$

$$= \int_{\mathbb{X}} \mathbb{1}\{F(\eta) \geq F(\eta - \delta_x) > 0\}\left(\frac{F(\eta)}{F(\eta)^{1-\gamma}} - \frac{F(\eta - \delta_x)}{F(\eta - \delta_x)^{1-\gamma}}\right)^2 d\eta(x)$$

$$+ \int_{\mathbb{X}} \mathbb{1}\{F(\eta) \geq F(\eta - \delta_x) = 0\}F(\eta)^{2\gamma} d\eta(x).$$

Since $1 - \gamma > 0$, we have that $F(\eta) \geq F(\eta - \delta_x)$ implies $F(\eta)^{1-\gamma} \geq F(\eta - \delta_x)^{1-\gamma}$. Hence, the above expression does not exceed

$$
\int_{\mathbb{X}} \mathbb{1}\{F(\eta) \geq F(\eta - \delta_x) > 0\} \left(\frac{F(\eta)}{F(\eta)^{1-\gamma}} - \frac{F(\eta - \delta_x)}{F(\eta)^{1-\gamma}} \right)^2 d\eta(x)
$$

$$
+ \int_{\mathbb{X}} \mathbb{1}\{F(\eta) \geq F(\eta - \delta_x) = 0\} F(\eta)^{2\gamma} d\eta(x)
$$

$$
= \frac{1}{F(\eta)^\alpha} \int_{\mathbb{X}} \left(D_x^+ F(\eta - \delta_x) \right)^2 d\eta(x).
$$

Quite similarly one obtains that on the event $\{F \neq 0\}$,

$$
\int_{\mathbb{X}} (D_x^- F^\gamma(\eta))^2 \, d\mu(x) \leq \frac{1}{F(\eta)^\alpha} \int_{\mathbb{X}} (D_x^- F(\eta))^2 \, d\mu(x).
$$

Hence, it follows that on the event $\{F \neq 0, V^+ \leq GF^\alpha\}$,

$$
V^+(F^\gamma) = \int_{\mathbb{X}} (D_x^- F^\gamma(\eta))^2 \, d\mu(x) + \int_{\mathbb{X}} (D_x^+ F^\gamma(\eta - \delta_x))^2 d\eta(x) \leq \frac{V^+}{F(\eta)^\alpha} \leq G.
$$

Moreover, it is easy to check that on the event $\{F = 0, V^+ \leq GF^\alpha\}$, one has that $V^+(F^\gamma) = 0 = V^+$. Therefore, by virtue of the assumption that almost surely $V^+ \leq GF^\alpha$, it follows that almost surely $V^+(F^\gamma) \leq G$. Applying Theorem 3.2 to the random variable F^γ yields the result. $\square$

Proof of Corollary 3.6. For $\alpha = 0$, the statement follows directly from Corollary 3.3 (ii), so let $\alpha \in (0, 2)$. Let $\gamma = 1 - \alpha/2$. Continuing in the same way as in the proof of Corollary 3.5 yields that almost surely

$$
V^+(F^\gamma) = \int_{\mathbb{X}} (D_x^- F^\gamma(\eta))^2 \, d\mu(x) + \int_{\mathbb{X}} (D_x^+ F^\gamma(\eta - \delta_x))^2 d\eta(x) \leq c.
$$

We conclude that Corollary 3.3 (ii) applies to F^γ. So F^γ is non-negative and has an exponentially decaying upper tail. Thus, by virtue of [18, Lemma 3.4], all moments of F^γ exist. In particular, F is integrable. As it was pointed out in [4, p. 1588], we can now write

$$
\mathbb{P}(F \geq \mathbb{E}F + r) = \mathbb{P}(F^\gamma \geq (r + \mathbb{E}F)^\gamma) \leq \mathbb{P}(F^\gamma - \mathbb{E}(F^\gamma) \geq (r + \mathbb{E}F)^\gamma - (\mathbb{E}F)^\gamma)
$$

$$
\leq \exp \left(-\frac{((r + \mathbb{E}F)^\gamma - (\mathbb{E}F)^\gamma)^2}{2c} \right).
$$

$\square$

We continue with the proof of Theorem 3.7. To get prepared for this, we first establish the following lemma.

Lemma 4.2. *Let $n \in \mathbb{N}$ and consider $F_n = \min(\max(F, -n), n)$ and $V_{(n)}^+ = V^+(F_n)$. Then for any real number $b \geq 0$, almost surely*

$$F(V_{(n)}^+ - b) \leq F_n(V^+ - b) \quad if \ \ F, F_n \geq 0,$$
$$F(V_{(n)}^+ - b) \geq F_n(V^+ - b) \quad if \ \ F, F_n \leq 0.$$

Proof. It is easy to see that $V_{(n)}^+ \leq V^+$. Hence, the desired statement holds on the event $\{F = F_n\}$. If $F \neq F_n$, then either $F_n = n < F$ or $F_n = -n > F$. The latter case implies $V_{(n)}^+ = 0$ and $F, F_n \leq 0$, hence the desired statement holds. So consider the case $F_n = n < F$ and let $A = F/n$. Then the desired inequality is equivalent to

$$AV_{(n)}^+ \leq V^+ + (A - 1)b.$$

Since $b \geq 0$ and $A > 1$, the above inequality is implied by $AV_{(n)}^+ \leq V^+$, i.e. by

$$A\left(\int_{\mathbb{X}} (n - F_n(\eta - \delta_x))_+^2 d\eta(x) + \int_{\mathbb{X}} (F_n(\eta + \delta_x) - n)_-^2 d\mu(x)\right)$$
$$\leq \int_{\mathbb{X}} (An - F(\eta - \delta_x))_+^2 d\eta(x) + \int_{\mathbb{X}} (F(\eta + \delta_x) - An)_-^2 d\mu(x).$$

To prove this, it suffices to conclude

(4.11) $$A(n - F_n(\eta - \delta_x))_+^2 \leq (An - F(\eta - \delta_x))_+^2,$$

(4.12) $$A(F_n(\eta + \delta_x) - n)_-^2 \leq (F(\eta + \delta_x) - An)_-^2.$$

We prove (4.11). If $F(\eta - \delta_x) > n$, then

$$A(n - F_n(\eta - \delta_x))_+^2 = 0 \leq (An - F(\eta - \delta_x))_+^2.$$

If $F(\eta - \delta_x) \leq n$, then

$$F(\eta - \delta_x) \leq F_n(\eta - \delta_x) =: m.$$

Now, since $|m| \leq n$, we have $An^2 - m^2 \geq 0$. This gives $(A^2 - A)n^2 + (1 - A)m^2 \geq 0$, thus

$$(An - m)^2 = A^2n^2 - 2Anm + m^2 \geq An^2 - 2Anm + Am^2 = A(n - m)^2.$$

Hence,

$$A(n - F_n(\eta - \delta_x))_+^2 \leq (An - F_n(\eta - \delta_x))_+^2 \leq (An - F(\eta - \delta_x))_+^2,$$

where the last inequality follows from $F(\eta - \delta_x) \leq F_n(\eta - \delta_x) < An$. This proves (4.11) and analogously one obtains (4.12). The result follows. $\square$

Proof of Theorem 3.7. For the case where F is bounded, we adapt the proof of [4, Theorem 5]. Here we can argue in the same way as in the beginning of the proof of Theorem 3.2 to obtain for any $u \in (0, \lambda]$,

$$\text{Ent}(e^{uF}) \leq \frac{1}{2}u^2\mathbb{E}(V^+e^{uF}).$$

Invoking the assumption on V^+ yields

$$u\mathbb{E}(Fe^{uF}) - \mathbb{E}(e^{uF})\log\mathbb{E}(e^{uF}) \leq \frac{1}{2}u^2(a\mathbb{E}(Fe^{uF}) + b\mathbb{E}(e^{uF})).$$

With $h(u) = \frac{1}{u}\log\mathbb{E}(e^{uF})$ this can be rearranged as

$$h'(u) \leq \frac{1}{2}(a\log(\mathbb{E}(e^{uF}))' + b).$$

Integrating this from 0 to λ gives

$$h(\lambda) - \mathbb{E}F \leq \frac{1}{2}\left(a\log(\mathbb{E}(e^{\lambda F}) + \lambda b\right).$$

Noting that $a\lambda < 2$ and rearranging the above inequality, we obtain the result for the bounded case.

For the unbounded case, consider for any $n \in \mathbb{N}$ the truncated random variables

$$F_n = \min(\max(F, -n), n).$$

It follows from the assumptions and Lemma 4.2 that almost surely

$$F(V_{(n)}^+ - b) \leq aF_nF \quad \text{if} \quad F \geq 0,$$
$$F(V_{(n)}^+ - b) \geq aF_nF \quad \text{if} \quad F \leq 0.$$

Note also that for $F = 0 = F_n$ we have $V_{(n)}^+ \leq V^+ \leq b$. Therefore, almost surely

$$V_{(n)}^+ \leq aF_n + b,$$

so the result holds for all F_n. By dominated convergence, the sequence $\mathbb{E}F_n$ is convergent, hence bounded above by some constant C. Moreover, we can choose a $\nu > 1$ such that $\nu\lambda < 2/a$. Thus, since we already proved that the result applies to all the F_n, we conclude

$$\sup_{n\in\mathbb{N}} \mathbb{E}[\exp(\lambda(F_n - \mathbb{E}F_n))^\nu] \leq \exp\left(\frac{\nu^2\lambda^2}{2 - a\nu\lambda}(aC + b)\right) < \infty.$$

By the Theorem of de la Vallée-Poussin this implies that the family of random variables

$$\{\exp(\lambda(F_n - \mathbb{E}F_n))\}_{n\in\mathbb{N}}$$

is uniformly integrable. Continuing as in the proof of Theorem 3.2 gives

$$\lim_{n\to\infty} \mathbb{E}\exp(\lambda(F_n - \mathbb{E}F_n)) = \mathbb{E}\exp(\lambda(F - \mathbb{E}F)) < \infty.$$

We note again that the result is already proved for the F_n and that $\mathbb{E}F_n \to \mathbb{E}F$ as $n \to \infty$. This concludes the proof of the first inequality.

The deviation inequality now follows using the inequality we just proved together with Markov's inequality and [4, Lemma 11]. $\qquad\square$

For the proof of Theorem 3.8 we use the following FKG inequality for Poisson point processes, taken from [17, Lemma 2.1], see also [43, Remark 1.5] and [22, Theorem 1.4].

Lemma 4.3. *Let F and G be bounded Poisson functionals and assume that*

$$D_x F(\xi), D_x G(\xi) \geq 0 \quad \text{for all} \quad (x, \xi) \in \mathbb{X} \times \mathbf{N}.$$

Then

$$\mathbb{E}(FG) \geq (\mathbb{E}F)(\mathbb{E}G).$$

It was also remarked in [17] that under conditions like $F, G \geq 0$ or $\mathbb{E}F^2, \mathbb{E}G^2 < \infty$, the above result easily extends to unbounded functionals by monotone convergence. For our purpose we need the following extension.

Corollary 4.4. *Let $F, G \geq 0$ be Poisson functionals. Assume that F is bounded and G is integrable. Moreover, assume that*

$$D_x F(\xi) \leq 0 \quad \text{and} \quad D_x G(\xi) \geq 0 \quad \text{for all} \quad (x, \xi) \in \mathbb{X} \times \mathbf{N}.$$

Then

$$\mathbb{E}(FG) \leq (\mathbb{E}F)(\mathbb{E}G).$$

Proof. Since F is bounded, it follows from $\mathbb{E}G < \infty$ that also $\mathbb{E}(FG) < \infty$. Now consider for any $n \in \mathbb{N}$ the truncations $G_n = \min(G, n)$. Then we have almost surely $G_n \to G$ and $FG_n \to FG$ as $n \to \infty$. By monotone convergence, $\mathbb{E}G_n \to \mathbb{E}G$ and $\mathbb{E}(FG_n) \to \mathbb{E}(FG)$ as $n \to \infty$. It follows from Lemma 4.3 that for any $n \in \mathbb{N}$,

$$\mathbb{E}(FG_n) \leq (\mathbb{E}F)(\mathbb{E}G_n).$$

The result follows. $\qquad\square$

The following proof is inspired by ideas from the proof of [4, Theorem 6].

Proof of Theorem 3.8. For any $n \in \mathbb{N}$ consider the truncations

$$F_n = \min(\max(F, -n), n).$$

Then the F_n are again non-decreasing. Let $\lambda < 0$. It follows from Proposition 3.1 with $I = \mathbb{X} \times \mathbf{N}$ that for any $u \in [\lambda, 0)$ we have

$$\text{Ent}(e^{uF_n}) \leq \mathbb{E}\left[e^{uF_n} \int_{\mathbb{X}} \psi(uD_x F_n) \, d\mu(x) \right].$$

Since $\psi(-z) \leq (1/2)z^2$ for $z \geq 0$, the right-hand side of the above expression does not exceed

$$\frac{1}{2}\mathbb{E}\left[e^{uF_n} \int_{\mathbb{X}} (uD_x F_n)^2 \, d\mu(x) \right] = \frac{1}{2}u^2 \, \mathbb{E}(e^{uF_n} V^-(F_n)).$$

We have $V^-(F_n) \leq V^-$ almost surely, hence $\mathbb{E}(e^{uF_n} V^-(F_n))$ in the above display can be upper bounded by $\mathbb{E}(e^{uF_n} V^-)$. Now, since F_n is non-decreasing and $u < 0$, the functional e^{uF_n} is non-increasing and bounded. Moreover, by assumption the functional V^- is non-decreasing and $\mathbb{E}V^- < \infty$. Hence, by Corollary 4.4 we have

$$\mathbb{E}(e^{uF_n} V^-) \leq \mathbb{E}(e^{uF_n}) \, \mathbb{E}V^-.$$

It follows that

$$h'(u) \leq \frac{1}{2}\mathbb{E}V^- \quad \text{where} \quad h(u) = \frac{1}{u}\log \mathbb{E}(e^{uF_n}).$$

Integrating from λ to 0 yields

$$\log \mathbb{E}[\exp(\lambda(F_n - \mathbb{E}F_n))] \leq \frac{1}{2}\lambda^2 \mathbb{E}V^-.$$

Since $\mathbb{E}V^- < \infty$, by Lemma 4.1 we have $\mathbb{E}|F| < \infty$. Thus, applying the Theorem of de la Vallée-Poussin similarly as in the proof of Theorem 3.2, we conclude that the inequality in the last display also holds for the random variable F. Using Markov's inequality and optimizing in λ yields the result. □

To prove the statement of Theorem 3.10 for bounded F, we adapt the proof of the product space version [26, Theorem 13]. To extend the result to unbounded F, Lemma 4.2 is used similarly as it was done in the proof of Theorem 3.7.

