



# Repelled point processes with application to numerical integration

Diala Hawat, Rémi Bardenet, Raphaël Lachièze-Rey

## ► To cite this version:

Diala Hawat, Rémi Bardenet, Raphaël Lachièze-Rey. Repelled point processes with application to numerical integration. 2023. hal-04179981

**HAL Id: hal-04179981**

**<https://hal.science/hal-04179981>**

Preprint submitted on 10 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Repelled point processes with application to numerical integration

DIALA HAWAT<sup>1,2,a</sup>, RÉMI BARDENET<sup>1,b</sup> and RAPHAËL LACHIEZE-REY<sup>2,c</sup>

<sup>1</sup>Université de Lille, CNRS, Centrale Lille, UMR 9189 – CRISTAL, Lille, France, <sup>a</sup>[diala.hawat@univ-lille.fr](mailto:diala.hawat@univ-lille.fr),

<sup>b</sup>[remi.bardenet@univ-lille.fr](mailto:remi.bardenet@univ-lille.fr)

<sup>2</sup>Université Paris Cité, MAP5, Paris, France, <sup>c</sup>[raphael.lachieze-rey@math.cnrs.fr](mailto:raphael.lachieze-rey@math.cnrs.fr)

Linear statistics of point processes yield Monte Carlo estimators of integrals. While the simplest approach relies on a homogeneous Poisson point process, more regularly spread point processes, such as scrambled low-discrepancy sequences or determinantal point processes, can yield Monte Carlo estimators with fast-decaying mean square error. Following the intuition that more regular configurations result in lower integration error, we introduce the repulsion operator, which reduces clustering by slightly pushing the points of a configuration away from each other. Our main theoretical result is that applying the repulsion operator to a homogeneous Poisson point process yields an unbiased Monte Carlo estimator with *lower* variance than under the original point process. On the computational side, the evaluation of our estimator is only quadratic in the number of integrand evaluations and can be easily parallelized without any communication across tasks. We illustrate our variance reduction result with numerical experiments and compare it to popular Monte Carlo methods. Finally, we numerically investigate a few open questions on the repulsion operator. In particular, the experiments suggest that the variance reduction also holds when the operator is applied to other motion-invariant point processes.

**Keywords:** Point processes; Monte Carlo integration; Coulomb interaction; Variance reduction

## Contents

1	Introduction . . . . .	2
2	Point processes and their intensity functions . . . . .	3
2.1	Spatial point processes . . . . .	3
2.2	The homogeneous Poisson point process (PPP) . . . . .	4
3	Repelled point processes . . . . .	4
3.1	The repulsion operator . . . . .	5
3.2	Properties of the repelled Poisson point process . . . . .	6
3.3	Main result . . . . .	8
3.4	Properties of the force . . . . .	9
3.5	Sampling from the repelled Poisson point process . . . . .	10
3.6	An experimental illustration of the variance reduction . . . . .	11
4	Application to numerical integration . . . . .	13
4.1	A few Monte Carlo methods . . . . .	13
4.2	Experimental comparison . . . . .	15
5	Other models and properties . . . . .	17
5.1	Iterating the repulsion . . . . .	17
5.2	Second-order properties of the repelled Poisson point process . . . . .	19
5.3	Other repelled point processes . . . . .	20
6	Conclusion . . . . .	22
	Acknowledgements . . . . .	23
	References . . . . .	23
A	Proofs . . . . .	24
A.1	Proofs of Proposition 5 . . . . .	25
A.2	Proof of Proposition 1 and Corollary 1 (motion-invariance) . . . . .	26

A.3	Proof of Proposition 2 (existence of the moments)	28
A.4	Proof of Theorem 3 (variance reduction)	40
B	Additional experiments	48
B.1	Further analysis of Section 4.2	48
B.2	Further analysis of Section 6	48

## 1. Introduction

Numerical integration – the task of approximating integrals using pointwise evaluations of the integrand – has a rich history, with (subjective) milestones such as the quadrature of Gauss (1815) or the Metropolis algorithm (Metropolis et al., 1953). Among methods that scale best to larger dimensions, Monte Carlo methods may be the most popular in applications; see e.g. (Owen, 2013). In its simplest form, crude Monte Carlo amounts to writing the target integral as an expectation and estimating it by a sample average formed with  $N$  i.i.d. samples. The performance of the estimator is typically measured by the mean squared error, which classical probability results show decreases as  $O(1/N)$ . This rate is slow, and variance reduction for Monte Carlo integration has been a rich topic of research, e. g. by leveraging auxiliary integrands with known integrals, a method known as *control variates*, see (Owen, 2013, Chapters 8 and 10) for classical results, and (South et al., 2022) and references therein for more recent work. Other methods directly improve the rate of convergence of the mean square error by introducing sophisticated dependence through weights in the sample average (Azaïs et al., 2018, Delyon and Portier, 2016, Leluc et al., 2023). Alternatively, one can consider expectations under more regularly spread probability distributions than i.i.d. draws, as in randomized quasi-Monte Carlo (Owen, 2008) or with determinantal point processes (Bardenet and Hardy, 2020, Belhadji et al., 2020, Coeurjolly et al., 2021). Our contribution is of the latter kind, taking an expectation under a well-spread random set of points. But while the cost of sampling the determinantal point process in (Bardenet and Hardy, 2020) is (at least) cubic in  $N$  (Gautier et al., 2019a), we aim at keeping the sampling cost quadratic, while still introducing enough dependence to force variance reduction.

Our inspiration comes from a set of results in stochastic geometry, on gravitation allocation (Chatterjee et al., 2010, Nazarov et al., 2007), hyperuniform point processes (Coste, 2021, Klatt et al., 2019, Torquato, 2018), and systems of points with Coulomb interactions (Serfaty, 2019). Loosely speaking and for our purpose, a point process is a random locally finite set of points in  $\mathbb{R}^d$ . A natural reference is the homogeneous Poisson point process, whose points in a compact set are i.i.d. uniform random variables, once we condition on the number of points in that compact set. At the other end of the spectrum, the counting statistics of *hyperuniform* point processes yield Monte Carlo estimators with mean square error decaying at a faster rate than the Poisson point process. For instance, Klatt et al. (2019) study the candidate hyperuniform point process that results from applying Lloyd’s algorithm iteratively to a Poisson point process. In a similar vein but leaving aside any iterative scheme for mathematical tractability, we propose to use gravitational allocation to yield, if not necessarily a faster rate, at least a provable variance reduction compared to Poisson. Imagining that the points of a Poisson point process repel each other through a Coulomb interaction, we perform a *single* step of a numerical scheme to integrate the corresponding differential equation. We obtain what we call the repelled Poisson point process, and we show that, for a quadratic cost, the variance of the resulting Monte Carlo estimator is smaller than under the initial Poisson point process. We note that *repelled point processes*, resulting from the same repulsion procedure applied to an arbitrary point process, are objects of independent interest.

The paper is structured as follows. Section 2 provides background information about point processes. Section 3 introduces the repulsion operator, a parametric operator that operates on locally finite sets of points in  $\mathbb{R}^d$ . We analyze the properties of this operator when applied to a homogeneous Poisson point

process, and present our main result, showing variance reduction compared to crude Monte Carlo. Additionally, we describe a sampling procedure and present an experimental illustration of the variance reduction. In Section 4, we put our method in context, by conducting a comparison with standard Monte Carlo methods on synthetic integration tasks. In Section 5, we explore additional aspects of the repulsion operator, such as iterating it several times, estimating the pair correlation function and structure factor of the repelled Poisson point process, as well as applying the repulsion operator to already repulsive point processes like the Ginibre point process. Section 6 concludes the paper with a few research directions. All proofs are gathered in Appendix A, while Appendix B gives extra simulation results.

## 2. Point processes and their intensity functions

In this section, we provide some background on point processes, with key results like the Slyvniak-Mecke theorem. We refer to (Stoyan et al., 2013) for details.

**Remark 1.** Throughout this manuscript, bold lowercase letters, like  $\mathbf{x}$ , indicate vectors in  $\mathbb{R}^d$ . The corresponding non-bold characters, like  $x$ , are scalars. In particular, we write  $\mathbf{x} = (x_1, \dots, x_d)$ . Whenever not confusing, we use the same letter in different fonts for a vector and its Euclidean norm, i.e.,  $r = \|\mathbf{r}\|_2$  and  $k = \|\mathbf{k}\|_2$ . Calligraphic letters like  $\mathcal{X}$  are used for point processes, i.e. random configurations of points. Configurations themselves are denoted with sans-serif letters like  $X$ . When the cardinality of a particular point process is almost surely constant, we sometimes write that point process as  $\mathcal{X}_N$ , with the value  $N$  of the cardinality as the index.

### 2.1. Spatial point processes

Let a configuration of  $\mathbb{R}^d$  be a locally finite set  $X$  of  $\mathbb{R}^d$ , that is, for any bounded Borel set  $B$  of  $\mathbb{R}^d$  the cardinality  $X(B)$  of  $X \cap B$  is finite. Endow the family  $\mathfrak{R}$  of such configurations with the  $\sigma$ -algebra generated by the mappings  $X \mapsto X(B)$ , for any bounded Borel set  $B$ . Formally, a *simple<sup>1</sup> spatial point process*, hereafter a *point process*, is a random element  $\mathcal{X}$  of  $\mathfrak{R}$ . The distribution of  $\mathcal{X}$  is determined by the system of *void probabilities*

$$\mathbb{T}_{\mathcal{X}}(K) \triangleq \mathbb{P}(\mathcal{X}(K) = 0), \quad (1)$$

as  $K$  ranges through the compact sets of  $\mathbb{R}^d$ .

In this paper, we work with *stationary* and *isotropic* point processes, also called *motion-invariant* point processes. By stationary, we mean that the law of the point process is invariant by translation: the law of  $\mathcal{X}$  is identical to that of  $\mathcal{X} + \mathbf{y} \triangleq \{\mathbf{x} + \mathbf{y}; \mathbf{x} \in \mathcal{X}\}$ , for all  $\mathbf{y} \in \mathbb{R}^d$ . Similarly, a point process is isotropic if its law is invariant by rotation.

Point processes are often described in terms of their intensity measures. The *first intensity measure*  $\mu^{(1)}$ , for instance, is defined by

$$\mu^{(1)}(B) \triangleq \mathbb{E}[\mathcal{X}(B)],$$

for any Borel set  $B \subset \mathbb{R}^d$ . When  $\mu^{(1)}$  has a density with respect to the Lebesgue measure,  $\mu^{(1)}(d\mathbf{x}) = \rho^{(1)}(\mathbf{x}) d\mathbf{x}$ , we call  $\rho^{(1)}$  the *intensity* of  $\mathcal{X}$ . If  $\mathcal{X}$  is stationary, then  $\mu^{(1)}$  is proportional to the Lebesgue measure and the intensity  $\rho^{(1)}$  is a positive constant  $\rho > 0$ , equal to the mean number of points of  $\mathcal{X}$  per unit volume.

<sup>1</sup>The term *simple* indicates that the point process almost surely consists of distinct points. For us, this is a direct consequence of defining configurations as sets. However, some authors avoid assuming simplicity; see e.g. (Last and Penrose, 2017, Chapter 6).

More generally, the  $n$ -th order intensity measure  $\mu^{(n)}$  of  $\mathcal{X}$  is defined by

$$\mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{X}}^{\#} f(\mathbf{x}_1, \dots, \mathbf{x}_n) \right] = \int_{\mathbb{R}^d \times \dots \times \mathbb{R}^d} f(\mathbf{x}_1, \dots, \mathbf{x}_n) \mu^{(n)}(d\mathbf{x}_1 \dots d\mathbf{x}_n), \quad (2)$$

where  $f$  is any non-negative bounded measurable function, and the summation is over all  $n$ -tuples of distinct points in  $\mathcal{X}$ ; see (Last and Penrose, 2017, Stoyan et al., 2013, Chapter 4). Again, when  $\mu^{(n)} = \rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \dots d\mathbf{x}_n$ ,  $\rho^{(n)}$  is called the  $n$ -th order intensity function. Intuitively, for any pairwise distinct points  $\mathbf{x}_1, \dots, \mathbf{x}_n$ ,  $\rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \dots d\mathbf{x}_n$  is the probability that  $\mathcal{X}$  has a point in each of the  $n$  infinitesimally small sets around  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , with respective volumes  $d\mathbf{x}_1, \dots, d\mathbf{x}_n$ .

## 2.2. The homogeneous Poisson point process (PPP)

Consider a compact set  $K$  of  $\mathbb{R}^d$ , and pick  $N$  i.i.d. points uniformly distributed in  $K$ . The point process  $\mathcal{X}_N$  formed by these  $N$  points is called the Binomial point process (BPP) of  $N$  points. Loosely speaking, when  $K$  is enlarged to fill out  $\mathbb{R}^d$  while maintaining  $N = \rho|K|$ , we obtain a limiting point process  $\mathcal{P}$  that is called the homogeneous Poisson point process (PPP) of intensity  $\rho > 0$ . We now list a few properties of  $\mathcal{P}$ .

First,  $\mathcal{P}$  is motion-invariant. Second, the random number  $\mathcal{P}(B)$  of points of  $\mathcal{P}$  in a bounded Borel set  $B$  has a Poisson distribution of mean  $\rho|B|$ , where  $|B|$  is the Lebesgue measure of  $B$ . In particular, the void probability (1) is  $\mathbb{T}_{\mathcal{P}}(B) = \exp(-\rho|B|)$ . Third,  $\mathcal{P}(B_1)$  and  $\mathcal{P}(B_2)$  are independent for any disjoint Borel sets  $B_1$  and  $B_2$ . This second fundamental property is known as *complete randomness*, and translates the intuition that the PPP has as little structure as possible; for more details see (Stoyan et al., 2013, Chapter 2). Fourth, all the moments of  $\mathcal{P}$  are determined by  $\rho$ , i.e., for any non-negative measurable function  $f$ ,

$$\mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{P}}^{\#} f(\mathbf{x}_1, \dots, \mathbf{x}_n) \right] = \rho^n \int_{\mathbb{R}^d \times \dots \times \mathbb{R}^d} f(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \dots d\mathbf{x}_n. \quad (3)$$

With the notation of Section 2.1, this reads  $\rho^{(n)} = \rho^n$ . Moreover, by the so-called extended Slivnyak-Mecke theorem, we have

$$\mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{P}}^{\#} h(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathcal{P} \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \right] = \rho^n \int_{\mathbb{R}^d \times \dots \times \mathbb{R}^d} \mathbb{E}[h(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathcal{P})] d\mathbf{x}_1 \dots d\mathbf{x}_n, \quad (4)$$

for any non-negative measurable function  $h$  on  $(\mathbb{R}^d)^n \times \mathfrak{R}$ ; see (Coeurjolly et al., 2017, Section 5.1). Equation (4) provides further evidence of the PPP's lack of dependency structure: informally, once conditioned on a finite number of points belonging to  $\mathcal{P}$ , the rest of  $\mathcal{P}$  has the same distribution as  $\mathcal{P}$ .

## 3. Repelled point processes

In this section, given a configuration  $\mathbf{X} \in \mathfrak{R}$  and a parameter  $\varepsilon > 0$ , we explain how to construct another configuration  $\Pi_{\varepsilon}(\mathbf{X})$ , called the *repelled configuration*. Keeping in mind our motivation for numerical integration, we want  $\Pi_{\varepsilon}$  to be (i) computationally cheap to apply. Moreover, when applied to a random

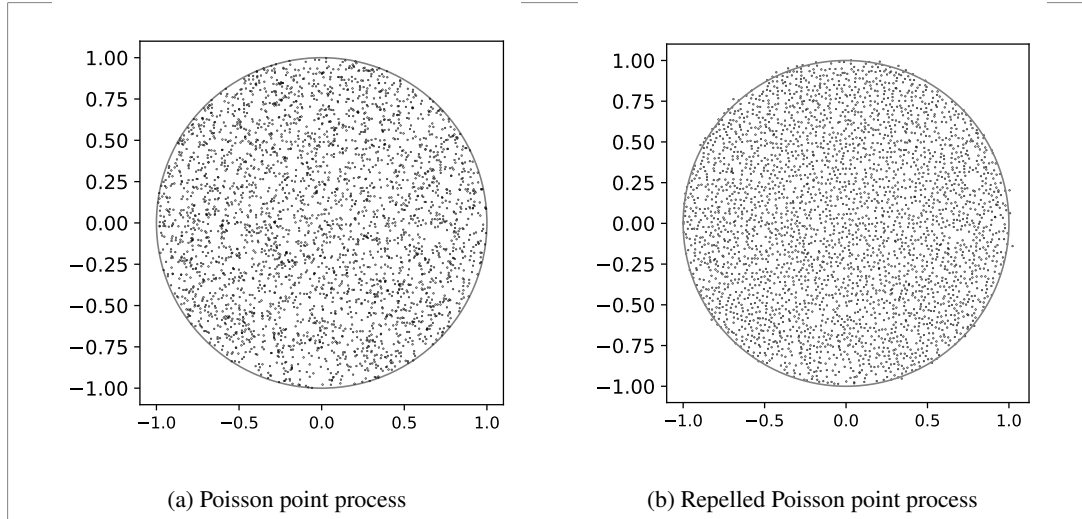


Figure 1: A sample from a PPP of intensity  $\rho = 1000$  and the corresponding repelled sample.

configuration,  $\Pi_\varepsilon$  should (ii) preserve stationarity, isotropy, and intensity, and (iii) reduce the variance of linear statistics.

In Section 3.1, we define the repulsion operator  $\Pi_\varepsilon$ . In Section 3.2, we detail the properties of  $\Pi_\varepsilon(\mathcal{P})$ , where  $\mathcal{P}$  is a homogeneous Poisson point process. Our main result is in Section 3.3, where we show that, for small enough  $\varepsilon$ , the variance of linear statistics under  $\Pi_\varepsilon \mathcal{P}$  is smaller than under  $\mathcal{P}$ . In particular, we give theoretical support for the choice of a particular value of  $\varepsilon$ , which is both explicit and independent of the considered linear statistic. Our repulsion operator is based on a stochastic process known as Coulomb force. We discuss basic properties of the latter in Section 3.4. In Section 3.5, we explain how to *approximately* sample from  $\Pi_\varepsilon \mathcal{X}$ , in time quadratic in the number of points of the point process  $\mathcal{X}$  in the observation window. Finally, Section 3.6 gives a first experimental illustration of our variance reduction result.

### 3.1. The repulsion operator

For  $\mathbf{x} \in \mathbb{R}^d$  and a configuration  $X \in \mathfrak{N}$ , consider the series (when it converges),

$$F_X(\mathbf{x}) \triangleq \sum_{\substack{\mathbf{z} \in X \setminus \{\mathbf{x}\} \\ \|\mathbf{x} - \mathbf{z}\|_2 \uparrow}} \frac{\mathbf{x} - \mathbf{z}}{\|\mathbf{x} - \mathbf{z}\|_2^d} = \lim_{R \rightarrow \infty} \sum_{\mathbf{z} \in X \setminus \{\mathbf{x}\} \cap B(\mathbf{x}, R)} \frac{\mathbf{x} - \mathbf{z}}{\|\mathbf{x} - \mathbf{z}\|_2^d}. \quad (F_1)$$

Several observations are in order. First, each term in the sum in  $(F_1)$  intuitively represents the Coulomb force felt by a charged particle at  $\mathbf{x}$  and due to a particle of the same charge placed at  $\mathbf{z}$ . In a dynamic setting, this force would repel  $\mathbf{x}$  away from  $\mathbf{z}$ . Second, as the series defining  $F_X(\mathbf{x})$  is not absolutely convergent, the order of the summation is important. Following Chatterjee et al., 2010, we consider the limit in an increasing ball centered at  $\mathbf{x}$ , i.e., the summands in  $(F_1)$  are arranged in order of increasing distance from  $\mathbf{x}$ . We will discuss in Section 3.4 rearranging the summation by increasing distance from the origin. Third, a fundamental insight, originally mentioned by Chandrasekhar (1943), states that if  $d \geq 3$  and  $\mathcal{P}$  is a homogeneous Poisson point process (PPP), then, for every  $\mathbf{x}$ , the series defining

$F_{\mathcal{P}}(\mathbf{x})$  converges almost surely. Further information regarding the characteristics of  $F_{\mathcal{P}}$  can be found in Section 3.4.

Call  $X \in \mathfrak{N}$  a *valid* configuration if, for all  $\mathbf{x}$ , the limit defining  $F_X(\mathbf{x})$  in  $(F_1)$  exists. For  $\varepsilon \in \mathbb{R}$ , we define the (Coulomb) repulsion<sup>2</sup> operator  $\Pi_{\varepsilon}$ , acting on valid configurations, through

$$\Pi_{\varepsilon} : X \mapsto \{\mathbf{x} + \varepsilon F_X(\mathbf{x}) : \mathbf{x} \in X\}. \quad (5)$$

There are two formal caveats to our definition (5). First, it only applies to valid configurations. Second, since by definition,  $\Pi_{\varepsilon}X$  is a set, it does not keep track of multiplicities, arising when several points in  $X$  are mapped to the same location by  $\Pi_{\varepsilon}$ . Anticipating a bit, Corollary 1 below shows that these two caveats are irrelevant when  $\Pi_{\varepsilon}$  is applied to a homogeneous Poisson point process  $\mathcal{P}$ . In particular,  $\mathcal{P}$  is almost surely a valid configuration, and for any two distinct points  $\mathbf{x}, \mathbf{y} \in \mathcal{P}$ , almost surely

$$\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) \neq \mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y}).$$

This guarantees that  $\Pi_{\varepsilon}\mathcal{P}$  is a simple point process, which we term the *repelled Poisson point process* (RPPP). We will occasionally consider the repelled point process  $\Pi_{\varepsilon}X$  of a more general point process  $X$ , although its existence needs to be discussed.

The first panel of Figure 1 displays a sample from a PPP of intensity  $\rho = 1000$  in  $d = 2$ , intersected with a disk-shaped observation window. Note that we plot the construction in dimension 2 for graphical convenience, but we are not making any convergence claim for  $(F_1)$  in  $d = 2$ . We illustrate the RPPP construction in the second panel of the figure. A detailed explanation of the simulation procedure will be provided in Section 3.5. At this stage, we simply observe that the repelled sample exhibits a reduced tendency for points to cluster together, compared to the Poisson sample.

### 3.2. Properties of the repelled Poisson point process

In this section, we state some properties of the repulsion operator when it is applied to a homogeneous Poisson point process.

**Proposition 1 (Motion-invariance).** *Let  $X$  be a point process that is almost surely valid, and  $\varepsilon \in \mathbb{R}$ . If  $X$  is motion-invariant, then  $\Pi_{\varepsilon}X$  is also motion-invariant.*

The proof of this proposition is deferred to Appendix A.2. Of particular importance to us is the following corollary.

**Corollary 1.** *Let  $d \geq 3$ , and  $\mathcal{P} \subset \mathbb{R}^d$  be a homogeneous Poisson point process of intensity  $\rho > 0$ . Then, for any  $\varepsilon \in \mathbb{R}$ , and any two distinct points  $\mathbf{x}, \mathbf{y} \in \mathcal{P}$ , we have, almost surely,*

$$\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) \neq \mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y}).$$

Moreover,  $\Pi_{\varepsilon}\mathcal{P}$  is a stationary and isotropic point process of intensity  $\rho$ .

The proof of this corollary is also deferred to Appendix A.2. According to Corollary 1,  $\Pi_{\varepsilon}\mathcal{P}$  is of intensity  $\rho$ , the same intensity as  $\mathcal{P}$ . Consequently, for any integrable function  $f$  of compact support

<sup>2</sup>While we generally speak of “repulsion”, note that when  $\varepsilon < 0$  the dynamics become attractive instead of repulsive.

$K$ , Equation (3) yields

$$\mathbb{E} \left[ \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} f(\mathbf{x}) \right] = \rho \int_K f(\mathbf{x}) d\mathbf{x}.$$

In particular,

$$\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f) \triangleq \frac{1}{\rho} \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} f(\mathbf{x}) \quad (6)$$

is an unbiased estimator of

$$I_K(f) \triangleq \int_K f(\mathbf{x}) d\mathbf{x}. \quad (7)$$

We shall also consider the so-called “self-normalized” estimator

$$\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}(f) \triangleq \frac{|K|}{\Pi_\varepsilon \mathcal{P}(K)} \mathbb{1}_{\{\Pi_\varepsilon \mathcal{P}(K) > 0\}} \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P} \cap K} f(\mathbf{x}), \quad (8)$$

where  $\Pi_\varepsilon \mathcal{P}(K)$  is the number of points of  $\Pi_\varepsilon \mathcal{P}$  in  $K$ . Compared to (6), (8) replaces  $\rho$  by an unbiased estimator. Self-normalized estimators are frequent in spatial statistics, and one can expect a (small) variance reduction in (8) at the price of a small bias.