Proof of Theorem 3.10. First consider the case where F is bounded. Let $\lambda < 0$ and $u \in [\lambda, 0)$. Then by Proposition 3.1 with $I = \emptyset$ we have

$$\text{Ent}(e^{uF}) \leq \mathbb{E}\left[e^{uF}\int_{\mathbb{X}}\phi(-uD_xF(\eta - \delta_x))d\eta(x)\right].$$

Moreover, since $0 \leq DF \leq 1$, we have $-uD_xF(\eta - \delta_x) \leq -u$ and since the map $z \mapsto \phi(z)/z^2$ is increasing, this implies

$$\int_{\mathbb{X}}\phi(-uD_xF(\eta - \delta_x))d\eta(x) = u^2\int_{\mathbb{X}}\frac{\phi(-uD_xF(\eta - \delta_x))}{u^2D_xF(\eta - \delta_x)^2}D_xF(\eta - \delta_x)^2 d\eta(x)$$

$$\leq u^2\int_{\mathbb{X}}\frac{\phi(-u)}{u^2}D_xF(\eta - \delta_x)^2 d\eta(x).$$

Now, since $V^+ \leq aF$, we obtain

$$\text{Ent}(e^{uF}) \leq \phi(-u)\mathbb{E}(e^{uF}V^+) \leq \phi(-u)a\mathbb{E}(Fe^{uF}).$$

Dividing by $u^2\mathbb{E}(e^{uF})$ and integrating from λ to 0 yields

$$\mathbb{E}F - \frac{1}{\lambda}\log\mathbb{E}(e^{\lambda F}) \leq -\frac{\phi(-\lambda)}{\lambda^2}a\log\mathbb{E}(e^{\lambda F}).$$

Since $1 - a\phi(-\lambda)/\lambda > 1$, this can be rearranged as

$$\log\mathbb{E}[\exp(\lambda(F - \mathbb{E}F))] \leq \lambda^2\frac{a\phi(-\lambda)/\lambda^2}{1 - a\phi(-\lambda)/\lambda}\mathbb{E}F.$$

Similarly as in the proof of Theorem 3.7, the above inequality can be extended to the case where F is unbounded. Here one should notice that according to Corollary 3.6, the condition $V^+ \leq aF$ guarantees $\mathbb{E}|F| < \infty$. It was pointed out in [26] that for any $\lambda < 0$,

$$\frac{a\phi(-\lambda)/\lambda^2}{1 - a\phi(-\lambda)/\lambda} \leq \frac{\max(a,1)}{2}.$$

Markov's inequality now gives

$$\mathbb{P}(F \leq \mathbb{E}F - r) \leq \mathbb{E}(e^{\lambda(F-\mathbb{E}F)})e^{\lambda r} \leq \exp\left(\lambda^2\frac{\max(a,1)}{2}\mathbb{E}F + \lambda r\right).$$

Optimizing in λ concludes the proof. □

5. Applications to U-Statistics

5.1. **General remarks.** The aim of the present section is to investigate the concentration properties of Poisson U-statistics. For this purpose, we need to specialize the very general framework that was in order so far. Throughout this section, the intensity measure μ on the space $\mathbb{X}$ is assumed to be non-atomic, that is, $\{x\} \in \mathcal{X}$ and $\mu(\{x\}) = 0$, for every $x \in \mathbb{X}$. This assumption is equivalent to the fact that the Poisson process η on $\mathbb{X}$ is *simple*, meaning that almost surely $\eta(\{x\}) \leq 1$ for all $x \in \mathbb{X}$. It is common practice in this setting to identify the simple point process η with its support, which now corresponds to a random set in $\mathbb{X}$. Plainly, the integral of a map $f : \mathbb{X} \to \mathbb{R}$ with respect to η is now exactly given by the (possibly infinite) sum

$$\int_{\mathbb{X}} f \, d\eta = \sum_{x \in \eta} f(x).$$

Remark 5.1. Consider a Poisson functional F together with some representative $\mathfrak{f} : \mathbf{N} \to \mathbb{R}$. For any $\xi \in \mathbf{N}$, we denote by $[\xi]$ the integer-valued measure uniquely determined by its value on singletons via the relation $[\xi](\{x\}) = \mathbb{1}\{\xi(\{x\}) > 0\}$, for all $x \in \mathbb{X}$. Then, since η is simple, we have that almost surely $F = \mathfrak{f}([\eta])$. It follows that another representative of F is given by $\mathfrak{f}' : \mathbf{N} \to \mathbb{R}$, where $\mathfrak{f}'(\xi) = \mathfrak{f}([\xi])$. Therefore, without loss of generality, we can assume that $F(\xi) = F([\xi])$ for all $\xi \in \mathbf{N}$, that is, given an arbitrary functional $\xi \mapsto F(\xi)$, in this section we will systematically select a representative of F that only depends on ξ via the mapping $\xi \mapsto [\xi]$. With this convention, one has that $D_x F(\xi) = D_x F([\xi])$, and also that $D_x F(\xi) = 0$ whenever $\xi(\{x\}) > 0$. Finally we observe that, again by virtue of the above convention and in accordance with the content of Remark 2.1, the fact that the quantity $D_x F(\xi)$ verifies some property $\mathcal{P}$ for every $x \in \mathbb{X}$ and every $\xi \in \mathbf{N}$ is equivalent to the fact that $\mathcal{P}$ is verified for all $(x, \xi) \in \mathbb{X} \times \mathbf{N}$ such that ξ charges each singleton with a mass at most equal to 1.

We now recall some relevant definitions. Let $f : \mathbb{X}^k \to \mathbb{R}_{\geq 0}$ be a symmetric measurable map and define the functional $S_f : \mathbf{N} \to [0, \infty]$ by

$$(5.13) \qquad\qquad S_f(\xi) = \sum_{\mathbf{x} \in \xi_{\neq}^k} f(\mathbf{x}).$$

A *(Poisson) U-statistic* F of order k with kernel f is a random variable such that almost surely $F = S_f(\eta)$. According to the Slyvniak-Mecke formula (2.4), the expectation of a U-statistic F is given by

$$\mathbb{E}F = \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} f(x_1, \ldots, x_k) d\mu(x_1) \cdots d\mu(x_k);$$

see e.g. [31, Section 3] for more details as well as for an introduction to U-statistics with kernels that may have arbitrary sign.

5.2. **Choice of a representative.** In order to apply results from Section 3 to a Poisson U-statistic F with kernel $f \geq 0$, we first need to choose a suitable representative of F as defined in Section 2. Whenever the considered U-statistic F is almost surely finite, we can choose as a representative of F the map $\mathfrak{f} : \mathbf{N} \to \mathbb{R}$ defined by $\mathfrak{f}(\xi) = S_f(\xi)$ if $S_f(\xi) < \infty$ and $\mathfrak{f}(\xi) = 0$ if $S_f(\xi) = \infty$.

In order to avoid technical problems arising from the choice of this representative, we will often assume that a given U-statistic with kernel f is *well-behaved*. By this we mean that there exists a measurable set $B \subseteq \mathbf{N}$ with $\mathbb{P}(\eta \in B) = 1$, such that

(i) $S_f(\xi) < \infty$ for all $\xi \in B$,
(ii) $\xi + \delta_x \in B$ whenever $\xi \in B$ and $x \in \mathbb{X}$,
(iii) $\xi - \delta_x \in B$ whenever $\xi \in B$ and $x \in \xi$.

If F is well-behaved, then we will choose as a representative of F the map $\mathfrak{f} : \mathbf{N} \to \mathbb{R}$ defined by $\mathfrak{f}(\xi) = S_f(\xi)$ if $\xi \in B$ and $\mathfrak{f}(\xi) = 0$ if $\xi \in B^c$. Then, for any $(x, \xi) \in \mathbb{X} \times \mathbf{N}$ one has $D_x F(\xi) = S_f(\xi + \delta_x) - S_f(\xi) < \infty$ if $\xi \in B$ and $D_x F(\xi) = 0$ if $\xi \in B^c$.

Note that by virtue of (2.3) and (5.13), the above choices of a representative imply $F(\xi) = F([\xi])$ for all $\xi \in \mathbf{N}$ which is consistent with Remark 5.1. Finally, note that U-statistics that arise in typical applications (in particular, all U-statistics considered in this paper) are usually well-behaved in the sense described above.

5.3. **General results.** We will use an explicit expression for the difference operator of a U-statistic that was established in [31]. The following result gathers together several results from [31, Lemma 3.3 and Theorem 3.6], in a form that is adapted to our setting.

Proposition 5.2. *Let the above assumptions and notation prevail, let F be a U-statistic with non-negative kernel f and let S_f be as in (5.13). Then, for any $\xi \in \mathbf{N}$ and $x \in \xi$, one has*

$$S_f(\xi) - S_f(\xi - \delta_x) = kF(x, \xi) \quad whenever \quad S_f(\xi) < \infty,$$

where, for any $\xi \in \mathbf{N}$ and every $x \in \xi$ such that $\xi(\{x\}) = 1$, the local version of F is defined as

$$(5.14) \qquad F(x, \xi) := \sum_{\boldsymbol{y} \in (\xi \setminus x)_{\neq}^{k-1}} f(x, \boldsymbol{y}),$$

where $\xi \setminus x$ is shorthand for the set obtained by deleting x from the support of ξ, and $F(x, \xi) = 0$ whenever $\xi(\{x\}) > 1$. Moreover, if $\mathbb{E}F^2 < \infty$, then $f \in L^1(\mu^k) \cap L^2(\mu^k)$.

As a direct consequence of the above result together with our canonical choices of a representative, described in Section 5.2, we obtain:

Corollary 5.3. *Let F be a U-statistic with non-negative kernel f. Then the following statements hold:*

(i) *If F is almost surely finite, then there exists a measurable set $B \subseteq \mathbf{N}$ that satisfies $\mathbb{P}(\eta \in B) = 1$ such that for any $\xi \in B$ and $x \in \xi$, the local version*

$F(x, \xi)$ *is finite and*

$$D_x F(\xi - \delta_x) = kF(x, \xi).$$

(ii) *If F is well-behaved, then there exists a measurable set $B \subseteq \mathbf{N}$ that satisfies $\mathbb{P}(\eta \in B) = 1$ such that the following holds:*

(a) *For any $\xi \in B$, $x \in \xi$ and $z \in \mathbb{X}$, the local versions $F(x, \xi)$ and $F(z, \xi + \delta_z)$ are finite, and moreover*

$$D_x F(\xi - \delta_x) = kF(x, \xi) \quad and \quad D_z F(\xi) = kF(z, \xi + \delta_z).$$

(b) *For any $\xi \in B^c$, $x \in \xi$ and $z \in \mathbb{X}$, one has*

$$D_x F(\xi - \delta_x) = 0 = D_z F(\xi).$$

The previous Corollary 5.3 implies that, if F is an almost surely finite U-statistic with kernel $f \geq 0$, then almost surely

$$V^+ = k^2 \sum_{x \in \eta} F(x, \eta)^2.$$

If F is in addition well-behaved, then almost surely

$$V^- = k^2 \int_{\mathbb{X}} F(x, \eta + \delta_x)^2 d\mu(x).$$

We have therefore the following consequences of Corollary 3.6 and Theorem 3.8.

Corollary 5.4. *Consider an almost surely finite U-statistic F of order k with non-negative kernel f. Assume that for some $\alpha \in [0, 2)$ and $c > 0$ we have almost surely*

$$\sum_{x \in \eta} F(x, \eta)^2 \leq cF^\alpha.$$

Then F is integrable and for all $r > 0$,

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\frac{((r + \mathbb{E}F)^{1-\alpha/2} - (\mathbb{E}F)^{1-\alpha/2})^2}{2ck^2}\right).$$

Corollary 5.5. *Consider a well-behaved U-statistic F of order k with non-negative kernel f. Assume that*

(5.15) $$V := \mathbb{E}\int_{\mathbb{X}} F(x, \eta + \delta_x)^2 d\mu(x) < \infty.$$

Then, for all $r > 0$ we have

$$\mathbb{P}(F \leq \mathbb{E}F - r) \leq \exp\left(-\frac{r^2}{2k^2V}\right).$$

Proof. We have $\mathbb{E}V^-(F) = k^2\mathbb{E}V < \infty$ and since F is well-behaved, it follows from Corollary 5.3 (ii) that $D_x F(\xi) \geq 0$ for any $(x, \xi) \in \mathbb{X} \times \mathbf{N}$. So the result follows from Theorem 3.8 together with Remark 3.9 once we proved that for any $(x, \xi) \in \mathbb{X} \times \mathbf{N}$ one has that $D_z D_x F(\xi) \geq 0$ holds for μ-a.e. $z \in \mathbb{X}$. Now, since the measure μ is non-atomic, it is sufficient if $D_z D_x F(\xi)$ is non-negative for any $(z, x, \xi) \in \mathbb{X} \times \mathbb{X} \times \mathbf{N}$

such that $x \neq z$. Let $B \subseteq \mathbf{N}$ be a measurable set as described in Subsection 5.2. For any $\xi \in B^c$ it is immediate that $D_z D_x F(\xi) = 0$, so let $\xi \in B$ and $x, z \in \mathbb{X}$ such that $x \neq z$. If $x \in \xi$ or $z \in \xi$, one also has clearly that $D_z D_x F(\xi) = 0$, so assume that $x, y \notin \xi$. Using Corollary 5.3 (ii) we can now write

$$
\begin{aligned}
D_z D_x F(\xi) &= D_x F(\xi + \delta_z) - D_x F(\xi) \\
&= kF(x, \xi + \delta_x + \delta_z) - kF(x, \xi + \delta_x) \\
&= k \sum_{\mathbf{y} \in (\xi + \delta_z)_{\neq}^{k-1}} f(x, \mathbf{y}) - k \sum_{\mathbf{y} \in \xi_{\neq}^{k-1}} f(x, \mathbf{y}) \\
&= k(k-1) \sum_{\mathbf{y} \in \xi_{\neq}^{k-2}} f(z, x, \mathbf{y}).
\end{aligned}
$$

Note that the above formula for DDF is just a special case of the formula for the iterated difference operator of a U-statistic that was established in the proof of [31, Lemma 3.5], and that the above computation is only carried out for the reader's convenience. The result follows from the last display since $f \geq 0$. $\qquad \square$

5.4. **Computing V in formula** (5.15). We will now provide a direct proof that condition (5.15) is equivalent to the fact that F is a square-integrable U-statistic and that one can obtain a rather explicit expression for V in terms of some set of auxiliary kernels built from f.