**Remark 2.** The self-normalized estimator (8) of  $I_K(f)$  is biased. Indeed

$$\mathbb{E} \left[ \widehat{I}_{s, \Pi_\varepsilon \mathcal{P}}(f) \right] = \mathbb{E} \left[ \frac{|K|}{\Pi_\varepsilon \mathcal{P}(K)} \mathbb{1}_{\{\Pi_\varepsilon \mathcal{P}(K) > 0\}} \mathbb{E} \left[ \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P} \cap K} f(\mathbf{x}) \mid \Pi_\varepsilon \mathcal{P}(K) \right] \right].$$

As  $\Pi_\varepsilon \mathcal{P}$  is a stationary point process, once conditioning on  $\Pi_\varepsilon \mathcal{P}(K)$  each point of  $\Pi_\varepsilon \mathcal{P} \cap K$  is uniformly distributed over  $K$ . Let  $(Y_i)_{i \geq 1}$  be random variables that follow the uniform distribution over  $K$ , we have

$$\begin{aligned} \mathbb{E} \left[ \widehat{I}_{s, \Pi_\varepsilon \mathcal{P}}(f) \right] &= \mathbb{E} \left[ \frac{|K|}{\Pi_\varepsilon \mathcal{P}(K)} \mathbb{1}_{\{\Pi_\varepsilon \mathcal{P}(K) > 0\}} \sum_{i=1}^{\Pi_\varepsilon \mathcal{P}(K)} \mathbb{E} [f(Y_i)] \right] \\ &= |K| \mathbb{E} \left[ \mathbb{1}_{\{\Pi_\varepsilon \mathcal{P}(K) > 0\}} \mathbb{E} [f(Y_1)] \right] \\ &= (1 - \mathbb{T}_{\Pi_\varepsilon \mathcal{P}}(K)) \int_K f(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

where the void probability  $\mathbb{T}_{\Pi_\varepsilon \mathcal{P}}$  is defined in (1). As  $K$  grows, the bias thus decreases. It is actually reasonable to expect that it vanishes exponentially fast with the size of  $K$ .

Before we investigate the variance of linear statistics under  $\Pi_\varepsilon \mathcal{P}$ , we need to ensure that the variance exists. Actually, the following result ensures that the RPPP has moments of any order.

**Proposition 2 (Existence of the moments).** *Let  $d \geq 3$  and  $\mathcal{P}$  be a homogeneous Poisson point process of intensity  $\rho > 0$  in  $\mathbb{R}^d$ . Let  $\varepsilon \in (-1, 1)$  and  $R > 0$ . For any positive integer  $m$*

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} \mathbb{1}_{B(0, R)}(\mathbf{x}) \right)^m \right] < \infty.$$

The proof is deferred to Appendix A.3, and we note that a quantitative upper bound of the expectation can be deduced from the proof. For now, a direct consequence of Proposition 2 is that for any continuous function  $f$  of compact support  $K$ , we have

$$\mathbb{V}\text{ar}\left[\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)\right] \leq \frac{\|f\|_\infty^2}{\rho^2} \mathbb{E}\left[\left(\sum_{\mathbf{x} \in \Pi_\varepsilon\mathcal{P}} \mathbf{1}_K(\mathbf{x})\right)^2\right] - I_K(f)^2 < \infty.$$

In the next section, we provide a more explicit expansion of the variance for small  $\varepsilon$ .

### 3.3. Main result

The following variance reduction result is the main theoretical finding of the present paper. Its proof is deferred to Appendix A.4.

**Theorem 3 (Variance reduction).** *Let  $d \geq 3$ ,  $\mathcal{P} \subset \mathbb{R}^d$  be a homogeneous Poisson point process of intensity  $\rho > 0$ , and  $\varepsilon \in (-1, 1)$ . For any function  $f \in C^2(\mathbb{R}^d)$  of compact support  $K$ , we have*

$$\mathbb{V}\text{ar}\left[\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)\right] = \mathbb{V}\text{ar}\left[\widehat{I}_\mathcal{P}(f)\right] (1 - 2d\kappa_d\rho\varepsilon) + O(\varepsilon^2), \quad (9)$$

where  $\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)$  is defined in (6),

$$\widehat{I}_\mathcal{P}(f) \triangleq \frac{1}{\rho} \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}), \quad (10)$$

and  $\kappa_d$  is the volume of the unit ball of  $\mathbb{R}^d$ .

Several remarks are in order.

**Remark 3.** Upon noting that  $\Pi_0\mathcal{P} = \mathcal{P}$ , Equation (9) implies a negative derivative of the variance of  $\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)$  at  $\varepsilon = 0$ . Actually,

$$\mathbb{V}\text{ar}\left[\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)\right] < \mathbb{V}\text{ar}\left[\widehat{I}_\mathcal{P}(f)\right] = \rho^{-1} I_K(f^2), \quad (11)$$

for a small enough stepsize  $\varepsilon > 0$ . Computing the second-order derivative of the variance is more challenging because the second-order moment of  $F_\mathcal{P}$  is not well-defined; see Section 3.4 and Appendix A.4.

**Remark 4.** Taking  $\varepsilon$  equal to

$$\varepsilon_0 \triangleq \frac{1}{2d\kappa_d\rho} \quad (12)$$

makes the term of order  $\varepsilon$  in (9) vanish. Note also that  $\varepsilon_0$  does not depend on the integrand  $f$ .

**Remark 5.** When  $\varepsilon < 0$ , we obtain a positive first-order derivative of the variance at  $\varepsilon = 0$ , so that, for  $|\varepsilon|$  small enough,

$$\mathbb{V}\text{ar}\left[\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f)\right] > \mathbb{V}\text{ar}\left[\widehat{I}_\mathcal{P}(f)\right].$$

This result is expected as the behavior of  $\Pi_\varepsilon$  shifts from repulsive to attractive.

**Remark 6.** A key element of the proof of Theorem 3 is the super-harmonicity of the Coulomb potential  $U_{\mathcal{P}}$ , which defines the force function  $F_{\mathcal{P}}$ . In other words, defining  $U_{\mathcal{P}}$  such that  $\nabla U_{\mathcal{P}}(\mathbf{x}) = F_{\mathcal{P}}(\mathbf{x})$ , we have

$$\Delta U_{\mathcal{P}}(\mathbf{x}) = \operatorname{div}(F_{\mathcal{P} \cap K}(\mathbf{x})) = d\kappa_d \sum_{\mathbf{z} \in \mathcal{P} \setminus \{\mathbf{x}\}} \delta_{\{\mathbf{z}\}}(\mathbf{x}) - \kappa_d \rho,$$

which is negative on  $\mathbb{R}^d \setminus \mathcal{P}$ . This property, combined with a tailored integration by parts, forms the main ingredient of the proof; see Remark 10 and Section 5.1.

**Remark 7.** Without further assumptions on the integrand, other types of interaction than Coulomb do not necessarily yield such a variance reduction if plugged into our repulsion operator. Relatedly, there are many links between Coulomb interaction and numerical integration beside our result. For instance, the so-called Fekete points, defined as maximizers of the Coulomb energy

$$\mathbf{x}_1, \dots, \mathbf{x}_N \mapsto \sum_{1 \leq i, j \leq N}^{\neq} \frac{1}{\|\mathbf{x}_i - \mathbf{x}_j\|_2^{d-2}}$$

on a compact, have been studied as a quadrature scheme, see e.g. Serfaty (2019) and references therein.

### 3.4. Properties of the force

In this section, we discuss key characteristics of the random function  $F_{\mathcal{P}}$ , when  $\mathcal{P}$  is a PPP. Additional properties can be found in the Appendices.

First, Chatterjee et al. (2010, Proposition 1) proved that when  $d \geq 3$ , almost surely, the series defining  $F_{\mathcal{P}}(\mathbf{x})$  converges simultaneously for all  $\mathbf{x}$  and defines a translation-invariant (in distribution) vector-valued random function, which is also almost surely continuously differentiable. The subsequent proposition provides further insights into the distribution of  $F_{\mathcal{P}}$ .

**Proposition 4.** *Let  $\mathcal{P}$  be a homogeneous Poisson point process of intensity  $\rho$  of  $\mathbb{R}^d$ , with  $d \geq 3$ . Then, for any  $\mathbf{x} \in \mathbb{R}^d$ ,  $F_{\mathcal{P}}(\mathbf{x})$  has a symmetric  $\alpha$ -stable distribution of index  $\alpha = \frac{d}{d-1}$ .*

This observation was mentioned by Chatterjee et al. (2010) and can be easily checked by observing that the union of  $n$  i.i.d. copies of  $\mathcal{P}$ , which is a PPP of intensity  $n\rho$ , is also a PPP of intensity  $\rho$  scaled by  $n^{-1/d}$ . The individual terms in  $F_{\mathcal{P}}(\mathbf{x})$  scale as a  $(d-1)$ -th power of the distance, so the sum of  $n$  i.i.d. copies of  $F_{\mathcal{P}}(\mathbf{x})$  has the same distribution as  $n^{(d-1)/d} F_{\mathcal{P}}(\mathbf{x})$ . Symmetry is obvious, as  $-\mathcal{P} = \mathcal{P}$  in distribution. Proposition 4 implies that  $\mathbb{E}[F_{\mathcal{P}}(\mathbf{x})] = 0$  and  $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{x})\|_2^v] < \infty$  iff  $v < \alpha$ . For more details about stable distributions, we refer to (Nolan, 2020, Section 1.5) and (Abdul-Hamid and Nolan, 1998).

Second, we have the following result regarding the distribution of the difference of forces.

**Proposition 5.** *Let  $\mathcal{P}$  be a homogeneous Poisson point process of  $\mathbb{R}^d$ . Then, for any two distinct points  $\mathbf{x}, \mathbf{y}$  of  $\mathbb{R}^d$ , the random vector  $F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y})$  is continuous, i.e., for any  $\mathbf{c} \in \mathbb{R}^d$ ,*

$$\mathbb{P}(F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}) = 0.$$

The proof of this proposition is deferred to Appendix A.1. We note that an additional result regarding the joint density of the vector  $(F_{\mathcal{P}}(\mathbf{x}), F_{\mathcal{P}}(\mathbf{y}))$  can be found in (Chatterjee et al., 2010, Theorem

10), where the authors demonstrate the existence of the joint density of  $(F_{\mathcal{P}}(\mathbf{x}), F_{\mathcal{P}}(\mathbf{y}))$  for  $\mathbf{x}$  and  $\mathbf{y}$  sufficiently far apart, further conditioning on having at least one point of  $\mathcal{P}$  within balls centered at  $\mathbf{x}$  and  $\mathbf{y}$ . They also derive an upper bound for the density.

Third, it is possible to derive an alternative expression for  $F_{\mathcal{P}}(\mathbf{x})$  that avoids the requirement of a different order of summation at each point  $\mathbf{x}$ . More precisely, ordering terms by their distance to the origin yields

$$F_{\mathcal{P}}(\mathbf{x}) = \sum_{\substack{\mathbf{z} \in \mathcal{P} \setminus \{\mathbf{x}\} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{\mathbf{x} - \mathbf{z}}{\|\mathbf{x} - \mathbf{z}\|_2^d} - \kappa_d \rho \mathbf{x}, \quad (F_2)$$

where  $\kappa_d$  is the volume of the unit ball of  $\mathbb{R}^d$ . Note the additional term in  $(F_2)$ , which compensates for fixing the order of summation. Chatterjee et al. (2010, Proposition 5) proved that the expressions  $(F_1)$  and  $(F_2)$  are equivalent when  $\mathcal{P}$  is a PPP of unit intensity. A similar proof with slight modifications holds in the general case when  $\rho \neq 1$ .

Finally, for a stationary point process  $\mathcal{X} \subset \mathbb{R}^d$ , and  $0 \leq q < p$ , we define the truncated force

$$F_{\mathcal{X}}^{(q,p)}(\mathbf{x}) \triangleq \sum_{\substack{\mathbf{z} \in \mathcal{X} \setminus \{\mathbf{x}\} \cap A^{(q,p)}(\mathbf{x}) \\ \|\mathbf{x} - \mathbf{z}\|_2 \uparrow}} \frac{\mathbf{x} - \mathbf{z}}{\|\mathbf{x} - \mathbf{z}\|_2^d}, \quad (13)$$

where  $A^{(q,p)}(\mathbf{x}) = B(\mathbf{x}, p) \setminus B(\mathbf{x}, q)$  is the annulus centered at  $\mathbf{x}$  with small radius  $q$  and big radius  $p$ . We will denote  $A^{(q,p)}(\mathbf{0})$  simply by  $A^{(q,p)}$ . Intuitively,  $F_{\mathcal{X}}^{(q,p)}(\mathbf{x})$  represents the total Coulomb force experienced by a charged particle at  $\mathbf{x}$  due to the influence of other particles of the same charge located in  $\mathcal{X} \cap A^{(q,p)}(\mathbf{x})$ . Note that the law of  $F_{\mathcal{P}}^{(q,p)}(\mathbf{x})$  is invariant under translation of  $\mathbf{x}$ , as was the case for its non-truncated counterpart. The truncated force is a useful tool for practical implementation, just like the truncated repelled point process

$$\Pi_{\varepsilon}^{(q,p)} \mathcal{X} \triangleq \{\mathbf{x} + \varepsilon F_{\mathcal{X}}^{(q,p)}(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}. \quad (14)$$

**Remark 8.** The proof of Theorem 3 holds even when we replace  $F_{\mathcal{P}}$  with its truncated version  $F_{\mathcal{P}}^{(0,p)}$ , as long as  $p$  is larger than the diameter of the support  $K$  of the integrand.