Definition 5.6. Let f be a symmetric element of $L^1(\mu^k)$, for some $k \geq 1$. For $i = 1, ..., k$, we define the kernels f_i as follows:

$$
(5.16) \qquad f_i(y_1, ..., y_i) := \binom{k}{i} \int_{\mathbb{X}^{k-i}} f(y_1, ..., y_i, z_1, ..., z_{k-i}) d\mu^{k-i}(z_1, ..., z_{k-i}),
$$

if the integral on the right-hand side is well defined, and $f_i(y_1, ..., y_i) = 0$ otherwise. Observe that, since f is in $L^1(\mu^k)$, then the class of those $(y_1, ..., y_i)$ such that the integral on the right-hand side of (5.16) is not defined has measure μ^i equal to zero, for every $i = 1, ..., k$. Plainly, each f_i is a symmetric mapping from $\mathbb{X}^i$ into $\mathbb{R}$ and $f_i \in L^1(\mu^i)$, for every $i = 1, ..., k$, and $f_k = f$ by definition.

The upcoming result provides necessary and sufficient conditions for the square-integrability of U-statistics. The equivalence of statement (i) and (ii) as well as the formula for the variance in the proposition below are essentially content from [36, Thoerem 4.11], and we only present proofs for these parts of the result for the sake of completeness. Note that, however, our arguments are based on classical results by Surgalis [40] and are therefore somewhat different from those in [36].

Proposition 5.7 (Characterization of square-integrable U-statistics). *Consider a well-behaved U-statistic F of order $k \geq 1$, with non-negative kernel $f \in L^1(\mu^k)$. Then, the following assertions are equivalent:*

    (i) *F is square-integrable;*
    (ii) *for every $i = 1, ..., k$, $f_i \in L^2(\mu^i) \cap L^1(\mu^i)$, where the kernels f_i have been introduced in Definition 5.6;*

(iii) $V < \infty$, where V is defined in (5.15).

If either one of conditions (i), (ii) or (iii) is verified, then

$$(5.17) \qquad k^2 V = \sum_{i=1}^{k} i i! \|f_i\|_{L^2(\mu^i)}^2 \quad and \quad \mathbb{V} F = \sum_{i=1}^{k} i! \|f_i\|_{L^2(\mu^i)}^2,$$

so that, in particular, $V \leq k^{-1} \times \mathbb{V} F$.

Proof. [Step 1: (i) $\to$ (ii), (iii)] According to [31, Theorem 3.6], if F is a U-statistic as in the statement and F is square-integrable, then necessarily $f_i \in L^1(\mu^i) \cap L^2(\mu^i)$ for every $i = 1, ..., k$, and moreover F admits the following representation:

$$F = \mathbb{E} F + \sum_{i=1}^{k} I_i(f_i),$$

where I_i denotes a multiple Wiener-Itô integral of order i, with respect to the compensated Poisson measure $\hat{\eta} = \eta - \mu$ (see e.g. [27, Chapter 5] for definitions). Note that, exploiting the standard orthonormality properties of multiple integrals, one has also that

$$\mathbb{V} F = \sum_{i=1}^{k} i! \|f_i\|_{L^2(\mu^i)}^2,$$

which corresponds to the first relation in (5.17). Combining [22, Theorem 3.3] with the previous discussion, one also infers that, if F is square-integrable, then a version of the add-one cost operator DF is given by

$$(5.18) \qquad D_x F = \sum_{i=1}^{k} i I_{i-1}(f_i(x, \cdot)),$$

where $I_{i-1}(f_i(x, \cdot))$ indicates a multiple Wiener-Itô integral of order $i-1$, with respect to $\hat{\eta} = \eta - \mu$, of the kernel $f_i(x, \cdot) : \mathbb{X}^{i-1} \to \mathbb{R}$, obtained from f_i (see Definition 5.6) by setting one of the variables in its argument equal to x; observe that, as usual, the right-hand side of (5.18) is implicitly set equal to zero on the exceptional set of those $x \in \mathbb{X}$ such that $f_i(x, \cdot) \notin L^2(\mu^{i-1})$ for at least one $i \in \{1, ..., k\}$. Exploiting the standard orthonormality properties of multiple integrals, one has therefore that

$$(5.19) \qquad \mathbb{E}[(D_x F)^2] = \sum_{i=1}^{k} i^2 (i-1)! \|f_i(x, \cdot)\|_{L^2(\mu^{i-1})}^2,$$

so that the conclusion (as well as the explicit expression of $k^2 V = \mathbb{E} \int_{\mathbb{X}} (D_x F)^2 d\mu(x)$ appearing in (5.17)) follows from an application of the Fubini Theorem.

[Step 2: (ii) $\to$ (i)] Assume that, for every $i = 1, ..., k$, $f_i \in L^2(\mu^i) \cap L^1(\mu^i)$. Then, according to [40, Theorem 4.1] the multiple integral $I_i(f_i)$ is a well-defined square-integrable random variable, and moreover

$$I_i(f_i) = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \sum_{(x_1, ..., x_j) \in \eta_{\neq}^j} f_i^{(j)}(x_1, ..., x_j),$$

where $f_i^{(j)} := \binom{k}{i} \times \binom{k}{j}^{-1} \times f_j$, and each (possibly infinite) sum in the previous expression converges in $L^1(\mathbb{P})$. Now write

$$u(j) := \sum_{(x_1,...,x_j)\in\eta_{\neq}^j} f_j(x_1, ..., x_j), \quad j = 0, ..., k.$$

The previous discussion yields that

$$\sum_{i=0}^{k} I_i(f_i) = \sum_{i=0}^{k} \binom{k}{i} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} u(j)$$

$$= \sum_{j=0}^{k} u(j) \left\{ \sum_{i=j}^{k} \binom{k}{i}\binom{i}{j}(-1)^{i-j} \right\} = u(k) = F,$$

where we have used the fact that the sum $\sum_{i=j}^{k} \binom{k}{i}\binom{i}{j}(-1)^{i-j}$ equals one if $j = k$, and vanishes otherwise. It follows that F is square-integrable, since it is equal to a finite sum of square-integrable random variables.

[Step 3: (iii) → (ii)] If $V < \infty$, then there exists a measurable set $B \subset \mathbb{X}$ such that $\mu(B^c) = 0$, and $\mathbb{E}(D_x F)^2 < \infty$, for every $x \in B$. Using [31, Lemma 3.5, Theorem 3.6] together with the fact that, since F is well-behaved, $D_x F$ is the (well-behaved) U-statistic of order $k-1$ defined in Proposition 5.2, we immediately deduce that, for $x \in B$, one has that (adopting the same notation as in Step 1) $f_i(x, \cdot) \in L^2(\mu^{i-1})$, and also

$$D_x F = \sum_{i=1}^{k} i I_{i-1}(f_i(x, \cdot)).$$

The conclusion follows by using once again (5.19) and the Fubini Theorem. $\qquad\square$

Remark 5.8. According e.g. to [28, Lemma 3.1], the condition $\mathbb{E} \int_{\mathbb{X}} (D_x F)^2 d\mu(x) < \infty$ is equivalent to the fact that F belongs to the domain of the *Malliavin derivative* associated with η. This fact is consistent with the fact that square-integrable U-statistics have a finite Wiener-Itô chaotic expansion, and therefore belong automatically to the domain of the Malliavin derivative.

An application of Proposition 5.7 to the estimation of lower tails for edge-counting in random geometric graphs (involving in particular U-statistics of order $k = 2$) is presented in Section 6.3.

6. Applications to edge counting

In this section, we let η denote a Poisson point process on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, with intensity given by a Borel measure μ (in particular, $\mu(K) < \infty$ for every compact set K). We also assume again that μ has no atoms, that is, $\mu(\{x\}) = 0$ for every $x \in \mathbb{R}^d$. For a fixed $\rho > 0$, we shall consider the graph $\mathfrak{G}$ (often called the *Gilbert graph*, or the *disk Graph* with radius ρ associated with η) obtained as follows: the vertex set of $\mathfrak{G}$ is given by the points in the support of η, and two vertices x, y are linked by an edge (in symbols, $x \leftrightarrow y$) whenever $0 < \|x - y\| \leq \rho$ (in particular, $\mathfrak{G}$ has no loops).

For technical reasons clarified below, we will assume for the rest of the section that the following condition on μ is verified: denoting $B(x, \rho)$ the closed ball of radius ρ centered in x,

$$(6.20) \qquad \int_{\mathbb{R}^d} \mu(B(x, \rho)) \, d\mu(x) < \infty.$$

Relation (6.20) is verified whenever $\mu(\mathbb{R}^d) < \infty$, but such a finiteness condition is not necessary for (6.20) to hold [1]. Note that, if μ is Borel and (6.20) is in order, then the mapping $x \mapsto \mu(B(x, \rho))$ is necessarily bounded. To see this, choose $\gamma > 0$ such that the ball $B(0, \rho)$ can be written as a union of $\lfloor 1/\gamma \rfloor$ many sets with diameter less than ρ. Then the pigeonhole principle yields that for any $y \in \mathbb{R}^d$ we can choose a set $C_y \subset B(y, \rho)$ satisfying: (i) $C_y \subseteq B(x, \rho)$ for all $x \in C_y$, and (ii) $\mu(C_y) \geq \gamma \mu(B(y, \rho))$. Now,

$$\int_{\mathbb{R}^d} \mu(B(x, \rho)) d\mu(x) \geq \sup_{y \in \mathbb{R}^d} \int_{C_y} \mu(B(x, \rho)) d\mu(x) \geq \gamma \sup_{y \in \mathbb{R}^d} \int_{C_y} \mu(B(y, \rho)) d\mu(x)$$
$$= \gamma \sup_{y \in \mathbb{R}^d} \mu(C_y) \mu(B(y, \rho)) \geq \gamma^2 \sup_{y \in \mathbb{R}^d} \mu(B(y, \rho))^2.$$

Originally introduced in 1959 by Gilbert in the seminal work [13], the disk graph $\mathfrak{G}$ is the archetypical example of a *random geometric graph*. Since then, the study of such an object has been at the center of a formidable collective effort, both at a theoretical and applied level. We refer the reader to the fundamental monograph [29] for a detailed overview of the literature on Gilbert graphs up to the year 2003. Recent developments that are relevant for our work are discussed e.g. in [6, 9, 19, 20, 31, 32, 37].

In this section, we will provide new concentration estimates for the random variable

$$N = N(\eta) := \#\big\{ \{x, y\} \subseteq \eta : x \leftrightarrow y \big\},$$

corresponding to the number of edges of $\mathfrak{G}$. It is immediately seen that N is a Poisson U-statistic of order 2 with positive kernel $f(x, y) = \frac{1}{2} \mathbb{1}\{\|x - y\| \leq \rho\}$. In particular, the Slivniak-Mecke formula (2.4) together with a standard use of the Fubini Theorem yields that the assumption (6.20) is actually equivalent to integrability of N and that

$$\mathbb{E} N = \frac{1}{2} \int_{\mathbb{R}^d} \mu(B(x, \rho)) \, d\mu(x).$$

We also see that assumption (6.20) implies that $N < \infty$ almost surely, yielding in turn that N is well-behaved.

---

[1] Consider for instance the measure μ on $\mathbb{R}^2$ having density $p(x) = (\|x\| + 1)^{-2}$ together with an arbitrary radius $\rho > 0$

6.1. **Preparation: optimal rates.** Let the above notation and assumptions prevail. In the forthcoming Section 6.2, we will provide estimates for the upper tail of N having the form

$$(6.21) \qquad \mathbb{P}(N \geq \mathbb{E}N + r) \leq \exp(-I(r)), \quad r > 0,$$

where $r \mapsto I(r)$ is a positive mapping verifying

$$(6.22) \qquad \lim_{r \to \infty} I(r) = \infty.$$

The next statement contains a universal necessary condition on the asymptotic behaviour of $I(r)$.

Proposition 6.1. *Let $I(r)$ verify (6.21) and (6.22). Then,*

$$\limsup_{r \to \infty} \frac{I(r)}{r^{1/2} \log r} \leq \frac{1}{\sqrt{2}}.$$

Proof. Let $x \in \mathbb{R}^d$ such that $q := \mu(B(x, \rho/2)) > 0$. Then $\hat{N} := \eta(B(x, \rho/2))$ is Poisson distributed with expectation q. Moreover, the distance between any $y, z \in B(x, \rho/2)$ is at most ρ, thus any two vertices in $B(x, \rho/2)$ are connected by an edge. This implies that almost surely $\hat{N}(\hat{N} - 1)/2 \leq N$. Hence, for any $r \geq 0$ we have

$$\mathbb{P}(N \geq \mathbb{E}N + r) \geq \mathbb{P}(\hat{N}^2 - \hat{N} \geq 2\mathbb{E}N + 2r) \geq \mathbb{P}(\hat{N} \geq h(r)),$$

where $h(r) := \sqrt{2\mathbb{E}N + \frac{1}{4} + 2r} + \frac{1}{2}$. Now, using that for sufficiently large r,

$$\mathbb{P}(\hat{N} \geq h(r)) \geq \frac{q^{\lceil h(r) \rceil} e^{-q}}{\lceil h(r) \rceil !} \geq \frac{q^{h(r)} e^{-q}}{h(r)^{h(r)}} = \exp(-h(r) \log(h(r)/q) - q),$$

we obtain

$$\liminf_{r \to \infty} \frac{\mathbb{P}(N \geq \mathbb{E}N + r)}{\exp(-h(r) \log(h(r)/q) - q)} \geq 1.$$

The above considerations yield that there exists a constant $C \geq 0$ such that, for sufficiently large r,

$$h(r) \log(h(r)/q) + q \geq I(r) - C.$$

Dividing this inequality by $I(r)$ and letting r diverge to infinity gives

$$(6.23) \qquad \liminf_{r \to \infty} \frac{h(r) \log(h(r)/q)}{I(r)} \geq 1.$$

The conclusion is obtained by observing that, as $r \to \infty$,

$$h(r) \log(h(r)/q) \sim (r/2)^{1/2} \log r.$$

$\square$

The following statement is an elementary consequence of Proposition 6.1.

Corollary 6.2. *Let $r \mapsto I(r)$ be a positive mapping verifying* (6.21)*, and assume that there exist constants $a, b > 0$ such that, as $r \to \infty$, $I(r) \sim b\,r^a$. Then, necessarily, $a \leq \frac{1}{2}$.*

6.2. **Deviation inequalities for the upper tail.** We will now deal with bounds on the upper tail of N. We start by observing that, for every $x \in \eta$, the local version $N(x, \eta)$, as defined in (5.14), is exactly given by the quantity $\deg(x)/2$, where $\deg(x) = \#\{y \in \eta : x \leftrightarrow y\}$ is the *degree* of the vertex x. Our aim in what follows is to show that, for some constant $c > 0$, one has that almost surely

$$(6.24) \qquad \sum_{x \in \eta} \deg(x)^2 \leq cN^{3/2}.$$

Hence, Corollary 3.6 yields the following deviation inequality for the upper tail:

$$(6.25) \qquad \mathbb{P}(N \geq \mathbb{E}N + r) \leq \exp\left(-\frac{((r + \mathbb{E}N)^{1/4} - (\mathbb{E}N)^{1/4})^2}{2c}\right).$$

The constant c in the above inequality is given by

$$c = \left(\frac{8\sqrt{2}}{3} + \frac{4}{D}\right)\mathfrak{p}(d),$$

where

$$D = \frac{4\sqrt{2}}{3}\mathfrak{p} + \sqrt{\frac{32}{9}\mathfrak{p}^2 + 4\mathfrak{p} - 1}$$

and $\mathfrak{p} = \mathfrak{p}(d)$ denotes the smallest integer $\mathfrak{p}$ such that the half ball $B = \{x \in \mathbb{R}^d : \|x\| \leq \rho, x_1 > 0\}$ can be written as a union $B = B_1 \cup \ldots \cup B_{\mathfrak{p}}$ of disjoint sets such that $\mathrm{diam}(B_i) \leq \rho$ for all $i = 1, ..., \mathfrak{p}$. Note that the value of $\mathfrak{p}$ depends on the dimension d of the surrounding Euclidean space. In the plane $\mathbb{R}^2$ one has for example that $\mathfrak{p} = 3$.

Observe that the right-hand side of (6.25) has the form $\exp(-I(r))$, where $I(r) \sim r^{1/2}/2c$, as $r \to \infty$. According to Corollary 6.2, the power $1/2$ for r is optimal in this situation. We will see in Section 7.1 that, by adopting an alternative approach, the rate of decay of $I(r)$ can indeed be improved by the square root of a logarithmic factor.

Also notice that by virtue of relation (6.24) together with Theorem 3.6, almost sure finiteness of N is equivalent to integrability of N. Hence, relation (6.20) actually holds if and only if N is almost surely finite.

We start by proving a geometric lemma, focussing on deterministic point configurations. The picture below illustrates the situation described in the proof of the upcoming result.
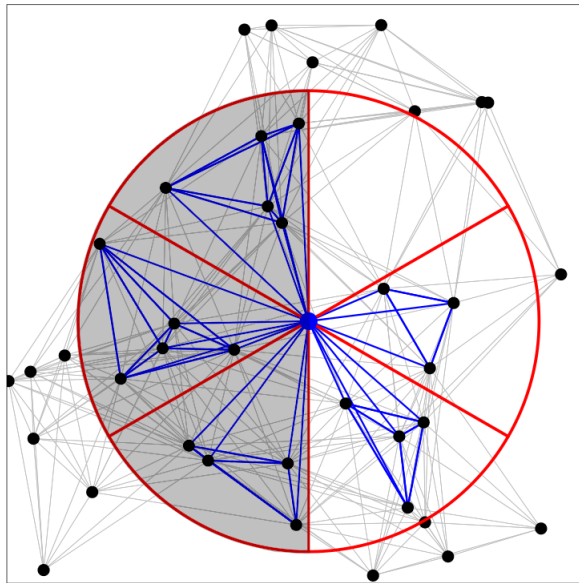
FIGURE 1. Partitioning of the half ball

Lemma 6.3. *Let $\xi \subset \mathbb{R}^d$ be a countable set. Denote the disk graph with radius $\rho > 0$ associated with ξ by $\mathfrak{G}_\xi$. For all $x \in \xi$, define the* right-degree *and the* left-degree *of x as*

$$\deg_r(x) = \#((x + B) \cap \xi) \quad and \quad \deg_l(x) = \#((x - B) \cap \xi).$$

Let T_ξ and N_ξ denote the number of triangles and edges in $\mathfrak{G}_\xi$, respectively. Then

$$2\mathfrak{p}T_\xi + \mathfrak{p}N_\xi \geq \sum_{x \in \xi} \deg_r(x)^2.$$

This inequality also holds for the left-degree instead of the right-degree.

Proof. For the rest of the proof, we write $\mathfrak{G}$, T and N, without the subscript ξ, to simplify the notation. Without loss of generality, we can assume that ξ only contains a finite number of non-isolated points in the topology of the graph; otherwise $N = \infty$ and the estimate in the statement is trivially satisfied. Let $x \in \xi$ and denote by $T(x)$ the number of those triangles $\{x, y, z\}$ in $\mathfrak{G}$ incident to x, and such that $x_1 < \min(y_1, z_1)$, where a_1 indicates the first coordinate of a given vector $a \in \mathbb{R}^d$. Moreover, for $i = 1, \ldots, \mathfrak{p}$ let n_i denote the number of elements of ξ contained in $B_i + x$. Then, since any two vertices contained in the same $B_i + x$ yield an edge and thus a triangle incident to x in the sense described above, we have

$$T(x) \geq \sum_{i=1}^{\mathfrak{p}} \binom{n_i}{2}.$$

In view of the relation $\sum_i n_i = \deg_r(x)$, the right-hand side of the previous expression can be further bounded from below, thus yielding the relation:

$$T(x) \geq \frac{1}{2}\mathfrak{p} \cdot \frac{\deg_r(x)}{\mathfrak{p}}\left(\frac{\deg_r(x)}{\mathfrak{p}} - 1\right).$$

Hence,

$$2\mathfrak{p}T(x) \geq \deg_r(x)^2 - \mathfrak{p}\deg_r(x).$$

Summing up over all $x \in \xi$ yields

$$2\mathfrak{p}\sum_{x \in \xi} T(x) + \mathfrak{p}\sum_{x \in \xi}\deg_r(x) \geq \sum_{x \in \xi}\deg_r(x)^2.$$

Observe that in the sum $\sum_{x \in \xi} T(x)$ each triangle is counted at most once, thus $\sum_{x \in \xi} T(x) \leq T$. Also, in the sum $\sum_{x \in \xi}\deg_r(x)$ each edge is counted at most once, so this sum is less than N. $\qquad\square$

We now deduce a bound on $\sum_{x \in \xi}\deg(x)^2$ for any countable point configuration ξ.

Corollary 6.4. *Let $\xi \subset \mathbb{R}^d$ be countable. Let $\mathfrak{G}_\xi$ be the disk graph (with some arbitrary radius $\rho > 0$) associated with ξ and denote the number of edges of $\mathfrak{G}_\xi$ by N_ξ. Then,*

$$\sum_{x \in \xi}\deg(x)^2 \leq \frac{8\sqrt{2}}{3}\mathfrak{p}N_\xi^{3/2} + 4\mathfrak{p}N_\xi \leq \left(\frac{8\sqrt{2}}{3} + 4\right)\mathfrak{p}N_\xi^{3/2}.$$

Proof. First we observe that without loss of generality it can be assumed that the first coordinates of all elements in ξ are distinct. Obviously, the combinatorial structure of $\mathfrak{G}_\xi$ is invariant under isometric transformations of the set ξ. Denote by $\langle \cdot, \cdot \rangle$ the Euclidean inner product on $\mathbb{R}^d$. If there exists a vector $a \in \mathbb{R}^d$ with $\|a\| = 1$ satisfying $\langle a, x \rangle \neq \langle a, y \rangle$ for all distinct $x, y \in \xi$, then the assumption in question can be achieved to hold by transforming the set ξ with respect to a linear isometry that takes a to the vector $(1, 0, \ldots, 0) \in \mathbb{R}^d$. Such a vector $a \in \mathbb{R}^d$ exists since the (countable) union of all orthogonal complements $(x-y)^\perp = \{z \in \mathbb{R}^d : \langle x-y, z \rangle = 0\}$ corresponding to vectors $x - y$ for distinct $x, y \in \xi$ has Lebesgue measure λ on $\mathbb{R}^d$ equal to zero:

$$\lambda\left(\cup_{(x,y)\in\xi_{\neq}^2}(x-y)^\perp\right) = 0.$$

So, it can be assumed that the elements of ξ have distinct first coordinates. In particular, for any $x \in \xi$ we have $\deg(x) = \deg_r(x) + \deg_l(x)$. Thus

$$\sum_{x \in \xi}\deg(x)^2 = \sum_{x \in \xi}(\deg_r(x) + \deg_l(x))^2 \leq 2\sum_{x \in \xi}(\deg_r(x)^2 + \deg_l(x)^2).$$

By Lemma 6.3, the latter expression does not exceed

$$8\mathfrak{p}T_\xi + 4\mathfrak{p}N_\xi,$$

where T_ξ stands for the number of triangles in $\mathfrak{G}_\xi$. The result follows by using an estimate taken from [34], where it was proven that

$$3T_\xi \leq \sqrt{2}N_\xi^{3/2}.$$

$\square$

The next statement is one of the main achievements in the present section.

Theorem 6.5. (i) *Let $\xi \subset \mathbb{R}^d$ be countable. Let $\mathfrak{G}_\xi$ be a disk graph with arbitrary radius ρ associated with ξ, and denote the number of edges of $\mathfrak{G}_\xi$ by N_ξ. Then,*

$$\sum_{x \in \xi} \deg(x)^2 \leq \left(\frac{8\sqrt{2}}{3} + \frac{4}{D} \right) \mathfrak{p}N_\xi^{3/2},$$

where

$$D = \frac{4\sqrt{2}}{3}\mathfrak{p} + \sqrt{\frac{32}{9}\mathfrak{p}^2 + 4\mathfrak{p} - 1}.$$

Moreover, an upper bound for $\mathfrak{p}(d)$ is given by

$$\mathfrak{p}(d) \leq \left\lceil 2\sqrt{d} \right\rceil^d.$$

(ii) *Now let η be the Poisson measure on $\mathbb{R}^d$ with non-atomic intensity μ considered in this section, and denote by $\mathfrak{G}_\eta = \mathfrak{G}$ the random disk graph (with arbitrary radius ρ) associated with η. Let $N_\eta = N$ be the number of edges of the random disk graph $\mathfrak{G}$. Then relation (6.24) holds almost surely, for $c = \left(\frac{8\sqrt{2}}{3} + \frac{4}{D} \right) \mathfrak{p}(d)$. In particular, the tail estimate (6.25) is verified.*

Proof. [Proof of (i)] By Corollary 6.4 we have

$$\sum_{x \in \xi} \deg(x)^2 \leq \left(\frac{8\sqrt{2}}{3} + \frac{4}{\sqrt{N_\xi}} \right) \mathfrak{p}N_\xi^{3/2}.$$

Observe (for instance by induction on n) that for any graph with n edges, the sum of the squared degrees is bounded from above by $n^2 + n$. Thus, we also have

$$\sum_{x \in \xi} \deg(x)^2 \leq N_\xi^2 + N_\xi = \left(\sqrt{N_\xi} + \frac{1}{\sqrt{N_\xi}} \right) N_\xi^{3/2}.$$

Now, since

$$\left(\frac{8\sqrt{2}}{3} + \frac{4}{\sqrt{N_\xi}} \right) \mathfrak{p}$$

is monotonically decreasing in N_ξ and

$$\sqrt{N_\xi} + \frac{1}{\sqrt{N_\xi}}$$

is monotonically increasing in N_ξ, the minimum of both functions is always less or equal to the value at the intersection of them. Computing this value yields the estimate for $\sum_{x \in \xi} \deg(x)^2$.

To conclude the bound for $\mathfrak{p}(d)$, note that the half ball B is contained in the cube $C = [-\rho, \rho]^d$. Moreover, the edge length of a cube with diameter ρ equals $\rho/\sqrt{d}$. Thus, we can cover an edge of the cube C with $\lceil 2\sqrt{d} \rceil$ many cubes of diameter ρ. The whole cube C can therefore be covered with $\lceil 2\sqrt{d} \rceil^d$ many cubes of diameter ρ, and this yields the bound for $\mathfrak{p}(d)$.

[Proof of (ii)] This follows directly from part (i) of the statement. $\square$

6.3. **Deviation inequalities for the lower tail.** We now focus on lower tails. In order to do that, we introduce the notation

$$K := \sup_{x \in \mathbb{R}^d} \mu(B(x, \rho)) < \infty,$$

and we define the parameter $\mathfrak{v}$ as

$$\mathfrak{v} := 2(K + 1) \, \mathbb{E}N.$$

Our main estimate is the following:

Theorem 6.6. *For every $r > 0$ one has the estimate*

$$(6.26) \qquad \qquad \mathbb{P}(N \leq \mathbb{E}N - r) \leq \exp\left(-\frac{r^2}{2\mathfrak{v}}\right).$$

Proof. We will freely use the notation and definitions introduced in Section 5.4. Since N is a well-behaved U-statistic of order 2 with kernel $f(x, y) = \frac{1}{2}\mathbb{1}\{\|x - y\| \leq \rho\}$, by virtue of Proposition 5.7 and Corollary 5.5 it is sufficient to show that, in this case, $4V \leq \mathfrak{v}$. In order to do that, observe first that $f_1(x) = \mu(B(x, \rho))$ and $f_2 = f$. As a consequence,

$$\|f_1\|^2_{L^2(\mu^1)} = \int_{\mathbb{R}^d} \mu(B(x, \rho))^2 \, d\mu(x) \leq 2K\mathbb{E}N;$$
$$\|f_2\|^2_{L^2(\mu^2)} = \tfrac{1}{2}\mathbb{E}N.$$

Plugging the above upper bounds into the formula for V contained in (5.17) yields the inequality $4V \leq \mathfrak{v}$, and therefore the desired conclusion. $\square$

6.4. **Comparison with the literature.** We shall now briefly compare the concentration inequalities for edge counting presented above with the results already existing in the literature. To the best of our knowledge, the only concentration inequalities known so far that apply to the setting of edge counting in random disk graphs over Poisson point configurations, are established in [32] and [11]. Both papers deal with the case where the intensity measure μ is finite.

Comparison with [11]. We shall use some computations from [32], where it is explained how the general results about stabilizing functionals from [11] apply to random graph statistics. In the special case of edge counting, using the notation

and assumptions of the present section, one deduces indeed from [32, Proposition 5.1] an estimate of the type

$$\mathbb{P}(|N - \mathbb{E}N| \geq r) \leq \exp(-I_0(r)), \quad \text{where } I_0(r) \sim ar^{1/3},$$

for some positive constant a, as $r \to \infty$. As far as the asymptotic behaviour of $r \mapsto I_0(r)$ is concerned, this result is worse than the estimates that one can obtain from (6.25) (where the argument of the exponential bound on the upper tail is asymptotic to $-r^{1/2}/2c$, and therefore optimal in the sense of Corollary 6.2), and than those given by (6.26) (where we have proved a Gaussian upper bound on the lower tail).