### 3.5. Sampling from the repelled Poisson point process

Let  $\mathcal{P}$  be a PPP of intensity  $\rho > 0$  and  $\Pi_{\varepsilon} \mathcal{P}$  be the associated RPPP. Let  $K \subset \mathbb{R}^d$  be compact, with diameter  $\text{diam}(K)$ . In this section, we propose two approaches to approximately sampling from  $\Pi_{\varepsilon} \mathcal{P} \cap K$ . By Corollary 1,  $\Pi_{\varepsilon} \mathcal{P}$  is stationary, and we henceforth assume that  $K \subset B(\mathbf{0}, \text{diam}(K)/2)$ .

Our first approach is simply to sample  $\Pi_{\varepsilon}^{(0,p)} \mathcal{P} \cap K$ , where  $p = \text{diam}(K)/2$ . In words, we use the points of  $\mathcal{P}$  that fall in the larger ball  $B(\mathbf{0}, \text{diam}(K))$  to displace the points of  $\mathcal{P} \cap K$ . Informally, for large  $\text{diam}(K)$ , we expect the resulting distribution to be close to that of  $\Pi_{\varepsilon} \mathcal{P} \cap K$  because, for each  $\mathbf{x} \in \mathcal{P} \cap K$ , we only neglected contributions to the force  $(F_1)$  from points at distance further than  $\text{diam}(K)/2$  from  $\mathbf{x}$ , and the magnitude of these contributions decreases fast. One downside of this approach is that it requires, for each  $\mathbf{x}$ , to order the points in  $\mathcal{P} \cap B(\mathbf{0}, p)$  by their distance to  $\mathbf{x}$ . While storing the initial sample of  $\mathcal{P} \cap B(\mathbf{0}, p)$  in an ad-hoc data structure like a KD-tree may help (Bentley, 1975), we empirically found it more computationally tractable to rely on the alternative expression  $(F_2)$  of the force.

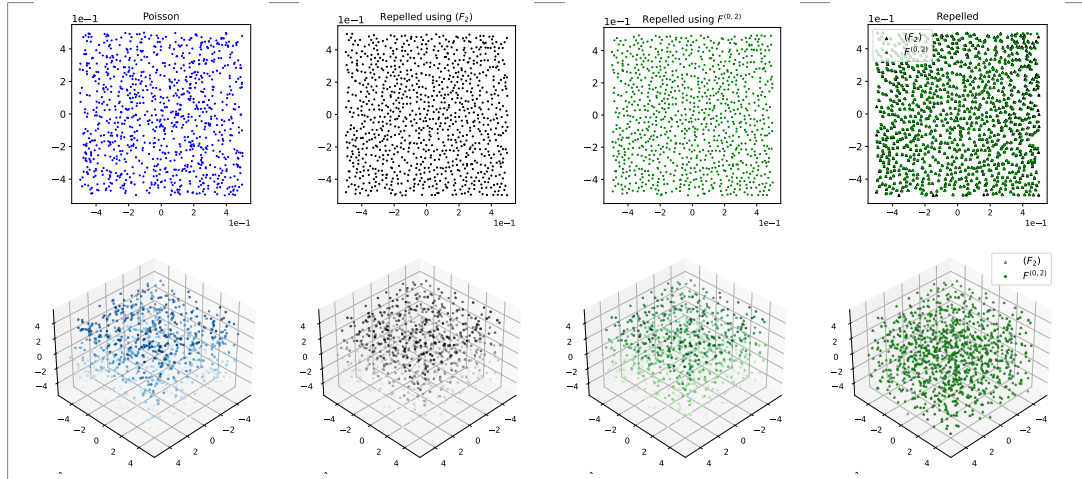


Figure 2: Poisson sample (first column) and the corresponding repelled samples, obtained using  $(F_2)$  (second column) and using  $F^{(0,2)}$  (third column). The first row corresponds to  $d = 2$  and the second row to  $d = 3$ , with  $\varepsilon$  set in each row to the value  $\varepsilon_0 = \varepsilon_0(d)$  in (12). The last column shows the two repelled samples superimposed.

Indeed, our second approach stems from the fact that the partial sums of  $(F_2)$  use the same points of  $\mathcal{P}$ , independently of  $\mathbf{x}$ . Moreover, the correction term in  $(F_2)$  partially takes into account the border effects of truncating the sum, so we propose to sample  $\mathcal{P} \cap B(\mathbf{0}, \text{diam}(K)/2)$  and use the points in the latter sample to displace the points of  $\mathcal{P} \cap K$  using  $(F_2)$ . We still need to compute the pairwise distances between the points of  $\mathcal{P} \cap B(\mathbf{0}, \text{diam}(K)/2)$ , which is  $O(N^2)$ , where  $N$  refers to the number of points of  $\mathcal{P}$  in  $K$ . However, we note that the points of  $\Pi_\varepsilon \mathcal{P} \cap K$  can be sampled concurrently, resulting in a reduction in computational time roughly proportional to the number of available processors. We provide a Python package, called [MCRPPy](https://github.com/dhawat/MCRPPy) and available on GitHub<sup>3</sup>, which implements our two simulation methods.

Figure 2 shows approximate samples of  $\Pi_{\varepsilon_0} \mathcal{P}$  in  $[-1/2, 1/2]^d$  obtained with the two aforementioned approaches, for  $d = 2$  (first row) and  $d = 3$  (second row). The corresponding PPP is of intensity 1000, and the initial samples are given in the first column. In the second column,  $(F_2)$  was used, while  $F_\mathcal{P}^{(0,2)}$  was used in the third column. The last column is a superposition of the samples obtained in columns 2 and 3, displaying very close agreement. Finally, note that Figures 1 and 3 were obtained using  $(F_2)$ , and we will keep using this simulation method in the next sections for sampling from the repelled point process.

### 3.6. An experimental illustration of the variance reduction

In this section, we present a numerical experiment to confirm the variance reduction found in Theorem 3. Additional experiments can be found in Section 4.

<sup>3</sup><https://github.com/dhawat/MCRPPy>

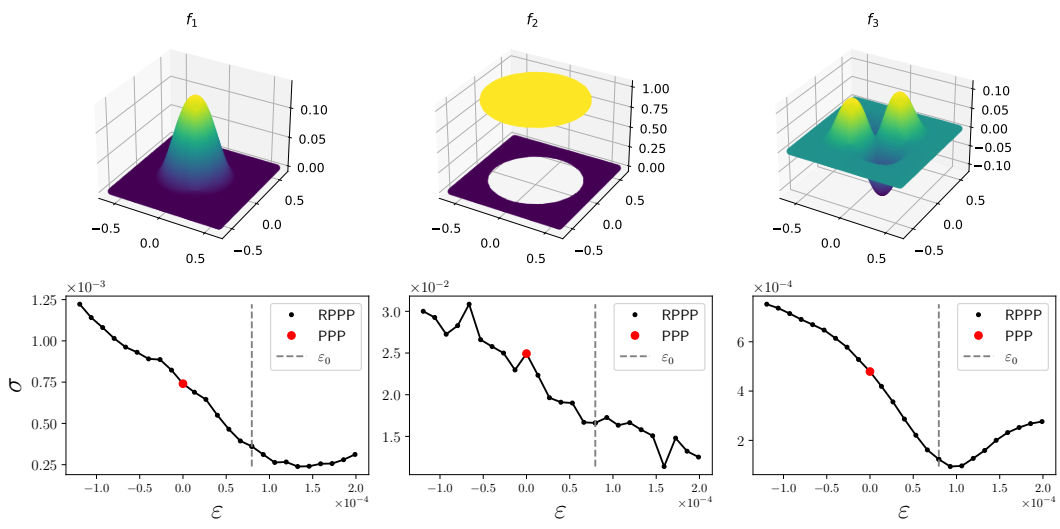


Figure 3: Estimated standard deviations of  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}$  with respect to  $\varepsilon$ , for  $f_1$ ,  $f_2$ , and  $f_3$ , in  $d = 3$ .

Let  $K = [-1/2, 1/2]^d$ . Consider the three following integrands, all supported in  $K$ ,

$$\begin{aligned}
 f_1(\mathbf{x}) &\triangleq \left(1 - 4\|\mathbf{x}\|_2^2\right)^2 \exp\left(\frac{-2}{1 - 4\|\mathbf{x}\|_2^2}\right) \mathbb{1}_{B(\mathbf{0}, 1/2)}(\mathbf{x}), \\
 f_2(\mathbf{x}) &\triangleq \mathbb{1}_{B(\mathbf{0}, 1/2)}(\mathbf{x}), \\
 \text{and } f_3(\mathbf{x}) &\triangleq \prod_{i=1}^d \cos^3(\pi x_i) \sin(\pi x_i) \mathbb{1}_K(\mathbf{x}).
 \end{aligned} \tag{15}$$

Both  $f_1$  and  $f_3$  satisfy the requirements of Theorem 3, while the indicator  $f_2$  is discontinuous on  $\partial B(\mathbf{0}, 1/2)$ . For each of these functions, Figure 3 shows the estimated standard deviations  $\widehat{\sigma}(\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}(\cdot))$  of the self-normalized estimator  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}$  (8) for varying values of  $\varepsilon$  in  $d = 3$ . We conducted the analysis using 50 independent samples of PPP  $\mathcal{P}$  of intensity  $\rho = 500$ .

The estimated standard deviations of  $\widehat{I}_{s, \mathcal{P} \cap K}$ , corresponding to  $\varepsilon = 0$ , are represented by the large red dots. The black dots indicates the values of  $\widehat{\sigma}(\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}(\cdot))$ . The dashed lines indicate  $\varepsilon_0$  (12). Note that within the range of  $\varepsilon$  employed, the number of points of  $(\Pi_\varepsilon \mathcal{P}) \cap K$  remains relatively stable, with an average ranging between 493 and 501.

First, we observe that for negative values of  $\varepsilon$ ,  $\widehat{\sigma}(\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}(\cdot))$  are greater than  $\widehat{\sigma}(\widehat{I}_{s, \mathcal{P} \cap K}(\cdot))$ , for the three functions. This behavior is expected because the operator  $\Pi_\varepsilon$  is attractive for negative values of  $\varepsilon$ . Second, for positive values of  $\varepsilon$ , up to  $\varepsilon_0$ , we observe that  $\widehat{\sigma}(\widehat{I}_{s, \Pi_\varepsilon \mathcal{P} \cap K}(\cdot))$  are lower than  $\widehat{\sigma}(\widehat{I}_{s, \mathcal{P} \cap K}(\cdot))$ . This result aligns with our theoretical expectations and provides evidence for the variance reduction in Theorem 3. Third, for  $f_1$  and  $f_3$  we observe an interesting trend when  $\varepsilon$  exceeds  $\varepsilon_0$ . The standard deviations decrease until reaching a minimum value. This minimum value, particularly for  $f_3$ , is relatively close to  $\varepsilon_0$ . However, after this minimum point, the standard deviations start to increase again. The behavior of  $f_2$  in this scenario appears to be more intricate and less predictable. Overall, it

appears that  $\varepsilon_0$  is a reasonable choice for  $\varepsilon$ , regardless of the integrand, although not necessarily the optimal threshold for a specific integrand.

## 4. Application to numerical integration

In this section, we benchmark the RPPP among a few key Monte Carlo methods, to provide context. The Python code for replicating this study can be found in [MCRPPy](https://github.com/dhawat/MCRPPy)<sup>4</sup>.