Comparison with [32]. Writing m for a median of the law of N, in [32, Theorem 5.2] one can find an estimate of the type

$$\mathbb{P}(|N - m| \geq r) \leq \exp(-I_1(r)), \quad \text{where } I_1(r) \sim br^{1/2},$$

as $r \to \infty$, for some positive constant b. Note that, as $r \to \infty$, the upper tail bound determined by I_1 has the same order as our estimate (6.25), whereas the lower tail estimate is worse than our Gaussian upper bound (6.26). We also stress that [32, Theorem 5.2] gives a concentration inequality around the median, and not the expectation. This might be a drawback for applications since the median of the edge count is harder to deal with than the expectation – which can be easily expressed using the Slivniak-Mecke formula. Finally, a major advantage of our results over those presented in [32, Theorem 5.2] is that the latter only applies to disk graphs built on finite intensity measure Poisson processes, whereas our tail estimates merely require that the number of edges is almost surely finite.

6.5. **Consistency with CLT.** In the following, we compare the deviation inequality for the upper tail with a CLT that was proven in [31]. Let η_1 be a Poisson point process in $\mathbb{R}^d$ with intensity measure μ_1 and fix some radius $\rho > 0$. For any $n \in \mathbb{N}$, let η_n be a Poisson point process with intensity measure $\mu_n = n\mu_1$ and denote the number of edges in the corresponding random geometric graph by N_n. Assume that $\mathbb{E}N_1 < \infty$ (and hence $\mathbb{E}N_n < \infty$ for all n). Then by [31, Theorem 5.2], the sequence of random variables N_n satisfies a central limit theorem, i.e. $(N_n - \mathbb{E}N_n)/\sqrt{\mathbb{V}N_n}$ converges to a standard Gaussian distribution, where $\mathbb{V}N_n$ stands for the variance of N_n. Therefore, as $n \to \infty$, the sequence of probabilities $\mathbb{P}(N_n \geq \mathbb{E}N_n + \sqrt{\mathbb{V}N_n}r)$, $n \geq 1$, converges to the quantity

$$\frac{1}{\sqrt{2\pi}} \int_r^\infty e^{-y^2/2} dy \leq \frac{e^{-r^2/2}}{r\sqrt{2\pi}}.$$

According to the next result, the asymptotic behavior of the upper tail deviation inequality is consistent with this Gaussian tail.

Theorem 6.7. *Let $c > 0$ be a constant satisfying (6.24). Then there exists a constant $C > 0$ and a sequence $(x_n)_{n \in \mathbb{N}}$ with $x_n \to \infty$ as $n \to \infty$ such that for any*

$n \in \mathbb{N}$,

$$\exp\left(-\frac{(((\mathbb{V}N_n)^{1/2}r + \mathbb{E}N_n)^{1/4} - (\mathbb{E}N_n)^{1/4})^2}{2c}\right) \leq \exp(-Cr^2) \quad \text{for all} \ \ r \in [0, x_n].$$

Proof. It was pointed out in [31] that there are constants $\alpha, \beta > 0$ such that

$$\mathbb{V}N_n \sim \alpha n^3,$$
$$\mathbb{E}N_n \sim \beta n^2.$$

Let $A = \sqrt{2Cc}$. Then the desired inequality is equivalent to

$$(\mathbb{V}N_n)^{1/2}r \geq \left(Ar + (\mathbb{E}N_n)^{1/4}\right)^4 - \mathbb{E}N_n$$
$$= A^4r^4 + 4A^3r^3(\mathbb{E}N_n)^{1/4} + 6A^2r^2(\mathbb{E}N_n)^{2/4} + 4Ar(\mathbb{E}N_n)^{3/4}.$$

This holds if and only if

$$(\mathbb{V}N_n)^{1/2}(\mathbb{E}N_n)^{-3/4} - 4A \geq A^4r^3(\mathbb{E}N_n)^{-3/4} + 4A^3r^2(\mathbb{E}N_n)^{-2/4} + 6A^2r(\mathbb{E}N_n)^{-1/4}.$$

Now, choose $C > 0$ such that $4A < \alpha^{1/2}\beta^{-3/4}$. Consider the equality corresponding to the inequality in the above display. For any $n \in \mathbb{N}$ let x_n be the (unique) positive solution of this equality in case such a solution exists, and let $x_n = 0$ otherwise. Then, since the right-hand side of the last display is increasing in r, the desired inequality holds for all $r \in [0, x_n]$. Moreover, the left-hand side converges to $\alpha^{1/2}\beta^{-3/4} - 4A > 0$ while $(\mathbb{E}N_n)^{-3/4}, (\mathbb{E}N_n)^{-2/4}, (\mathbb{E}N_n)^{-1/4} \to 0$, as $n \to \infty$. From this it follows that $x_n \to \infty$ as $n \to \infty$. $\qquad\square$

7. Another look at U-Statistics of order two

In this section, we develop a different approach for obtaining deviation inequalities for the upper tail of U-statistics of order 2, that is partially inspired by the results from [16, 33]. Throughout this section, we let the assumptions of Section 5 prevail; in particular, the intensity μ of η is a non-atomic positive measure on $(\mathbb{X}, \mathcal{X})$. We begin by generalizing [33, Theorem 3] to Poisson processes with possibly non-finite intensity measure:

Theorem 7.1. *Consider a countable family $\{f_j\}_{j \in J}$ of functions $\mathbb{X} \to [0, 1]$ and let*

$$G = \sup_{j \in J} \sum_{x \in \eta} f_j(x).$$

Assume that $\mathbb{E}G < \infty$. Then for any $\lambda > 0$ we have

$$\log \mathbb{E}[\exp(\lambda(G - \mathbb{E}G))] \leq \phi(\lambda)\mathbb{E}G,$$

where $\phi(\lambda) = e^\lambda - \lambda - 1$.

Proof. First note that by monotone convergence, we can assume without loss of generality that $|J| < \infty$. For each $n \in \mathbb{N}$ let $G_n = \min(G, n)$. Then $\mathbb{E}(e^{\lambda G_n}) < \infty$, hence Proposition 3.1 with $I = \emptyset$ gives

$$(7.27) \qquad \text{Ent}(e^{\lambda G_n}) \leq \mathbb{E}\left(e^{\lambda G_n}\sum_{x \in \eta}\phi(-\lambda D_x G_n(\eta - \delta_x))\right).$$

Consider some realization of η. Since we assumed $|J| < \infty$, it follows that for some $j^* \in J$ we have

$$G(\eta) = \sum_{x \in \eta} f_{j^*}(x).$$

Now, for any $x \in \eta$ we have

$$0 \leq D_x G(\eta - \delta_x) \leq f_{j^*}(x) \leq 1.$$

Moreover, if $G(\eta - \delta_x) \geq n$, then $D_x G_n(\eta - \delta_x) = 0$ and if $G(\eta - \delta_x) < n$, then

$$D_x G_n(\eta - \delta_x) = D_x G(\eta - \delta_x) - \max(0, G(\eta) - n).$$

From this we obtain

$$\sum_{x \in \eta} D_x G_n(\eta - \delta_x) \leq \left(\sum_{x \in \eta} f_{j^*}(x)\right) - \max(0, G(\eta) - n)$$
$$= G(\eta) - \max(0, G(\eta) - n) = G_n(\eta).$$

Since $\phi(-\lambda z) \leq \phi(-\lambda)z$ for $\lambda > 0$ and $0 \leq z \leq 1$, it follows from the above considerations that

$$\text{Ent}(e^{\lambda G_n}) \leq \phi(-\lambda)\mathbb{E}\left(e^{\lambda G_n}\sum_{x \in \eta} D_x G_n(\eta - \delta_x)\right) \leq \phi(-\lambda)\mathbb{E}(e^{\lambda G_n}G_n).$$

Continuing in the same way as in the proof of [25, Theorem 10] gives

$$(7.28) \qquad \log \mathbb{E}[\exp(\lambda(G_n - \mathbb{E}G_n))] \leq \phi(\lambda)\mathbb{E}G_n \leq \phi(\lambda)\mathbb{E}G.$$

Now, since $\lambda > 0$ and $\mathbb{E}G < \infty$, by monotone convergence we have

$$\lim_{n \to \infty} \mathbb{E}[\exp(\lambda(G_n - \mathbb{E}G_n))] = \mathbb{E}[\exp(\lambda(G - \mathbb{E}G))].$$

Invoking (7.28) yields the result. $\qquad \square$

We continue with an analytic lemma that is used in the proof of the upcoming theorem, but which might be of independent interest in similar situations. The proof is inspired by the proof of [23, Corollary 2.12].

Lemma 7.2. *For any $z > 0$,*

$$\sup_{\lambda > 0}\left[\lambda z - e^{\lambda^2} + 1\right] \geq \frac{\sqrt{\log(z+1)}z^{3/2}}{4\sqrt{z}+8}.$$

Proof. We begin with a preparing observation. Let $a > 1$ and consider the map $\lambda \mapsto 1 - e^{\lambda^2} + a\lambda^2$. Then this map takes 0 to 0 and it has a unique local extremum (a maximum) at $\sqrt{\log(a)}$ on the positive reals. Hence, whenever $1 - e^{x^2} + ax^2 \geq 0$ for some $x > 0$, we have $1 - e^{\lambda^2} \geq -a\lambda^2$ for any $\lambda \in (0, x]$.

Now assume that $0 < z \leq 4$. Then we take $\lambda = z/4$. Since $\lambda \leq 1$ and $1 - e + 2 \geq 0$, the above observation implies

$$1 - e^{\lambda^2} \geq -2\lambda^2.$$

Hence, we have

$$\lambda z - e^{\lambda^2} + 1 \geq \lambda z - 2\lambda^2 = \frac{z^2}{8} \geq \frac{\sqrt{\log(z+1)}z^{3/2}}{4\sqrt{z}+8}.$$

For $4 < z \leq 12$ we take $\lambda = z/8$. Then $\lambda \leq 3/2$ and since $1 - e^{9/4} + 9 \geq 0$, the initial observation gives

$$1 - e^{\lambda^2} \geq -4\lambda^2.$$

Thus,

$$\lambda z - e^{\lambda^2} + 1 \geq \lambda z - 4\lambda^2 = \frac{z^2}{16} \geq \frac{\sqrt{\log(z+1)}z^{3/2}}{4\sqrt{z}+8}.$$

It remains to prove that the result holds for $z > 12$. Here we can take $\lambda = \sqrt{\log(z)}$ and see that the supremum is lower bounded by

$$z(\sqrt{\log(z)} - 1) + 1.$$

Now, since $\sqrt{\log(z)} - \sqrt{\log(z+1)}$ is increasing for $z \geq 12$, we have

$$\sqrt{\log(z)} \geq \sqrt{\log(z+1)} + A,$$

where $A = \sqrt{\log(12)} - \sqrt{\log(13)} < 0$. Hence,

$$z(\sqrt{\log(z)} - 1) + 1 \geq z(\sqrt{\log(z+1)} + A - 1) + 1.$$

We claim that

$$z(\sqrt{\log(z+1)} + A - 1) + 1 \geq \frac{1}{4}z(\sqrt{\log(z+1)}) \quad \text{for} \quad z > 12.$$

This will imply the result since

$$\frac{1}{4}z(\sqrt{\log(z+1)}) \geq \frac{1}{4}\frac{z(\sqrt{\log(z+1)})\sqrt{z}}{\sqrt{z}+2} = \frac{\sqrt{\log(z+1)}z^{3/2}}{4\sqrt{z}+8}.$$

To prove the claim, define

$$h(z) = z(\sqrt{\log(z+1)} + A - 1) + 1 - \frac{1}{4}z(\sqrt{\log(z+1)})$$

$$= \frac{3}{4}z\sqrt{\log(z+1)} + z(A - 1) + 1.$$

The derivative of h satisfies

$$h'(z) = \frac{3}{4}\left(\frac{z + 2\log(z+1)(z+1)}{2\sqrt{\log(z+1)}(z+1)}\right) + A - 1 \geq \frac{3}{4}\sqrt{\log(z+1)} + A - 1.$$

Since the right-hand side is increasing in z and positive for $z = 12$, we have that $h'(z) \geq 0$ for any $z > 12$. So h is increasing for $z > 12$ and since also $h(12) \geq 0$, it follows that $h(z) \geq 0$ for all $z > 12$. This proves the claim and concludes the proof of the result. $\square$

Theorem 7.3. *Let F be a U-statistic of order 2 with kernel $f \geq 0$ such that $\mathbb{E}F < \infty$. Assume that there is a countable family $\{g_j\}_{j \in J}$ of functions $\mathbb{X} \to [0,1]$ and a constant $c > 0$ such that*

$$G = \sup_{j \in J} \sum_{x \in \eta} g_j(x)$$

satisfies almost surely

$$\sup_{y \in \eta} \sum_{x \in \eta \setminus y} f(y, x) \leq cG$$

and $\mathbb{E}G < \infty$. Then for any $r > 0$ we have

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\mathbb{E}(G)\,\chi\left(\frac{\sqrt{\mathbb{E}F + r} - \sqrt{\mathbb{E}F}}{\sqrt{4c}\,\mathbb{E}G}\right)\right),$$

where

$$\chi(z) = \frac{\sqrt{\log(z+1)}z^{3/2}}{4\sqrt{z} + 8}.$$

Proof. The assumptions imply that almost surely

$$V^+ = 4\sum_{y \in \eta}\left(\sum_{x \in \eta \setminus y} f(y, x)\right)^2 \leq 4cGF.$$

By Corollary 3.5 and Theorem 7.1 this gives for any $\lambda > 0$,

$$\log \mathbb{E}[\exp(\lambda(\sqrt{F} - \mathbb{E}\sqrt{F}))] \leq \inf_{\theta \in (0, 2/\lambda)} \frac{\lambda\theta}{2 - \lambda\theta} \log \mathbb{E}\left[\exp\left(\frac{4c\lambda}{\theta}G\right)\right]$$

$$\leq \inf_{\theta \in (0, 2/\lambda)} \frac{\lambda\theta}{2 - \lambda\theta}\mathbb{E}G(\exp(4c\lambda/\theta) - 1)$$

$$\leq \mathbb{E}G(\exp(4c\lambda^2) - 1).$$

Let $r > 0$. Then, using the above computation and Markov's inequality, we obtain for any $\lambda > 0$,

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \mathbb{P}(e^{\lambda(\sqrt{F} - \mathbb{E}\sqrt{F})} \geq e^{\lambda(\sqrt{\mathbb{E}F + r} - \sqrt{\mathbb{E}F})})$$

$$\leq \exp\left(\mathbb{E}G(\exp(4c\lambda^2) - 1) - \lambda(\sqrt{\mathbb{E}F + r} - \sqrt{\mathbb{E}F})\right).$$

Hence, writing $z = (\sqrt{\mathbb{E}F + r} - \sqrt{\mathbb{E}F})/(\sqrt{4c}\mathbb{E}G)$ and substituting λ by $\lambda/\sqrt{4c}$, we obtain

$$\mathbb{P}(F \geq \mathbb{E}F + r) \leq \exp\left(-\mathbb{E}G \sup_{\lambda > 0} \left[\lambda z - \exp(\lambda^2) + 1\right]\right).$$

The result now follows from Lemma 7.2. $\qquad\square$

7.1. **Length power functionals.** As an application, we now focus on length power functionals in random geometric graphs. These estimates contain as special case the edge counting statistics that we studied in Section 6. We will see in particular that one can take advantage of the upper tail estimate stated in Theorem 7.3 in order to provide an alternate bound to that appearing in (6.25), which actually displays a *strictly faster* rate of decay in r.

As we did in Section 6, we consider a Poisson measure on $\mathbb{R}^d$ with σ-finite and non-atomic Borel intensity measure μ. We let $\rho > 0$ be some radius and consider again the disk graph $\mathfrak{G}(\eta)$ associated with η. For any $\alpha \in [0,1]$ the *length power functional* $L^{(\alpha)}$ is the U-statistic of order 2 with kernel

$$f_\alpha(x,y) = \tfrac{1}{2}\mathbb{1}\{\|x - y\| \leq \rho\}\|x - y\|^\alpha.$$

Note that $L^{(0)} = N$, the number of edges in $\mathfrak{G}(\eta)$, and $L^{(1)}$ is just the (total edge) length of the graph. One easily sees that

$$\sup_{y \in \eta} \sum_{x \in \eta \setminus y} f_\alpha(x,y) \leq 2^{d-1}\rho^\alpha \sup_{j \in \mathbb{N}^d} \sum_{x \in \eta} \mathbb{1}\{x \in [0, 2\rho]^d + 2\rho j\}.$$

It is also straightforward to check that, if $\mathbb{E}N < \infty$, then the expectation of the right-hand side of the above inequality is finite. We stress also that, if $\mathbb{E}N < \infty$, then $L^{(\alpha)}$ is trivially well-behaved for every $\alpha \in [0,1]$; see Section 5.2. The following consequence of Theorem 7.3 therefore holds.

Corollary 7.4. *Let the above notation prevail, fix $\alpha \in [0,1]$, assume that $\mathbb{E}N < \infty$, and let*

$$G = \sup_{j \in \mathbb{N}^d} \eta([0, 2\rho]^d + 2\rho j).$$

Then, $\mathbb{E}G < \infty$ and, for all $r > 0$,

$$(7.29) \qquad \mathbb{P}(L^{(\alpha)} \geq \mathbb{E}L^{(\alpha)} + r) \leq \exp\left(-\mathbb{E}(G)\,\chi\left(\frac{\sqrt{\mathbb{E}L^{(\alpha)} + r} - \sqrt{\mathbb{E}L^{(\alpha)}}}{\sqrt{2^{d+1}\rho^\alpha}\,\mathbb{E}G}\right)\right),$$

where $\chi(z)$ is as in Theorem 7.3.

Remark 7.5. The right-hand side of (7.29) has the form $\exp(-I(r))$, where $I(r) \sim b\sqrt{r \log r}$, for some $b > 0$, as $r \to \infty$. Such a rate of decay is better than the one we can deduce from [32, Proposition 5.1] (which is indeed a translation of the results from [11]), that applies to the case where μ is a multiple of the restriction of the Lebesgue measure to a convex body, and implies an upper bounds of the form $\exp(-I_0(r))$, with $I_0(r) \sim b_0 r^{1/3}$. Our result provides also a rate of decay that is faster than the one appearing in [32, Theorem 5.5], where the bound has the form

$\exp(-I_1(r))$, with $I_1(r) \sim b_1 r^{1/2}$. It is remarkable that, in the case $\alpha = 0$ and as far as the rate of decay (as $r \to \infty$) is concerned, the estimate (7.29) is also strictly better than (6.25), and that this comes at the cost of somewhat more complicated constants. Finally, we observe that the asymptotic relation $I(r) \sim b\sqrt{r \log r}$ is consistent with Proposition 6.1.

7.2. **Length in more general graph models.** In the following, we consider a slightly more general model of random geometric graphs. For this, let η be a Poisson point process on $\mathbb{R}^d$ with σ-finite and non-atomic Borel intensity measure μ, and let $\rho : \mathbb{R}^d \to \mathbb{R}_+$ be given by

$$(7.30) \qquad \rho(x) = \left(\frac{1}{\|x\| + 1} \right)^{\gamma},$$

for some $\gamma > 0$. We define $\mathfrak{H}(\eta)$ as the graph with vertex set η and an edge between vertices $x, y \in \eta$ whenever $0 < \|x - y\| \leq \rho(x) + \rho(y)$. Note that the graph $\mathfrak{H} = \mathfrak{H}(\eta)$ is obtained by implementing the following two-step procedure: (a) for every $x \in \eta$, draw the closed ball $B(x, \rho(x))$, centered at x and with radius $\rho(x)$, and (b) connect two distinct points $x, y \in \eta$ with an edge, if and only if

$$B(y, \rho(y)) \cap B(x, \rho(x)) \neq \emptyset.$$

In other words, $\mathfrak{H}$ is the *intersection graph* of the balls centered at the points of η with (decaying) radii given by $\rho(x), x \in \eta$. We will see that this model allows situations where $\mathfrak{H}$ has almost surely infinitely many edges but still a finite length. Interestingly, even if there are infinitely many edges, the length can have an exponentially decaying upper tail.

Let L be the length of $\mathfrak{H}$. Then L is a U-statistic of order 2 with kernel

$$f_L(x, y) = \tfrac{1}{2}\|x - y\|\mathbb{1}\{\|x - y\| \leq \rho(x) + \rho(y)\}.$$

Moreover, if μ guarantees that almost surely $L < \infty$, then L is even well-behaved. As a consequence of Theorem 7.3 we obtain the following result.

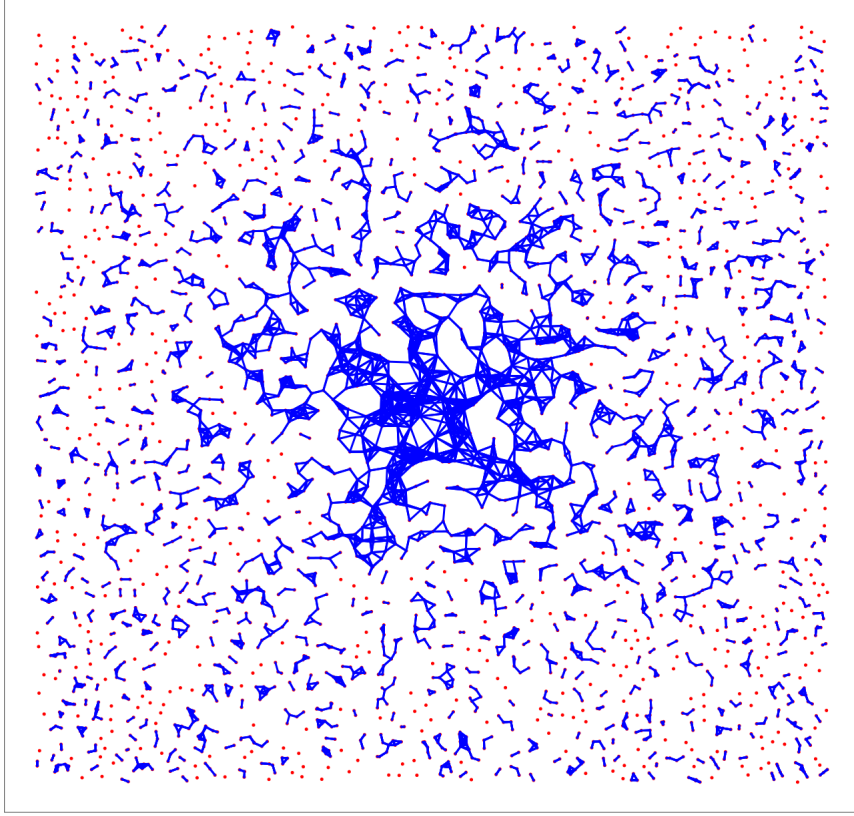**Corollary 7.6.** *Let $c = 3^{\gamma} + 1$ and define*

$$(7.31) \qquad G = \sup_{x \in \mathbb{Q}^d} \rho(x)\eta(B(x, c\rho(x))).$$

Assume that $\mathbb{E}G < \infty$ and $\mathbb{E}L < \infty$. Then for any $r \geq 0$,

$$\mathbb{P}(L \geq \mathbb{E}L + r) \leq \exp\left(-\mathbb{E}(G)\, \chi\left(\frac{\sqrt{\mathbb{E}L + r} - \sqrt{\mathbb{E}L}}{\sqrt{2c}\, \mathbb{E}G} \right) \right),$$

where χ is defined as in Theorem 7.3.

Before we carry out the proof of the above result, we present an illustration of how the considered graph might look like in the plane.

FIGURE 2. A realisation of the intersection graph $\mathfrak{H}(\eta)$

Proof. Let $x, y \in \mathbb{R}^d$ be such that $\|x - y\| \leq \rho(x) + \rho(y)$. Then we have $|\|x\| - \|y\|| \leq 2$ since ρ is upper bounded by 1. Hence,

$$(7.32) \qquad \frac{\rho(y)}{\rho(x)} = \left(\frac{\|x\| + 1}{\|y\| + 1} \right)^\gamma \leq \left(\frac{\|y\| + 3}{\|y\| + 1} \right)^\gamma \leq 3^\gamma.$$

Therefore, $\|x - y\| \leq (3^\gamma + 1)\rho(x)$. It follows that the local version of L satisfies

$$L(x, \eta) = \tfrac{1}{2} \sum_{y \in \eta \setminus x} \|x - y\| \mathbb{1}\{\|x - y\| \leq \rho(x) + \rho(y)\}$$

$$(7.33) \qquad \leq \tfrac{3^\gamma + 1}{2} \sum_{y \in \eta \setminus x} \rho(x) \mathbb{1}\{\|x - y\| \leq (3^\gamma + 1)\rho(x)\}.$$

For $x, y \in \mathbb{R}^d$ we define $g_x(y) = \rho(x)\mathbb{1}\{\|x - y\| \leq (3^\gamma + 1)\rho(x)\}$. Then the above reasoning gives that, almost surely,

$$\sup_{x \in \eta} \sum_{y \in \eta \setminus x} f_L(x, y) \leq \tfrac{3^\gamma + 1}{2} \sup_{x \in \mathbb{Q}^d} \sum_{y \in \eta} g_x(y).$$

Let

$$G = \sup_{x \in \mathbb{Q}^d} \sum_{y \in \eta} g_x(y).$$

We see that Theorem 7.3 applies to L whenever $\mathbb{E}L, \mathbb{E}G < \infty$ and this concludes the proof. $\qquad\square$

Next we will prove a sufficient condition for the finiteness of the expectations appearing in Corollary 7.6 for the case where the Poisson process η is homogeneous, that is, when the intensity measure of η has the form $t \times \lambda$, where λ is the Lebesgue measure.

Proposition 7.7. *Assume that η is a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity $t\lambda$, $t > 0$. Let L be the length of $\mathfrak{H}$ and define the random variable G as in (7.31). Then G and L are integrable, provided that*

$$(7.34) \qquad \int_{\mathbb{R}^d} \rho(x)^{d+1} dx < \infty.$$

Proof. First observe, quite similarly as it was done in (7.32), that for any $x \in \mathbb{Q}^d$ and $\hat{x} \in B(x, (3^\gamma + 1)\rho(x))$,

$$\rho(x) \le (3^\gamma + 2)^\gamma \rho(\hat{x}).$$

Hence, writing $c = 3^\gamma + 1$ and $c' = (3^\gamma + 2)^\gamma$, we have

$$\sum_{y \in \eta} g_x(y) = \rho(x) \sum_{y \in \eta} \mathbb{1}\{\|x - y\| \le c\rho(x)\}$$

$$\le c'\rho(\hat{x}) \sum_{y \in \eta} \mathbb{1}\{\|\hat{x} - y\| \le 2cc'\rho(\hat{x})\}.$$

It follows that $\mathbb{E}G < \infty$ if the expectation of the following is finite:

$$\sup_{x \in \eta} \sum_{y \in \eta \setminus x} \rho(x) \mathbb{1}\{\|x - y\| \le 2cc'\rho(x)\}$$

$$(7.35) \qquad \le \sum_{(x,y) \in \eta^2_{\neq}} \rho(x) \mathbb{1}\{\|x - y\| \le 2cc'\rho(x)\}.$$

Using the Slivniak-Mecke formula (2.4), we obtain that the expectation of the latter expression equals

$$t^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(x) \mathbb{1}\{\|x - y\| \le 2cc'\rho(x)\} \ dx \ dy = t^2 (2cc')^d \kappa_d \int_{\mathbb{R}^d} \rho(x)^{d+1} dx,$$

where κ_d denotes the Lebesgue measure of the unit ball in $\mathbb{R}^d$. So we have $\mathbb{E}G < \infty$ provided that (7.34) holds. This condition also guarantees $\mathbb{E}L < \infty$ (to see this, just perform a computation similar to the one above where the estimate (7.33) is used instead of (7.35)). $\qquad\square$

We initially claimed that $\mathfrak{H}$ can have a.s. infinitely many edges but still $\mathbb{E}L < \infty$. The following result, when combined with Proposition 7.7, substantiates this statement.

Proposition 7.8. *Assume that η is a homogeneous Poisson point process in $\mathbb{R}^d$ with intensity $t > 0$. Denote by N the number of edges in $\mathfrak{H}$. Then almost surely $N = \infty$, provided that*

$$(7.36) \qquad \int_{\mathbb{R}^d} \rho(x)^d dx = \infty.$$

Remark 7.9. The phenomenon described above is remarkable since it allows for situations where the U-statistic

$$L = \sum_{(x,y) \in \eta_{\neq}^2} f_L(x,y)$$

is indeed (as opposed to the edge counting statistics) almost surely an infinite series, i.e. we have almost surely $f_L(x,y) > 0$ for infinitely many $(x,y) \in \eta_{\neq}^2$. Intuitively, one might expect that in this situation strong concentration properties for L are more difficult to establish. However, as we have seen above, our method works without problems and yields exponential tail bounds for the graph length regardless of whether finitely or infinitely many edges are present.