### 4.1. A few Monte Carlo methods

Let  $f$  be a continuous function supported in  $K = [-1/2, 1/2]^d$ . Our goal is to estimate the integral  $I_K(f)$  in (7). The simple (or crude) Monte Carlo method employs a Binomial point process (BPP)  $\mathcal{B}_N$  supported on  $K$ , see Section 2.2, to estimate  $I_K(f)$  as follows

$$\widehat{I}_{\text{MC}}(f) = \frac{1}{N} \sum_{\mathbf{x} \in \mathcal{B}_N} f(\mathbf{x}). \quad (16)$$

The number of points used is fixed to  $N$ .  $\widehat{I}_{\text{MC}}(f)$  is an unbiased estimator of  $I_K(f)$  with a variance equal to  $N^{-1} \mathbb{V}\text{ar}(f(\mathbf{u}))$  where  $\mathbf{u}$  is a uniformly drawn point of  $K$ .

**Remark 9.** As mentioned in Section 2.2, when the number of points of a BPP increases appropriately with the size of the observation window, the BPP converges to a homogeneous Poisson point process (PPP). Consequently, one can think of the estimator  $\widehat{I}_{\text{MC}}(f)$  as a self-normalized version of the estimator  $\widehat{I}_{\mathcal{P}}(f)$  (10), where  $\mathcal{P}$  represents a PPP with suitable intensity. However, by setting the intensity of  $\mathcal{P}$  equal to  $N$ , we obtain the following variance for  $\widehat{I}_{\mathcal{P}}(f)$

$$\mathbb{V}\text{ar} \left[ \widehat{I}_{\mathcal{P}}(f) \right] = N^{-1} \int_K f^2(\mathbf{x}) d\mathbf{x},$$

which is larger than the variance of  $\widehat{I}_{\text{MC}}(f)$

$$\mathbb{V}\text{ar} \left[ \widehat{I}_{\text{MC}}(f) \right] = N^{-1} \left( \int_K f^2(\mathbf{x}) d\mathbf{x} - \left( \int_K f(\mathbf{x}) d\mathbf{x} \right)^2 \right).$$

These variances are only equal if the integral of  $f$  is zero. Otherwise, fixing the number of points is preferable to using a random number of points.

Much research has gone into reducing the variance of  $\widehat{I}_{\text{MC}}$ ; see e.g. (Owen, 2013). Control variate methods, for instance, rely on incorporating a function  $h$  that is computationally cheap to evaluate and possesses a known integral. The Monte Carlo control variate estimator is defined as

$$\widehat{I}_{\text{MCCV}}(f) = \widehat{I}_{\text{MC}}(f) - c(\widehat{I}_{\text{MC}}(h) - I_K(h)), \quad (17)$$

<sup>4</sup><https://github.com/dhawat/MCRPPy>

where  $c$  is a free parameter to be tuned. First,  $\widehat{I}_{\text{MCCV}}(f)$  is an unbiased estimator of  $I_K(f)$ . Second, the value of  $c$  that minimizes the variance of the estimator is

$$\frac{\text{Cov}(f(\mathbf{u}), h(\mathbf{u}))}{\text{Var}[h(\mathbf{u})]},$$

where  $\mathbf{u}$  is a uniform random vector on  $K$  (Owen, 2013, Section 8.9). For this specific value of  $c$ ,

$$\text{Var}[\widehat{I}_{\text{MCCV}}(f)] = \text{Var}[\widehat{I}_{\text{MC}}(f)] \left(1 - \text{Corr}^2(f(\mathbf{u}), h(\mathbf{u}))\right).$$

Hence, any function  $h$  that correlates with  $f$  helps reducing the variance, even one that correlates negatively. Finding an appropriate  $h$  is usually challenging. We opt in this paper for a polynomial regression of  $f$  as our choice for  $h$ , and we estimate the coefficient  $c$  by

$$\hat{c} = \frac{\sum_{\mathbf{x} \in \mathcal{B}'} f(\mathbf{x})(h(\mathbf{x}) - \bar{h})}{\sum_{\mathbf{x} \in \mathcal{B}'} (h(\mathbf{x}) - \bar{h})^2}, \quad (18)$$

where  $\mathcal{B}'$  is a BPP on  $K$ , which is independent of the point process used to compute  $\widehat{I}_{\text{MC}}$  in (17), and  $\bar{h}$  is the mean of  $h$  over  $\mathcal{B}'$ . Under the independence assumption, using  $\hat{c}$  instead of  $c$  does not affect the unbiasedness of  $\widehat{I}_{\text{MCCV}}(f)$  (Owen, 2013, Example 8.3). Finally, note that the construction of the regressor  $h$  and the evaluation of  $\hat{c}$  require extra integrand evaluations compared to crude Monte Carlo, which we neglect here for simplicity.

On the other hand, Monte Carlo methods that try to improve the rate of convergence of the variance have been proposed, starting with grid-based stratification (Owen, 2013, Chapter 10). For a recent example, replacing the BPP of crude Monte Carlo with one of a particular family of determinantal point processes (DPPs) has been shown to enhance the convergence rate of the variance beyond  $O(N^{-1})$ ; see (Bardenet and Hardy, 2020, Coeurjolly et al., 2021). For our comparison, we use the unbiased estimator given in (Bardenet and Hardy, 2020, Section 2.1) and implemented in DPPy (Gautier et al., 2019b),

$$\widehat{I}_{\text{MCDPP}}(f) = \sum_{\mathbf{x} \in \mathcal{D}_N} \frac{f(\mathbf{x}/2)}{2^d \kappa_N(\mathbf{x}, \mathbf{x})}, \quad (19)$$

where  $\mathcal{D}_N$  is the multivariate Legendre ensemble and  $\kappa_N$  its kernel. According to Bardenet and Hardy (2020), if  $f$  is  $C^1$  and supported on an open set that is bounded away from the boundary of the hypercube  $K$ ,  $\widehat{I}_{\text{MCDPP}}(f)$  is an unbiased estimator of  $I_K(f)$  and its variance is  $O(N^{-1-1/d})$ , which is faster than the usual  $O(N^{-1})$ . One of the limitations of DPP-based methods lies in the computational complexity associated with sampling from DPPs, which is at least cubic in the number  $N$  of integrand evaluations.

We conclude with an instance of the Randomized Quasi-Monte Carlo method, which is an attempt at getting both the convenience of variance statements and the error reduction of stratified deterministic quadrature. The estimator used is

$$\widehat{I}_{\text{RQMC}}(f) = \frac{1}{N} \sum_{\mathbf{x} \in \mathcal{S}_N} f(\mathbf{x}), \quad (20)$$

where  $\mathcal{S}_N$  is obtained by applying a suitable random perturbation to a low-discrepancy (deterministic) sequence (Owen, 2013, Chapter 17). In particular, each point of  $\mathcal{S}_N$  is uniformly distributed in  $K$ , so that (20) is an unbiased estimator of  $I_K(f)$ . Under strong regularity assumptions on  $f$  (at least all mixed partial derivatives of  $f$  of order less than  $d$  should be continuous on  $K$ ), the variance of

(20) is  $O(\log(N)^{d-1}N^{-3})$  (Owen, 2013, Theorem 17.5). Despite the  $d$ -dependence of the rate and the strong smoothness assumptions,  $\widehat{I}_{\text{RQMC}}$  is remarkably efficient in small-to-moderate dimensions, and sampling is computationally cheap compared to, e.g., DPPs. In this paper, we take  $\mathcal{S}_N$  to be a scrambled Sobol sequence<sup>5</sup> (Sobol, 1967).

## 4.2. Experimental comparison

This section focuses on examining and comparing the performance of  $\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f)$ , for  $\varepsilon = \varepsilon_0$  (12), with the Monte Carlo estimators outlined in Section 4.1.

While  $\widehat{I}_{\text{MC}}$ ,  $\widehat{I}_{\text{MCCV}}$ ,  $\widehat{I}_{\text{MCDPP}}$ , and  $\widehat{I}_{\text{RQMC}}$  use a constant number of points across trials,  $\widehat{I}_{\Pi_{\varepsilon_0} \mathcal{P}}$  is the only method where the number of points is not fixed, neither in the computational budget nor in the (smaller) number of integrand evaluations. In an effort to conduct a fair comparison, we replace the PPP in the estimator  $\widehat{I}_{\Pi_{\varepsilon_0} \mathcal{P}}$  with a BPP, which has a fixed number of points. The resulting estimator is referred to as  $\widehat{I}_{\text{MCRB}}$ . Next, we sample  $M$  realizations from the repelled point processes as described in Section 3.5, and find the average number of points obtained within the  $M$  trials. We set the number  $N$  of points used in the other methods to this average. Note that we give an unfair advantage to the estimator  $\widehat{I}_{\text{MCCV}}$  by not accounting for the evaluations of the integrand necessary to estimate the coefficient (18).

We use  $M = 100$  samples from each of the point processes and we make sure  $N$  ranges roughly from 50 to 1000. We examine the functions  $f_1$ ,  $f_2$ , and  $f_3$  defined in Section 3 by Equation (15). Their integrals are

$$I_K(f_2) = \kappa_d/2^d, \quad I_K(f_3) = 0,$$

while the precise value of  $I_K(f_1)$  is unknown.

Figure 4 displays the log of the estimated standard deviation of the estimators  $\widehat{I}_{\text{MC}}$ ,  $\widehat{I}_{\text{MCRB}}$ ,  $\widehat{I}_{\text{MCCV}}$ , and  $\widehat{I}_{\text{RQMC}}$ , plotted against the log number of points  $\log(N)$ , for the functions  $f_1$ ,  $f_2$ , and  $f_3$  and  $d$  in  $\{2, 3, 4, 5, 7\}$ .  $\widehat{I}_{\text{MCDPP}}$  is only examined for  $d \in \{2, 3\}$  due to its high cost compared to other methods. Lines correspond to ordinary least square linear regressions (OLS). The slopes and standard deviations of the slopes are indicated in the legend.

First, as expected, the estimated variances of  $\widehat{I}_{\text{MCRB}}$  are generally lower than those of  $\widehat{I}_{\text{MC}}$ .  $\widehat{I}_{\text{MCRB}}$  outperforms  $\widehat{I}_{\text{MC}}$  in most scenarios, except for  $f_4$  with  $d = 7$  where the estimated variances of  $\widehat{I}_{\text{MC}}$  and  $\widehat{I}_{\text{MCRB}}$  are comparable. Interestingly, in this case, the estimated variances of all methods are in the same ballpark. Second,  $\widehat{I}_{\text{MCRB}}$  outperforms  $\widehat{I}_{\text{MCDPP}}$  in dimension 3. Although the convergence rate for the variance of  $\widehat{I}_{\text{MCDPP}}$  is faster than  $O(N^{-1})$ , it seems that the numerator in the variance increases with the dimension. Moreover,  $\widehat{I}_{\text{MCDPP}}$  is computationally demanding. Fourth, it appears that  $\widehat{I}_{\text{MCRB}}$  outperforms  $\widehat{I}_{\text{MCCV}}$  in most cases.  $\widehat{I}_{\text{MCRB}}$  and  $\widehat{I}_{\text{RQMC}}$  seem the main competitors.

$\widehat{I}_{\text{RQMC}}$  consistently performs well. For  $d \leq 3$ , the estimated variances of  $\widehat{I}_{\text{RQMC}}$  are lower than those of  $\widehat{I}_{\text{MCRB}}$  for  $N$  large enough, and the slope of the variance of  $\widehat{I}_{\text{RQMC}}$  is steeper than that of  $\widehat{I}_{\text{MCRB}}$ . However,  $\widehat{I}_{\text{RQMC}}$ 's performance decreases significantly as  $d$  increases. Interestingly, for  $f_3$  and  $d = 4$ , the estimated variances of  $\widehat{I}_{\text{RQMC}}$  are larger than those of  $\widehat{I}_{\text{MCRB}}$ , but the slope for  $\widehat{I}_{\text{RQMC}}$  remains steeper than for  $\widehat{I}_{\text{MCRB}}$ , letting  $\widehat{I}_{\text{RQMC}}$  catch up as  $N$  grows large. The same trend is observed for  $f_1$  in  $d = 7$  up to a large value of  $N$ , while the opposite trend is observed for  $f_4$  in  $d = 5$ . When comparing the estimated variances of  $\widehat{I}_{\text{RQMC}}$  and  $\widehat{I}_{\text{MCRB}}$  for  $f_1$  and  $f_2$  across different dimensions, it appears that  $\widehat{I}_{\text{RQMC}}$ 's performance declines more rapidly than  $\widehat{I}_{\text{MCRB}}$ 's as the dimension  $d$  increases.

<sup>5</sup>We used the method `Sobol` from the Python package `scipy` to sample from  $\mathcal{S}_N$  (Virtanen et al., 2020).

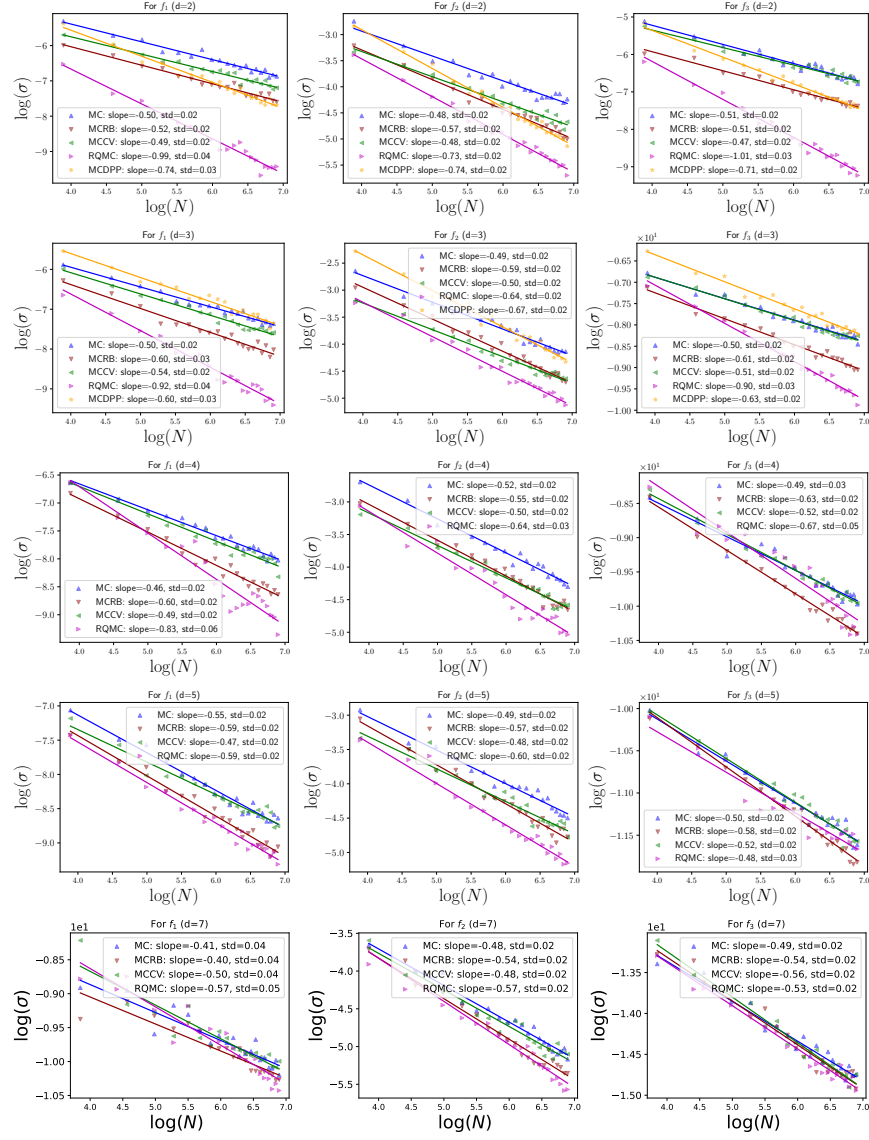


Figure 4: Estimated standard deviations of various Monte Carlo methods for  $f_1$ ,  $f_2$ , and  $f_3$  across different dimensions  $d \in \{2, 3, 4, 5, 7\}$ .

Finally, we conducted additional analyses in Appendix B.1, and computed confidence intervals for the slopes corresponding to  $\hat{I}_{\text{MCRB}}$  illustrated in Figure 4. We also compared the error of the Monte Carlo estimators under investigation. In short, data suggests that the slope is the usual Monte Carlo rate or slightly faster. Additionally, there is no substantial evidence of bias in  $\hat{I}_{\text{MCRB}}$ .

## 5. Other models and properties

This section contains numerical investigations of several intriguing questions that arise from the repulsion operator. In Section 5.1, we analyze the behavior of a PPP when the repulsion operator is applied several times. Then, in Section 5.2, we estimate the second-order properties of the RPPP. Finally, in Section 5.3, we explore the behavior of the repulsion operator when applied to two point processes that are more regular than the Poisson point process.

### 5.1. Iterating the repulsion

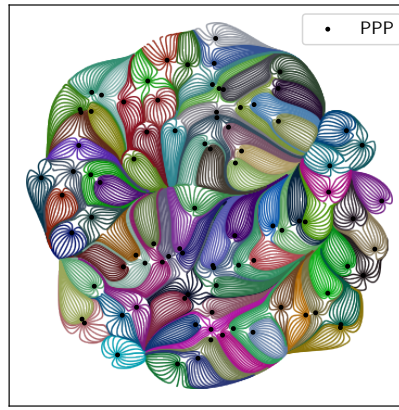


Figure 5: Illustration of the gravitational allocation from Lebesgue to a realization (black points) of a PPP in a disk. Each set of curves sharing the same color illustrates a gravitational cell, which indicates the points of the space allocated to the point of the PPP that belongs to that particular colored region. The code to generate this picture can be found in [MCRPPy](#). Note that the image is purely for illustrative purposes, and no claims are being made regarding the existence of a gravitational allocation from the Lebesgue measure to a PPP in dimension  $d = 2$ .

Our repulsion operator  $\Pi_\varepsilon$  in (5) can be seen as performing one step of a discretization scheme for a system of differential equations describing gravitational allocation. In this section, we carry the analogy with gravitational allocation further by iterating the application of  $\Pi_\varepsilon$ .

To provide more context, consider a PPP  $\mathcal{P}$  of unit intensity, let  $\mathbf{x} \in \mathbb{R}^d \setminus \mathcal{P}$ , and consider the differential equation

$$\frac{\partial}{\partial t} Y_{\mathbf{x}}(t) = -F_{\mathcal{P}}(Y_{\mathbf{x}}(t)) \quad \text{with} \quad Y_{\mathbf{x}}(0) = \mathbf{x}. \quad (21)$$

The solution  $t \mapsto Y_{\mathbf{x}}(t)$  of the differential equation (21), defined up to some positive time  $\tau_{\mathbf{x}} \in (0, \infty]$ , is called a flow curve of the gravitational allocation of  $\mathcal{P}$  to the Lebesgue measure ([Chatterjee et al., 2010](#)); see Figure 5 for an illustration of the gravitational allocation of a PPP of unit intensity. Remark that for  $\varepsilon < 0$ ,  $\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})$  is the first step in a naive numerical scheme discretizing the differential equation (21), with a stepsize equal to  $-\varepsilon$ . Similarly, each point of  $\Pi_\varepsilon \mathcal{P}$  can be viewed as the initial discretization step of a differential equation akin to (21), with stepsize  $-\varepsilon$ . The catch is that  $\Pi_\varepsilon \mathcal{P}$  is applied to the

points of  $\mathcal{P}$  itself, not to points in  $\mathbb{R}^d \setminus \mathcal{P}$  as in (21). Loosely speaking, the image of  $\mathbf{x} \in \mathcal{P}$  in  $\Pi_\varepsilon \mathcal{P}$  can be seen as a first step of the numerical discretization of a flow curve, in a gravitational allocation from the reduced Palm measure of  $\mathcal{P}$  to the Lebesgue measure. However, to fully formalize and understand this allocation, further technical details are required and are out of the scope of this paper.

Consider now performing  $M$  steps of the same numerical scheme, i.e., for each  $\mathbf{x}_0 \in \mathcal{P}$ , define

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \varepsilon F_{\mathcal{P}}(\mathbf{x}_{t-1}), \quad t = 1, \dots, M. \quad (22)$$

Call  $\Pi_{\varepsilon,t} \mathcal{P}$  the set of  $t$ -th iterates (22) of the points of  $\mathcal{P}$ . We expect that for  $\varepsilon < 0$  and large  $t$ ,

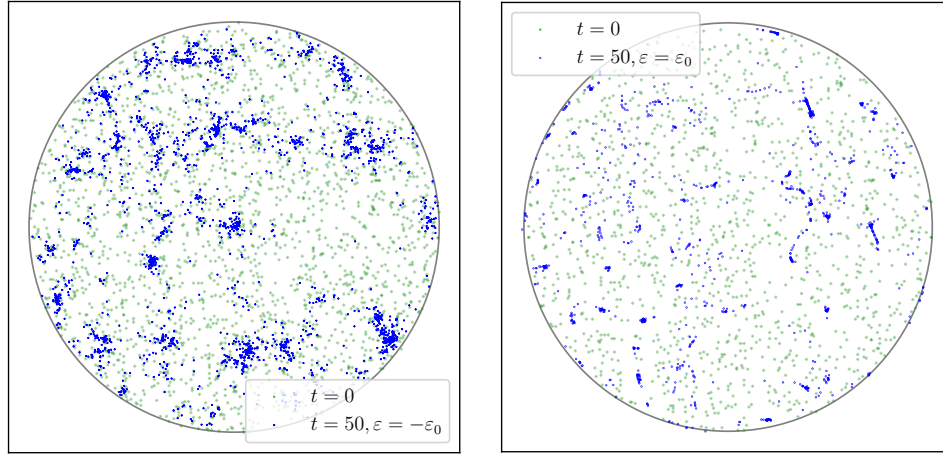


Figure 6: The green points represent a sample from a PPP  $\mathcal{P}$  of unit intensity, the blue points correspond to  $\Pi_{-\varepsilon_0,50} \mathcal{P}$  (left), and  $\Pi_{+\varepsilon_0,50} \mathcal{P}$  (right).

the distribution of the points of  $\Pi_{\varepsilon,t} \mathcal{P}$  will not be more regular than  $\mathcal{P}$ . This assertion is supported by the clustered arrangement of the points of  $\Pi_{-\varepsilon_0,50} \mathcal{P}$  observed in the left panel of Figure 6; see also Figure 13. However, it is important to note that some points of  $\Pi_{-\varepsilon_0,50}$  are situated outside the observation window and are thus not visible in Figure 6.

Following the arguments of Chatterjee et al. (2010), one can prove that the differential equation (21) defines an allocation rule. So, in particular, for almost any  $\mathbf{x} \in \mathbb{R}^d \setminus \mathcal{P}$ , each curve  $Y_{\mathbf{x}}(t)$  will eventually end at a point of  $\mathcal{P}$ , almost surely, as  $t$  goes to  $\tau_{\mathbf{x}}$ . Similarly, in an “ideal” discretization scenario, we would expect that  $\mathbf{x}_t$  ends at a point of  $\mathcal{P}$  as  $t \rightarrow \tau_{\mathbf{x}_0}$ . However, during our experiments, we observed that certain points have moved away from the observation window and the remaining points clustered together within it. The movement of certain points away from the observation window can be attributed to the naive discretization scheme (22), where we employed a fixed stepsize  $\varepsilon$ , while the points of  $\mathcal{P}$  are singular points of  $F_{\mathcal{P}}$ . As a result, when an  $\mathbf{x}_t$  is in close proximity to a point in  $\mathcal{P}$ , the force acting on it becomes considerable, which in turn compels  $\mathbf{x}_{t+1}$  to escape the observation window. Applying a truncated version of the force might help prevent this phenomenon.

The scenario where  $\varepsilon > 0$  and  $M$  is large is less straightforward. This case can be associated with the reverse dynamics of Equation (21). Indeed, for  $\mathbf{x} \in \mathbb{R}^d \setminus \mathcal{P}$ , the trajectory of  $Y_{\mathbf{x}}(t)$  halts at some point of the boundary of an allocation cell of the gravitational allocation from Lebesgue to  $\mathcal{P}$ ; see Figure 5. By the same analogy as before, we can expect that for  $\mathbf{x}_0 \in \mathcal{P}$  and  $\varepsilon$  small enough, as  $t \rightarrow \infty$ , each

point  $\mathbf{x}_i$  will eventually approach the boundary of an allocation cell of the gravitational allocation from Lebesgue to the reduced Palm measure of  $\mathcal{P}$ . It is hard to dig deeper without further investigation. Interestingly, in the right panel of Figure 6, we observe a peculiar clustering behavior of the points in  $\Pi_{\varepsilon_0, 50}\mathcal{P}$  characterized by points appearing to overlap or superimpose with each other; see also Figure 13.

## 5.2. Second-order properties of the repelled Poisson point process

The second-order characteristics of a point process, such as the pair correlation function and the structure factor, offer valuable insights into the regularity of the point process (Baddeley et al., 2015), variance reduction (Pilleboue et al., 2015), and hyperuniformity (Coste, 2021, Torquato, 2018). In this section, we estimate the pair correlation function and the structure factor of the RPPP.