Proof of Proposition 7.8. Note that

$$N = \tfrac{1}{2} \sum_{(x,y) \in \eta_{\neq}^2} \mathbb{1}\{\|x - y\| \leq \rho(x) + \rho(y)\}.$$

For all $x \in \mathbb{N}^d$ define the cube

$$Q_x = [0, \tfrac{1}{\sqrt{d}}\rho(x + \mathbf{1})]^d \subset \mathbb{R}^d,$$

where $\mathbf{1} = (1, \dots, 1) \in \mathbb{N}^d$. Observe that for any $x \in \mathbb{N}^d$, we can place

$$\left\lfloor \frac{\sqrt{d}}{\rho(x + \mathbf{1})} \right\rfloor^d =: r(x)$$

many disjoint translated copies of Q_x into the cube $x + [0, 1]^d$. Denote these copies by $Q_x^1, \dots Q_x^{r(x)}$. Observe also that for the diameter of each Q_x^i one has $\mathrm{diam}(Q_x^i) = \rho(x+\mathbf{1})$. Hence, any two distinct vertices $x, y \in \eta$ within the same cube are connected by an edge. Therefore,

$$N \geq \sum_{x \in \mathbb{N}^d} \sum_{i=1}^{r(x)} \binom{\eta(Q_x^i)}{2}.$$

Now, the $\eta(Q_x^i)$ are independent Poisson random variables. Thus, by the second Borel-Cantelli lemma, the right-hand side in the above display is almost surely non-finite if

$$\sum_{x \in \mathbb{N}^d} \sum_{i=1}^{r(x)} \mathbb{P}(\eta(Q_x^i) \geq 2) = \infty.$$

The expectation of $\eta(Q_x^i)$ is given by

$$\lambda_x = \frac{t}{d^{d/2}} \rho(x + \mathbf{1})^d.$$

Using this we obtain

$$\sum_{x \in \mathbb{N}^d} \sum_{i=1}^{r(x)} \mathbb{P}(\eta(Q_x^i) \geq 2) = \sum_{x \in \mathbb{N}^d} r(x)(1 - e^{-\lambda_x} - \lambda_x e^{-\lambda_x}).$$

Since $1 - e^{-z} - ze^{-z} \geq (1/4)z^2$ for $z \in [0, 1]$ and since $\lambda_x \to 0$ as $\|x\| \to \infty$, the above series is non-finite if

$$\sum_{x \in \mathbb{N}^d} r(x)\lambda_x^2 = \infty \quad \text{or equivalently} \quad \sum_{x \in \mathbb{N}^d} \rho(x + \mathbf{1})^d = \infty.$$

It is easy to see that the above is implied by condition (7.36). $\qquad\square$

8. Concentration for the convex distance in Poisson-based models

The convex distance for product spaces that was introduced by M. Talagrand in [41] has proved to be a very useful tool in the context of concentration inequalities – see e.g. [10, Chapter 11], [39, Chapter 6], [42, Chapter 2] and the references therein. In the recent paper [30] by M. Reitzner, this notion has been adapted for models based on Poisson point processes with finite intensity measure. For both the product space and the Poisson space version, the method of using the convex distance to establish concentration properties is based on an isoperimetric inequality. First applications of this method for Poisson-based models are worked out in [32, 21] where concentration inequalities for Poisson U-statistics are presented.

The proof of the convex distance isoperimetric inequality in [30] uses an approximation of the Poisson process by binomial processes. The goal in this section is to give an alternative proof for this inequality. Apart from slightly worse constants, we entirely recover Reitzner's result [30, Theorem 1.1] with the tools developed in the present work. In particular, we only use methods from Poisson process theory, thus answering the question proposed in [30] of whether such a direct proof is possible. Moreover, the assumptions on the space $\mathbb{X}$ for our results are less restrictive than in [30] where only locally compact second countable Hausdorff spaces are considered. As before, we consider a measurable space $(\mathbb{X}, \mathcal{X})$ and we will assume throughout the present section that $\{x\} \in \mathcal{X}$ for all $x \in \mathbb{X}$.

The upcoming presentation is based on [5, Section 2] and [4, pp. 1594-1595] where the convex distance for product spaces is recovered using the entropy method. The results presented in the following are derived by adapting the reasoning from the latter references to the setting of Poisson processes considered in this section.

8.1. **Convex distance for Poisson processes.** To introduce the convex distance for Poisson point processes, let $\mathbf{N}_{\text{fin}} \subset \mathbf{N}$ denote the space of finite integer-valued measures on $\mathbb{X}$ which is equipped with the σ-algebra $\mathcal{N}_{\text{fin}}$ obtained by restricting

$\mathcal{N}$ to $\mathbf{N}_{\mathrm{fin}}$. We will write $\xi(x) = \xi(\{x\})$ whenever $\xi \in \mathbf{N}_{\mathrm{fin}}$ and $x \in \mathbb{X}$ in order to simplify notations. For any two measures $\xi, \nu \in \mathbf{N}_{\mathrm{fin}}$, we define the measure $\xi \setminus \nu$ by

$$\xi \setminus \nu = \sum_{x \in \xi} (\xi(x) - \nu(x))_+ \delta_x,$$

where $x \in \xi$ indicates that x belongs to the support of ξ. The *convex distance* $d_T(\xi, A)$ is now defined for any measurable set $A \in \mathcal{N}_{\mathrm{fin}}$ and $\xi \in \mathbf{N}_{\mathrm{fin}}$ by

$$d_T(\xi, A) = \sup_{\|u\|_\xi \leq 1} \inf_{\nu \in A} \int_{\mathbb{X}} u \, d(\xi \setminus \nu),$$

where the supremum ranges over all measurable maps $u : \mathbb{X} \to \mathbb{R}$ such that $\|u\|_\xi \leq 1$ and $\|\cdot\|_\xi$ denotes the 2-norm with respect to the measure ξ. It is immediate from the above definition that

$$(8.37) \qquad d_T(\xi, A) = \sup_{\|u\|_\xi \leq 1} \inf_{\nu \in A} \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+.$$

The following result gives an alternative characterization for the convex distance which will be crucial for our proof of the isoperimetric inequality later on.

Proposition 8.1. *Let $A \in \mathcal{N}_{\mathrm{fin}}$ and denote by $\mathcal{M}(A)$ the set of probability measures on A. Then, for any $\xi \in \mathbf{N}_{\mathrm{fin}}$ we have*

$$d_T(\xi, A) = \max_{\|u\|_\xi \leq 1} \min_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+]$$

$$= \min_{\zeta \in \mathcal{M}(A)} \max_{\|u\|_\xi \leq 1} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+],$$

where, here and for the rest of the section, we use the shorthand notation

$$\mathbb{E}_{\zeta(\nu)}[h(\nu)] = \int_A h(\nu) d\zeta(\nu),$$

for every positive measurable mapping $h : A \to \mathbb{R}_+$.

Proof. Here we adapt arguments from the proof of [4, Proposition 13]. We begin by proving that

$$(8.38) \qquad d_T(\xi, A) = \sup_{\|u\|_\xi \leq 1} \inf_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+].$$

For any $\nu \in A$ consider the probability measure $\zeta_\nu \in \mathcal{M}(A)$ that is concentrated on ν. Then

$$\sum_{x \in \xi} u(x) \mathbb{E}_{\zeta_\nu(\nu')}[(\xi(x) - \nu'(x))_+] = \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+.$$

Hence, for any u with $\|u\|_\xi \leq 1$, we have

$$\inf_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+] \leq \inf_{\nu \in A} \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+.$$

On the other hand, for all $\zeta \in \mathcal{M}(A)$ we have

$$\inf_{\nu \in A} \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+ \leq \mathbb{E}_{\zeta(\nu)} \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+$$
$$= \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+].$$

Thus,

$$\inf_{\nu \in A} \sum_{x \in \xi} u(x)(\xi(x) - \nu(x))_+ \leq \inf_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+].$$

This establishes equation (8.38).

We aim at applying Sion's minimax theorem [38, Corollary 3.3]. To get prepared for this, first note that the supremum in (8.38) can obviously by performed with respect to those functions $u : \mathbb{X} \to \mathbb{R}$ satisfying $u(x) = 0$ whenever $x \notin \xi$. Note also that these functions form a finite dimensional real vector space (whose dimension is given by $\#\{x \in \xi\}$) which will be denoted by U. So the supremum is actually taken over

$$U_{\leq 1} = \{u \in U : \|u\|_\xi \leq 1\}$$

which is a convex and compact subset of U. Denote by Q the finite set of maps $q : \mathbb{X} \to \mathbb{N}_0$ satisfying $q(x) \leq \xi(x)$ for all $x \in \xi$ and $q(x) = 0$ whenever $x \notin \xi$. Moreover, define the map I by

$$I : A \to Q, \nu \mapsto (x \mapsto (\xi(x) - \nu(x))_+).$$

Then, for any $\zeta \in \mathcal{M}(A)$ and $x \in \xi$ we have

$$\mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+] = \mathbb{E}_{I\zeta(q)}[q(x)],$$

where $I\zeta$ denotes the pushforward measure of ζ with respect to I. Now, instead of taking the infimum in $\mathcal{M}(A)$ we can also minimize in the set of pushforward measures $I\mathcal{M}(A)$. The set $I\mathcal{M}(A)$ coincides with the set of probability measures on $I(A)$, denoted by $\mathcal{M}I(A)$. Observe that $\mathcal{M}I(A)$ is a convex and compact subset in the finite dimensional real vector space of all signed measures on $I(A)$ which we denote by $\mathcal{S}I(A)$. Obviously, the map

$$U \times \mathcal{S}I(A) \to \mathbb{R}, \ (u, \zeta) \mapsto \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(q)}[q(x)],$$

is both linear in u and ζ. Hence, it is also upper semicontinuous and quasi-concave in u and lower semicontinuous and quasi-convex in ζ. According to the above considerations, the assumptions of Sion's theorem are satisfied and we obtain

$$\sup_{\|u\|_\xi \leq 1} \inf_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+] = \sup_{u \in U_{\leq 1}} \inf_{\zeta \in \mathcal{M}I(A)} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(q)}[q(x)]$$

$$= \inf_{\zeta \in \mathcal{M}I(A)} \sup_{u \in U_{\leq 1}} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(q)}[q(x)] = \inf_{\zeta \in \mathcal{M}(A)} \sup_{\|u\|_\xi \leq 1} \sum_{x \in \xi} u(x) \mathbb{E}_{\zeta(\nu)}[(\xi(x) - \nu(x))_+].$$

Since both $U_{\leq 1}$ and $\mathcal{M}I(A)$ are compact, the suprema and infima are actually maxima and minima. $\qquad\square$

8.2. **Convex distance inequality.** In what follows, we will give the announced new proof of the convex distance inequality for Poisson point processes. The result we aim to prove is the following:

Theorem 8.2. *Let η be a Poisson point process in $\mathbb{X}$ with finite intensity measure μ. Let $A \in \mathbb{N}_{\mathrm{fin}}$ be arbitrary. Then*

$$\mathbb{P}(\eta \in A)\mathbb{E}(e^{d_T(\eta,A)^2/10}) \leq 1.$$

In particular, for any $r \geq 0$,

$$\text{(8.39)} \qquad\qquad \mathbb{P}(\eta \in A)\mathbb{P}(d_T(\eta, A) \geq r) \leq e^{-r^2/10}.$$

Note that in [30, Theorem 1.1] (under more restrictive assumptions on the space $(\mathbb{X}, \mathcal{X}, \mu)$), an inequality stronger than (8.39) is proved, where the constant $1/10$ is replaced by $1/4$. To get prepared for the proof of Theorem 8.2, we first establish the following result. This is interesting in its own right since it particularly states that the variance of the convex distance is bounded by 1.

Proposition 8.3. *Let η be a Poisson point process on $\mathbb{X}$ with finite intensity measure μ. Then for any $A \in \mathbb{N}_{\mathrm{fin}}$ almost surely*

$$V^+(d_T(\eta, A)) = \int_{\mathbb{X}} (D_x d_T(\eta - \delta_x, A))^2 d\eta(x) \leq 1.$$

In particular, $\mathbb{V}d_T(\eta, A) \leq 1$.

Proof. For this proof we adapt arguments from the proof of [4, Proposition 13]. According to Proposition 8.1 we can choose a map $\hat{u} : \mathbb{X} \to \mathbb{R}$ with $\|\hat{u}\|_\xi \leq 1$ and a probability measure $\hat{\zeta}$ on A satisfying

$$d_T(\xi, A) = \sum_{x \in \xi} \hat{u}(x)\mathbb{E}_{\hat{\zeta}(\nu)}[(\xi(x) - \nu(x))_+].$$

Then, for any $z \in \xi$, we have

$$d_T(\xi - \delta_z, A) \geq \min_{\zeta \in \mathcal{M}(A)} \sum_{x \in \xi - \delta_z} \hat{u}(x)\mathbb{E}_{\zeta(\nu)}[((\xi - \delta_z)(x) - \nu(x))_+].$$

Choose some $\tilde{\zeta} \in \mathcal{M}(A)$ that achieves the minimum in the above right-hand side. Then

$$d_T(\xi, A) \leq \sum_{x \in \xi} \hat{u}(x)\mathbb{E}_{\tilde{\zeta}(\nu)}[(\xi(x) - \nu(x))_+].$$

It follows that

$$d_T(\xi, A) - d_T(\xi - \delta_z, A)$$
$$\leq \sum_{x \in \xi} \hat{u}(x) \mathbb{E}_{\tilde{\zeta}(\nu)}[(\xi(x) - \nu(x))_+] - \sum_{x \in \xi - \delta_z} \hat{u}(x) \mathbb{E}_{\tilde{\zeta}(\nu)}[((\xi - \delta_z)(x) - \nu(x))_+]$$
$$= \hat{u}(z) \mathbb{E}_{\tilde{\zeta}(\nu)}[(\xi(z) - \nu(z))_+ - (\xi(z) - \nu(z) - 1)_+]$$
$$= \hat{u}(z) \mathbb{E}_{\tilde{\zeta}(\nu)}[\mathbb{1}\{\xi(z) > \nu(z)\}] \leq \hat{u}(z).$$

This yields

$$\int_{\mathbb{X}} (d_T(\xi, A) - d_T(\xi - \delta_z, A))^2 d\xi(z) \leq \int_{\mathbb{X}} \hat{u}(z)^2 \xi(z) = \|\hat{u}\|_\xi^2 \leq 1.$$

Hence, almost surely

$$\int_{\mathbb{X}} (D_x d_T(\eta - \delta_x, A))^2 d\eta(x) \leq 1.$$

Applying the Poincare Inequality for Poisson processes (see e.g. [43, Remark 1.4]) yields $\mathbb{V} d_T(\eta, A) \leq 1$. $\square$

As a final ingredient for the upcoming proof of the convex distance inequality, we derive the following consequence of the Cauchy-Schwarz Inequality.