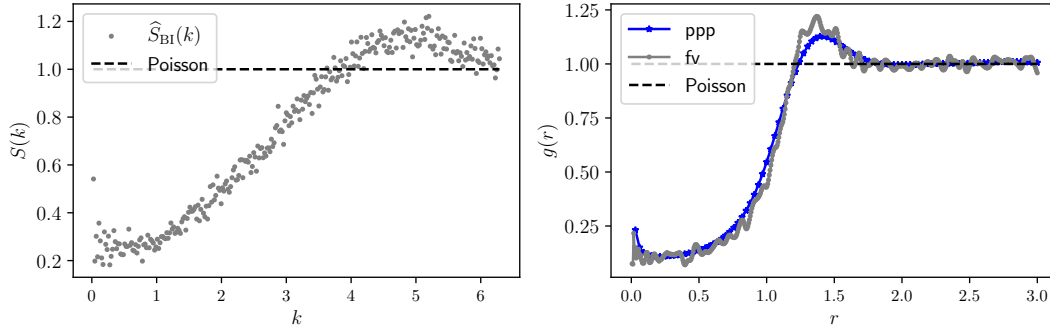


Figure 7: Estimated structure factor (left) and pair correlation function (right) of an RPPP of  $\mathbb{R}^2$  of intensity  $1/\pi$ . The labels ‘pdp’ and ‘fv’ correspond to two different estimators of the pair correlation function; see e.g. (Hawat et al., 2023).

The pair correlation function  $g$  of a stationary point process  $X$  of intensity  $\rho > 0$  is a function characterizing the probability of finding a pair of points of the point process at a certain distance. It is important to note that not all point processes have a pair correlation function in the strict sense, as it can sometimes be a measure instead. This is the case for a perturbed lattice, for example. However, if  $g$  exists, it is given by

$$\mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in X}^{\neq} h(\mathbf{x}, \mathbf{y}) \right] \triangleq \int_{\mathbb{R}^d \times \mathbb{R}^d} h(\mathbf{x} + \mathbf{y}, \mathbf{y}) \rho^2 g(\mathbf{x}) d\mathbf{x} d\mathbf{y},$$

for any nonnegative measurable bounded function  $h$  with compact support. When  $g - 1$  is integrable we can further define the structure factor  $S$  of  $X$  by

$$S(\mathbf{k}) = 1 + \rho \mathcal{F}(g - 1)(\mathbf{k}), \quad (23)$$

where  $\mathcal{F}$  denotes the Fourier transform. Assuming that  $\mathcal{X}$  is also isotropic,<sup>6</sup> both  $g$  and  $S$  become radial functions, meaning that  $g(\mathbf{r})$  and  $S(\mathbf{k})$  depend only on the norm of  $\mathbf{r}$  and  $\mathbf{k}$  respectively. In that case, we abusively write  $g(\mathbf{r}) = g(r)$ , and  $S(\mathbf{k}) = S(k)$  where  $x \triangleq \|\mathbf{x}\|_2$ . For a PPP, the structure factor and the pair correlation function are both constant and equal to 1. A pair correlation function  $g(r)$  below 1 indicates repulsion between the points of the corresponding point process at scale  $r$ , whereas a pair correlation function greater than 1 is a sign of attraction. The structure factor, on the other hand, is used as an indicator of hyperuniformity:  $S(\mathbf{0}) = 0$  implies that the point process is hyperuniform, that is, the variance of the number of points that fall in a ball grows slower than the volume of that ball. For the RPPP, with  $\varepsilon = \varepsilon_0$  we expect to observe  $g(r)$  less than 1 for a certain range of  $r$ , as the repulsion operator introduces repulsion between the points. However, having intuition on hyperuniformity is more difficult, and a statistical diagnostic using an estimator of  $S$  is required.

Figure 7 shows the estimated pair correlation function and structure factor of an RPPP of intensity  $1/\pi$  in dimension  $d = 2$ . The estimations were obtained using a sample observed within a centered ball of radius  $r = 150$ . Bartlett's isotropic estimator  $S_{\text{BI}}$  was used to estimate the structure factor on the corresponding set of allowed wavenumbers  $k$ ; see (Hawat et al., 2023, Sections 3.2.1 and 5.3). The estimators `pcf.ppp`, and `pcf.fv` (Baddeley et al., 2015, Sections 7.4.4 and 7.4.5) of the R library `spatstat` were used to estimate the pair correlation function through the Python package `structure-factor`; see (Hawat et al., 2023, Section 5.3). Dashed black lines represent the structure factor and pair correlation function of the PPP. As expected,  $g$  is smaller than 1 up to a certain value of  $r$ , indicating repulsion between the points at small scales. However,  $S_{\text{BI}}(k)$  is greater than 0.2 at small  $k$ , suggesting that the RPPP may not be hyperuniform.

### 5.3. Other repelled point processes

In this section, we perform numerical investigations to determine whether the variance reduction identified in Theorem 3 remains valid when the initial points are more evenly distributed than a Poisson point process, such as in the case of the Ginibre ensemble or a scrambled Sobol sequence.

The Ginibre ensemble (GPP) is a motion-invariant point process of  $\mathbb{R}^2$  of intensity  $1/\pi$ . It can be defined (and approximately sampled) as the limit of the set of eigenvalues of matrices filled with i.i.d. standard complex Gaussian entries, as the size of the matrix goes to infinity (Hough et al., 2009, Theorem 4.3.10). Interestingly, the GPP is a hyperuniform point process, and the variance of the number of points in  $B(\mathbf{0}, R)$  scales like  $O(R^{d-1})$  as  $R$  goes to infinity, instead of scaling like the volume of the ball, as for the PPP. By Beck (1987), this is the smallest possible growth rate. Figure 8 displays a sample of a GPP and the corresponding repelled sample with  $\varepsilon = \varepsilon_0$ , observed within  $B(\mathbf{0}, 50)$ . The sampling methodology follows the description provided in Section 3.5 for the PPP. Another model of interest is the scrambled Sobol sequence, which is a typical example of a randomized low-discrepancy sequence, as mentioned in Section 4.1. Figure 9 displays a sample of the scrambled Sobol sequence and the corresponding repelled sample, with  $\varepsilon = \varepsilon_0$ , observed within  $B(\mathbf{0}, 50)$ . Like for Ginibre, the repelled Sobol sequence displays a high level of homogeneity.

We now examine the behavior of the variance of  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{G} \cap K}$  and  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{S} \cap K}$  w.r.t.  $\varepsilon$ , where  $\mathcal{G}$  represents a GPP and  $\mathcal{S}$  denotes a scrambled Sobol sequence. We use the functions  $f_1$ ,  $f_2$  and  $f_3$  defined in Equation (15). Figure 10 illustrates the estimated standard deviations of  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{G} \cap K}$  (first row) for  $d = 2$  and  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{S} \cap K}$  for  $d = 2$  (second row), as well as for  $d = 3$  (last row), for various values of  $\varepsilon$ . We conducted the experiments using 50 independent samples of  $\mathcal{G}$  and  $\mathcal{S}$  of intensity  $\rho = 500$ . GPP's

<sup>6</sup>The assumptions of stationarity and isotropy, in this case, can be straightforwardly weakened to assuming that the intensity measure is invariant to translations or that the pair correlation function only depends on the inter-point distance.

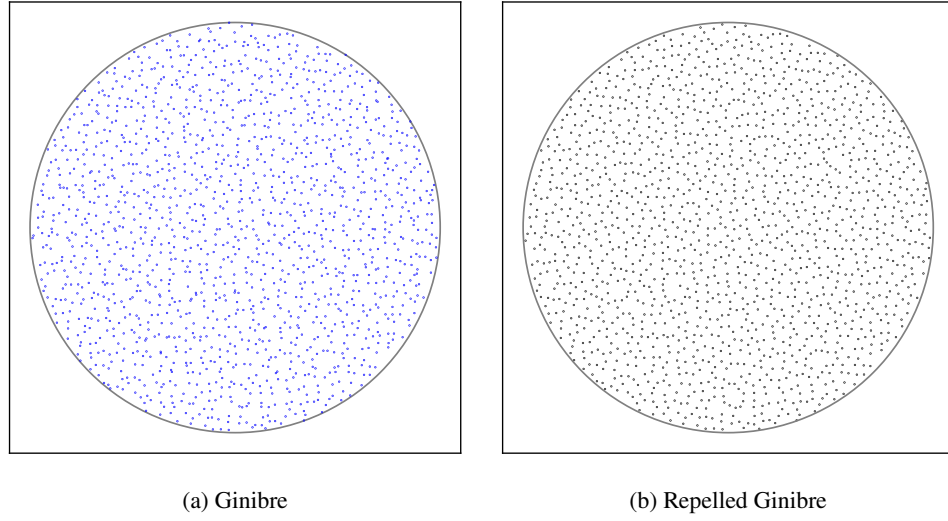


Figure 8: A sample from the Ginibre ensemble (left) and the obtained repelled sample (right) with  $\varepsilon = \varepsilon_0$ .

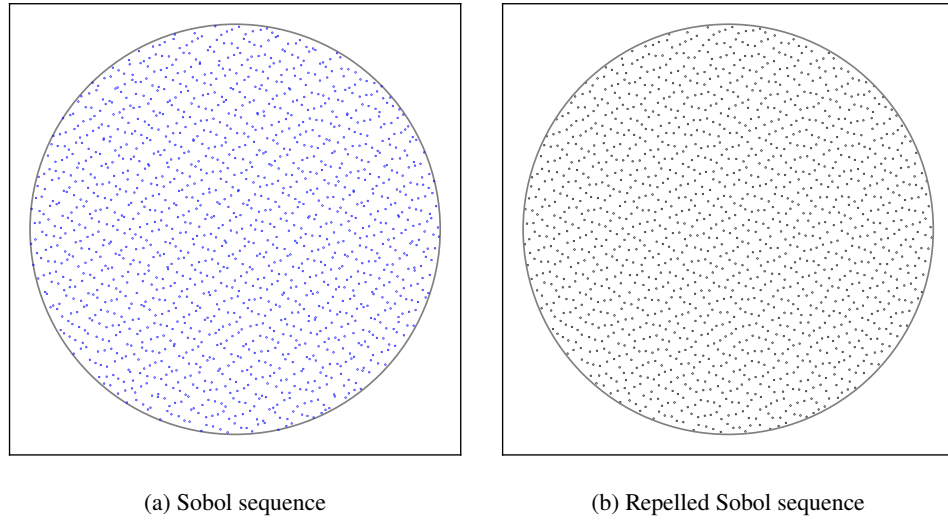


Figure 9: A sample from the scrambled Sobol sequence (left) and the obtained repelled sample (right) with  $\varepsilon = \varepsilon_0$ .

samples were rescaled to achieve  $\rho = 500$ . The estimated standard deviations of  $\widehat{I}_{s, \mathcal{G} \cap K}$  and  $\widehat{I}_{s, \mathcal{S} \cap K}$  are indicated by the large red dots in Figure 3, while the black dots correspond to the estimated standard deviations of  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{G} \cap K}$  and  $\widehat{I}_{s, \Pi_\varepsilon \mathcal{S} \cap K}$ . The dashed lines indicate the value of  $\varepsilon_0$  defined in Equation (12). For the  $C^2$  functions  $f_1$  and  $f_3$ , we observe a behavior similar to the RPPP results depicted in

Figure 3, indicating a variance reduction within a certain range of positive values of  $\varepsilon$ . However, for  $f_2$ , the variance decreases as  $\varepsilon$  increases for  $\Pi_\varepsilon \mathcal{G}$  in a manner similar to Figure 3, while a more intricate behavior is observed for  $\Pi_\varepsilon \mathcal{S}$ . These observations allow one to conjecture that the repulsion operator may produce variance reduction for smooth functions for a wide range of point processes.