Lemma 8.4. *Let $\xi \in \mathbf{N}_{\text{fin}}$ and consider the measure space $(\mathbb{X}, \mathcal{X}, \xi)$. Then for any measurable map $h : \mathbb{X} \to \mathbb{R}$,*

(8.40) $$\sup_{\|u\|_\xi \leq 1} \int_{\mathbb{X}} u(x) h(x) d\xi(x) = \|h\|_\xi.$$

Proof. Note that h is of course square-integrable with respect to ξ. Hence, by the Cauchy-Schwarz Inequality, for any u such that $\|u\|_\xi \leq 1$,

$$\int_{\mathbb{X}} u(x) h(x) d\xi(x) \leq \|u\|_\xi \|h\|_\xi \leq \|h\|_\xi.$$

We see that the left-hand side in (8.40) is less or equal to the right-hand side. Moreover, we can take $u = h/\|h\|_\xi$ to conclude that the right-hand side is less or equal to the left-hand side. $\square$

Proof of Theorem 8.2. Here we adapt arguments from the proofs of [5, Lemma 1 and Corollary 1]. We will prove below that

(8.41) $$0 \leq D_x(d_T(\xi, A)^2) \leq 2 \text{ for any } (x, \xi) \in \mathbb{X} \times \mathbf{N}_{\text{fin}}$$

(8.42) $$\text{and almost surely } V^+(d_T(\eta, A)^2) \leq 4 d_T(\eta, A)^2.$$

Hence, if follows from Theorem 3.7 and Theorem 3.10 (where the latter result is applied to the Poisson functional $\frac{1}{2} d_T(\eta, A)^2$) that

(i) For any $\lambda \in (0, 1/2)$,

$$\log \mathbb{E}(\exp(\lambda(d_T(\eta, A)^2 - \mathbb{E} d_T(\eta, A)^2))) \leq \frac{2\lambda^2 \mathbb{E} d_T(\eta, A)^2}{1 - 2\lambda},$$

(ii) For any $r \geq 0$,

$$\mathbb{P}(d_T(\eta, A)^2 \leq \mathbb{E}d_T(\eta, A)^2 - r) \leq \exp\left(-\frac{r^2}{8\mathbb{E}d_T(\eta, A)^2}\right).$$

Taking $\lambda = 1/10$ we obtain from (i) that

$$\mathbb{E}e^{d_T(\eta, A)^2/10} \leq \exp\left(\frac{\mathbb{E}d_T(\eta, A)^2}{8}\right).$$

Moreover, since $\eta \in A$ implies $d_T(\eta, A) = 0$, it follows from (ii) with $r = \mathbb{E}d_T(\eta, A)^2$ that

$$\mathbb{P}(\eta \in A) \leq \mathbb{P}\left(d_T(\eta, A)^2 \leq \mathbb{E}d_T(\eta, A)^2 - \mathbb{E}d_T(\eta, A)^2\right) \leq \exp\left(-\frac{\mathbb{E}d_T(\eta, A)^2}{8}\right).$$

So, the result follows once we have proven (8.41) and (8.42). To prove (8.42), first observe that $d_T(\cdot, A)$ is a non-decreasing functional. Using this and Proposition 8.3 we compute

$$V^+(d_T(\eta, A)^2) = \int_{\mathbb{X}} (d_T(\eta, A)^2 - d_T(\eta - \delta_x, A)^2)^2 \, d\eta(x)$$

$$= \int_{\mathbb{X}} (d_T(\eta, A) - d_T(\eta - \delta_x, A))^2 (d_T(\eta, A) + d_T(\eta - \delta_x, A))^2 \, d\eta(x)$$

$$\leq \int_{\mathbb{X}} (d_T(\eta, A) - d_T(\eta - \delta_x, A))^2 4d_T(\eta, A)^2 \, d\eta(x)$$

$$= 4d_T(\eta, A)^2 \int_{\mathbb{X}} (D_x d_T(\eta - \delta_x, A))^2 \, d\eta(x) \leq 4d_T(\eta, A)^2.$$

It remains to prove (8.41). For this, let $(z, \xi) \in \mathbb{X} \times \mathbf{N}_{\text{fin}}$. Then, according to Proposition 8.1, we can write

$$d_T(\xi, A) = \max_{\|u\|_\xi \leq 1} \sum_{x \in \xi} u(x) \mathbb{E}_{\hat{\zeta}(\nu)}[(\xi(x) - \nu(x))_+]$$

$$= \max_{\|u\|_\xi \leq 1} \int_{\mathbb{X}} u(x) \mathbb{E}_{\hat{\zeta}(\nu)} \left[\left(1 - \frac{\nu(x)}{\xi(x)}\right)_+\right] \, d\xi(x)$$

for some probability measure $\hat{\zeta}$ on A. By virtue of Lemma 8.4, the latter expression equals

$$\sqrt{\int_{\mathbb{X}} \left(\mathbb{E}_{\hat{\zeta}(\nu)} \left[\left(1 - \frac{\nu(x)}{\xi(x)}\right)_+\right]\right)^2 \, d\xi(x)}.$$

Invoking Proposition 8.1 and again Lemma 8.4, we also obtain

$$d_T(\xi + \delta_z, A) \leq \max_{\|u\|_{\xi+\delta_z} \leq 1} \sum_{x \in \xi + \delta_z} u(x) \mathbb{E}_{\hat{\zeta}(\nu)}[((\xi + \delta_z)(x) - \nu(x))_+]$$

$$= \sqrt{\int_{\mathbb{X}} \left(\mathbb{E}_{\hat{\zeta}(\nu)} \left[\left(1 - \frac{\nu(x)}{(\xi + \delta_z)(x)}\right)_+\right]\right)^2 \, d(\xi + \delta_z)(x)}.$$

From this it follows that

$$D_z d_T^2 \leq \left(\mathbb{E}_{\hat{\zeta}(\nu)} \left[\left(1 - \frac{\nu(z)}{\xi(z)+1} \right)_+ \right] \right)^2 (\xi(z)+1) - \left(\mathbb{E}_{\hat{\zeta}(\nu)} \left[\left(1 - \frac{\nu(z)}{\xi(z)} \right)_+ \right] \right)^2 \xi(z)$$

where the subtrahend vanishes whenever $\xi(z) = 0$. Clearly, if $\xi(z) = 0$, then the right-hand side in the above display is less or equal to 1. So assume that $\xi(z) > 0$. Then, using the abbreviations $G(\nu, z) = (\xi(z) - \nu(z) + 1)_+$ and $G'(\nu, z) = (\xi(z) - \nu(z))_+$, one observes that the right-hand side in the last display can be upper bounded by

$$\frac{(\mathbb{E}_{\hat{\zeta}(\nu)} G(\nu, z))^2 - (\mathbb{E}_{\hat{\zeta}(\nu)} G'(\nu, z))^2}{\xi(z)+1} = \frac{\mathbb{E}_{\hat{\zeta}(\nu)}[G(\nu, z) - G'(\nu, z)]\, \mathbb{E}_{\hat{\zeta}(\nu)}[G(\nu, z) + G'(\nu, z)]}{\xi(z)+1}$$

$$\leq \frac{\mathbb{E}_{\hat{\zeta}(\nu)}[G(\nu, z) + G'(\nu, z)]}{\xi(z)+1}$$

$$= \mathbb{E}_{\hat{\zeta}(\nu)} \left[\frac{(\xi(z) - \nu(z) + 1)_+ + (\xi(z) - \nu(z))_+}{\xi(z)+1} \right]$$

$$\leq 2.$$

It follows that $D_z(d_T(\xi, A)^2) \leq 2$.

The functional $d_T(\cdot, A)$ is non-decreasing and non-negative, thus $D_z(d_T(\xi, A)^2) \geq 0$. This concludes the proof. $\qquad\square$

References

[1] C. Ané and M. Ledoux. "On logarithmic Sobolev inequalities for continuous time random walks on graphs". In: *Probab. Theory Related Fields* 116.4 (2000), pp. 573–602.

[2] S. G. Bobkov and M. Ledoux. "On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures". In: *J. Funct. Anal.* 156.2 (1998), pp. 347–365.

[3] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities.* A non-asymptotic theory of independence, With a foreword by Michel Ledoux. Oxford University Press, Oxford, 2013, pp. x+481.

[4] S. Boucheron, G. Lugosi, and P. Massart. "Concentration inequalities using the entropy method". In: *Ann. Probab.* 31.3 (2003), pp. 1583–1614.

[5] S. Boucheron, G. Lugosi, and P. Massart. "On concentration of self-bounding functions". In: *Electron. J. Probab.* 14 (2009), no. 64, 1884–1899.

[6] S. Bourguin and G. Peccati. "Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering". In: *Electron. J. Probab.* 19 (2014), no. 66, 42.

[7] J.-Ch. Breton, Ch. Houdré, and N. Privault. "Dimension free and infinite variance tail estimates on Poisson space". In: *Acta Appl. Math.* 95.3 (2007), pp. 151–203.

[8] D. Chafaï. "Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue". In: *ESAIM Probab. Stat.* 10 (2006), 317–339 (electronic).

[9] L. Decreusefond, E. Ferraz, H. Randriambololona, and A. Vergne. "Simplicial homology of random configurations". In: *Adv. in Appl. Probab.* 46.2 (2014), pp. 325–347.

[10] D. P. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press, Cambridge, 2009, pp. xvi+196.

[11] P. Eichelsbacher, M. Raič, and T. Schreiber. "Moderate deviations for stabilizing functionals in geometric probability". In: *Ann. Inst. Henri Poincaré Probab. Stat.* 51.1 (2015), pp. 89–128.

[12] P. Eichelsbacher and Ch. Thäle. "New Berry-Esseen bounds for non-linear functionals of Poisson random measures". In: *Electron. J. Probab.* 19 (2014), no. 102, 25.

[13] E. N. Gilbert. "Random graphs". In: *Ann. Math. Statist.* 30 (1959), pp. 1141–1144.

[14] M. Heveling and M. Reitzner. "Poisson-Voronoi approximation". In: *Ann. Appl. Probab.* 19.2 (2009), pp. 719–736.

[15] Ch. Houdré and N. Privault. "Concentration and deviation inequalities in infinite dimensions via covariance representations". In: *Bernoulli* 8.6 (2002), pp. 697–720.

[16] Ch. Houdré and P. Reynaud-Bouret. "Exponential inequalities, with constants, for U-statistics of order two". In: *Stochastic inequalities and applications.* Vol. 56. Progr. Probab. Birkhäuser, Basel, 2003, pp. 55–69.

[17] S. Janson. "Bounds on the distributions of extremal values of a scanning process". In: *Stochastic Process. Appl.* 18.2 (1984), pp. 313–328.

[18] O. Kallenberg. *Foundations of modern probability.* Second. Probability and its Applications (New York). Springer-Verlag, New York, 2002, pp. xx+638.

[19] R. Lachièze-Rey and G. Peccati. "Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs". In: *Electron. J. Probab.* 18 (2013), no. 32, 32.

[20] R. Lachièze-Rey and G. Peccati. "Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics". In: *Stochastic Process. Appl.* 123.12 (2013), pp. 4186–4218.

[21] R. Lachièze-Rey and M. Reitzner. *U-statistics in stochastic geometry.* Version 2. Mar. 26, 2015. arXiv: `1503.00110v2 [math.PR]`.

[22] G. Last and M. D. Penrose. "Poisson process Fock space representation, chaos expansion and covariance inequalities". In: *Probab. Theory Related Fields* 150.3-4 (2011), pp. 663–690.

[23] M. Ledoux. "Concentration of measure and logarithmic Sobolev inequalities". In: *Séminaire de Probabilités, XXXIII.* Vol. 1709. Lecture Notes in Math. Springer, Berlin, 1999, pp. 120–216.

[24] M. Ledoux. *The concentration of measure phenomenon.* Vol. 89. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2001, pp. x+181.

[25] P. Massart. "About the constants in Talagrand's concentration inequalities for empirical processes". In: *Ann. Probab.* 28.2 (2000), pp. 863–884.

[26] A. Maurer. "Concentration inequalities for functions of independent variables". In: *Random Structures Algorithms* 29.2 (2006), pp. 121–138.

[27] G. Peccati and M. S. Taqqu. *Wiener chaos: moments, cumulants and diagrams.* Vol. 1. Bocconi & Springer Series. A survey with computer implementation, Supplementary material available online. Springer, Milan; Bocconi University Press, Milan, 2011, pp. xiv+274.

[28] G. Peccati and Ch. Thäle. "Gamma limits and U-statistics on the Poisson space". In: *ALEA Lat. Am. J. Probab. Math. Stat.* 10.1 (2013), pp. 525–560.

[29] M. D. Penrose. *Random geometric graphs.* Vol. 5. Oxford Studies in Probability. Oxford University Press, Oxford, 2003, pp. xiv+330.

[30] M. Reitzner. "Poisson point processes: large deviation inequalities for the convex distance". In: *Electron. Commun. Probab.* 18 (2013), no. 96, 7.

[31] M. Reitzner and M. Schulte. "Central limit theorems for U-statistics of Poisson point processes". In: *Ann. Probab.* 41.6 (2013), pp. 3879–3909.

[32] M. Reitzner, M. Schulte, and Ch. Thäle. *Limit theory for the Gilbert graph.* Version 1. Oct. 17, 2013. arXiv: `312.4861` [`math.PR`].

[33] P. Reynaud-Bouret. "Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities". In: *Probab. Theory Related Fields* 126.1 (2003), pp. 103–153.

[34] I. Rivin. "Counting cycles and finite dimensional L^p norms". In: *Adv. in Appl. Math.* 29.4 (2002), pp. 647–662.

[35] R. Schneider and W. Weil. *Stochastic and integral geometry.* Probability and its Applications (New York). Springer-Verlag, Berlin, 2008, pp. xii+693.

[36] M. Schulte. "Malliavin-Stein Method in Stochastic Geometry". PhD thesis. Universität Osnabrück, 2013.

[37] M. Schulte and Ch. Thäle. "The scaling limit of Poisson-driven order statistics with applications in geometric probability". In: *Stochastic Process. Appl.* 122.12 (2012), pp. 4096–4120.

[38] M. Sion. "On general minimax theorems". In: *Pacific J. Math.* 8 (1958), pp. 171–176.

[39] J. M. Steele. *Probability theory and combinatorial optimization.* Vol. 69. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997, pp. viii+159.

[40] D. Surgailis. "On multiple Poisson stochastic integrals and associated Markov semigroups". In: *Probab. Math. Statist.* 3.2 (1984), pp. 217–239.

[41] M. Talagrand. "Concentration of measure and isoperimetric inequalities in product spaces". In: *Inst. Hautes Études Sci. Publ. Math.* 81 (1995), pp. 73–205.

[42] T. Tao. *Topics in random matrix theory.* Vol. 132. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012, pp. x+282.

[43] L. Wu. "A new modified logarithmic Sobolev inequality for Poisson point processes and several applications". In: *Probab. Theory Related Fields* 118.3 (2000), pp. 427–438.

Sascha Bachmann
Institut für Mathematik
Universität Osnabrück
49069 Osnabrück, Germany
Email: sascha.bachmann@uos.de

Giovanni Peccati
Unité de Recherche en Mathématiques
Faculté des Sciences, de la Technologie
et de la Communication, Université du Luxembourg
Email: giovanni.peccati@gmail.com
Webpage: https://sites.google.com/site/giovannipeccati/Home