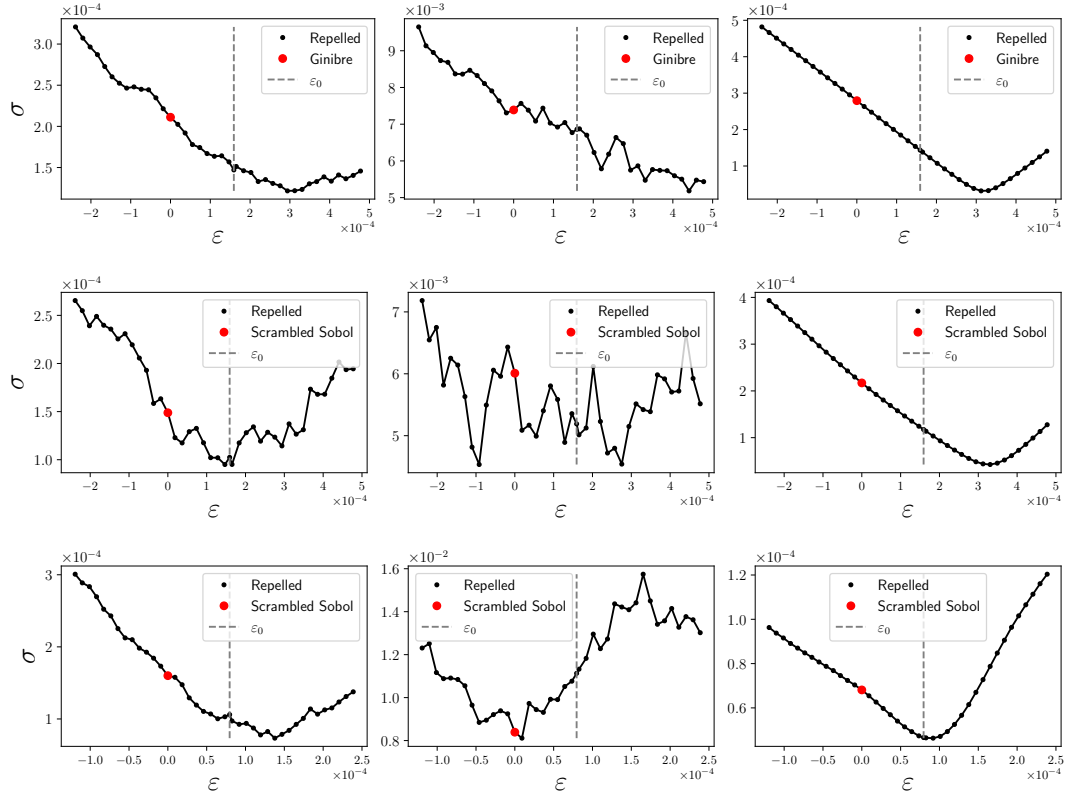


Figure 10: Estimated standard deviations of  $\hat{I}_{s, \Pi_\varepsilon \mathcal{G} \cap K}$  (Ginibre) and  $\hat{I}_{s, \Pi_\varepsilon \mathcal{S} \cap K}$  (scrambled Sobol) with respect to  $\varepsilon$  for  $f_1$  (first column),  $f_2$ , (second column) and  $f_3$  (last column). The first row shows the obtained results for the Ginibre ensemble in  $d = 2$ , the second row for the scrambled Sobol sequence in  $d = 2$  and the last one for the scrambled Sobol sequence in  $d = 3$ .

## 6. Conclusion

Motivated by variance reduction for Monte Carlo methods and inspired by gravitational allocation, we have introduced the repulsion operator  $\Pi_\varepsilon$ . For small  $\varepsilon > 0$ , this operator intuitively makes point processes more regular by slightly pushing their points apart from each other using a force function controlled by  $\varepsilon$ . We have provided a detailed theoretical study of  $\Pi_\varepsilon \mathcal{P}$ , where the repulsion operator is applied to a homogeneous Poisson point process  $\mathcal{P}$ . In particular, we have proved variance reduction for smooth linear statistics and suggested a practical value for the parameter  $\varepsilon$ . Numerical experiments

support the variance reduction and make the repelled Poisson point process a promising alternative to crude Monte Carlo integration if one can afford the quadratic cost of computing pairwise distances. We have found no numerical evidence of hyperuniformity for the repelled Poisson point process through the estimation of its structure factor, while the variance decay in experiments on smooth functions is compatible with a slightly faster than crude Monte Carlo rate.

Exploratory experiments suggest that variance reduction is also achieved when applying the repulsion operator to other point processes, such as the (hyperuniform) Ginibre point process. Proving this is a natural next step for future work, as well as proving variance reduction when applying the repulsion operator to non-homogeneous point processes. Once such a result will have been obtained under sufficiently weak assumptions, applying the repulsion operator could become a default postprocessing in many Monte Carlo integration tasks. In another direction, it would be interesting to explore the attractive case of the operator  $\Pi_\varepsilon$ , achieved by selecting a negative parameter  $\varepsilon$ . This approach yields clustering point processes, where points tend to aggregate near the points of the original point process, offering insights into spatial clustering phenomena.

## Acknowledgements

We thank David Dereudre for valuable discussions along this project and Manjunath Krishnapur for helpful insights about gravitational allocations. This work is supported by ERC-2019-STG-851866 and ANR-20-CHIA-0002.

## References

- H. Abdul-Hamid and J. P. Nolan. Multivariate stable densities as functions of one dimensional projections. *Journal of Multivariate Analysis*, 67(1):80–89, 1998. ISSN 0047-259X.
- S. Axler, P. Bourdon, and W. Ramey. *Harmonic Function Theory*. Springer New York, NY, 2 edition, 2001. ISBN 9780511801334.
- R. Azaïs, B. Delyon, and F. Portier. Integral estimation based on Markovian design. *Advances in Applied Probability*, 50(3):833–857, 2018. ISSN 00018678.
- A. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns : Methodology and Applications with R*. Chapman and Hall/CRC, nov 2015. ISBN 9780429161704.
- R. Bardenet and A. Hardy. Monte Carlo with determinantal point processes. *The Annals of Applied Probability*, 30(1), feb 2020. ISSN 1050-5164.
- J. Beck. Irregularities of distribution. I. *Acta Mathematica*, 159(0):1–49, 1987. ISSN 0001-5962.
- A. Belhadji, R. Bardenet, and P. Chainais. Kernel interpolation with continuous volume sampling. In *International Conference on Machine Learning (ICML)*, 2020.
- J. L. Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, sep 1975. ISSN 0001-0782.
- S. Chandrasekhar. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15:1–89, Jan 1943.
- S. Chatterjee, R. Peled, Y. Peres, and D. Romik. Gravitational allocation to Poisson points. *Annals of mathematics*, 172(1):617–671, 2010.
- J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. A tutorial on Palm distributions for spatial point processes. *International Statistical Review*, 85:404–420, 2017.
- J.-F. Coeurjolly, A. Mazoyer, and P.-O. Amblard. Monte Carlo integration of non-differentiable functions on  $[0, 1]^d$ ,  $d = 1, \dots, d$ , using a single determinantal point pattern defined on  $[0, 1]^d$ . *Electronic Journal of Statistics*, 15(2):6228 – 6280, 2021.
- S. Coste. Order, fluctuations, rigidities. 2021.
- B. Delyon and F. Portier. Integral approximation by kernel smoothing. *Bernoulli*, 22(4):2177 – 2208, 2016.
- C. F. Gauss. *Methodus Nova Integralium Valores per Approximationem Inveniendi*. H. Dieterich, 1815.

- G. Gautier, R. Bardenet, and M. Valko. On two ways to use determinantal point processes for Monte Carlo integration. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32. Curran Associates, Inc., 2019a.
- G. Gautier, G. Polito, R. Bardenet, and M. Valko. DPPy: DPP Sampling with Python. *Journal of Machine Learning Research - Machine Learning Open Source Software (JMLR-MLOSS)*, 2019b.
- D. Hawat, G. Gautier, R. Bardenet, and R. Lachièze-Rey. On estimating the structure factor of a point process, with applications to hyperuniformity. *Statistics and Computing*, 33, 2023.
- J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Society, 2009. ISBN 978-0-8218-4373-4.
- M. A. Klatt, J. Lovrić, D. Chen, S. C. Kapfer, F. M. Schaller, P. W. A. Schönhöfer, B. S. Gardiner, A.-S. Smith, G. E. Schröder-Turk, and S. Torquato. Universal hidden order in amorphous cellular geometries. *Nature Communications*, 10(1):811, dec 2019. ISSN 2041-1723.
- G. Last and M. Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.
- R. Leluc, F. Portier, J. Segers, and A. Zhuman. Speeding up Monte Carlo integration: Control neighbors for optimal convergence. 2023.
- N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- B. S. Mityagin. The zero set of a real analytic function. *Mathematical Notes*, 107:529–530, MArch 2020.
- N. Mohd Razali and B. Yap. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Model. Analytics*, 2, 01 2011.
- F. Nazarov, M. Sodin, and A. Volberg. Transportation to random zeroes by the gradient flow. *GAFA Geometric And Functional Analysis*, 17(3):887–935, 2007.
- J. Nolan. *Univariate Stable Distributions, Models for Heavy Tailed Data*. Springer Cham, 01 2020. ISBN 978-3-030-52914-7.
- A. B. Owen. Local antithetic sampling with scrambled nets. *The Annals of Statistics*, 36(5):2319–2343, 2008. ISSN 00905364.
- A. B. Owen. *Monte Carlo theory, methods and examples*. 2013.
- A. Pilleboue, G. Singh, D. Coeurjolly, M. Kazhdan, and V. Ostromoukhov. Variance analysis for monte carlo integration. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, 34(4):14, aug 2015.
- S. Serfaty. Systems of points with Coulomb interactions. *Proc. ICM 2018*, pages 935–977, 2019.
- S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965. ISSN 00063444.
- I. M. Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *Ussr Computational Mathematics and Mathematical Physics*, 7:86–112, 1967.
- L. F. South, C. J. Oates, A. Mira, and C. Drovandi. Regularized zero-variance control variates. *Bayesian Analysis*, 1(1):1–24, 2022.
- D. Stoyan, W. S. Kendall, S. N. Chiu, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, aug 2013. ISBN 9780470664810.
- S. Torquato. Hyperuniform states of matter. *Physics Reports*, 745:1–95, jun 2018. ISSN 03701573.
- P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

## Appendix A: Proofs

In this section, we present the proofs of the results mentioned in Sections 3.2, 3.3 and 3.4. The proof of Proposition 5 can be found in Section A.1. The motion invariance results stated by Proposition 1

and Corollary 1 are proven in Section A.2. The most intricate proof, that of Proposition 2, is presented in Section A.3. Finally, Section A.4 contains the proof of the variance reduction result formulated in Theorem 3, which is the main outcome of the paper.

### A.1. Proofs of Proposition 5

In this section, our objective is to establish Proposition 5. This proposition will play a crucial role in establishing the first part of Corollary 1.

**Proof of Proposition 5.** Let  $\{B_n\}_{n \geq 1}$  be a sequence of disjoint balls, each with the same volume  $|B_1| = 1/\rho$ , and consider the collection of events  $\{\Omega_n\}_{n \geq 1}$  defined by

$$\Omega_n = \{\mathcal{P}(B_n) = 1\}.$$

As  $\mathcal{P}$  is a PPP, the events  $\Omega_1, \Omega_2, \dots$  are independent, and we have

$$\sum_{n \geq 1} \mathbb{P}(\Omega_n) = \sum_{n \geq 1} \exp(-1) = \infty.$$

By Borel-Cantelli,  $\mathbb{P}(\limsup \Omega_n) = 1$ . Thus almost surely, infinitely many  $\Omega_n$  occur, and for  $\mathbf{c} \in \mathbb{R}^d$ ,

$$\begin{aligned} \{F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}\} &\subset \bigcup_{n \geq 1} \{F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}, \Omega_n\} \\ &\subset \bigcup_{n \geq 1} \{\mathbf{X}_n = \mathbf{c} - \mathbf{Y}_n, \Omega_n\}, \end{aligned}$$

where  $\mathbf{X}_n = F_{\mathcal{P} \cap B_n}(\mathbf{x}) - F_{\mathcal{P} \cap B_n}(\mathbf{y})$ , and  $\mathbf{Y}_n \triangleq F_{\mathcal{P} \cap B_n^c}(\mathbf{x}) - F_{\mathcal{P} \cap B_n^c}(\mathbf{y})$ . In particular,

$$\mathbb{P}(F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}) \leq \sum_{n \geq 1} \mathbb{P}(\mathbf{X}_n = \mathbf{c} - \mathbf{Y}_n, \Omega_n). \quad (24)$$

Now, for all  $n$ , conditionally on  $\Omega_n$ ,  $\mathbf{X}_n$  and  $\mathbf{Y}_n$  are independent random vectors and we further claim that  $\mathbf{X}_n$  is continuous, we thus get  $\mathbb{P}(\mathbf{X}_n = \mathbf{c} - \mathbf{Y}_n, \Omega_n) = 0$ . By (24), we conclude that

$$\mathbb{P}(F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}) = 0.$$

Finally, the claim that  $\mathbf{X}_n$  is continuous conditionally on the event  $\Omega_n$  can be easily verified using harmonic function theory. For instance, for  $i \in 1, \dots, d$ , let

$$\begin{aligned} g_i: B_n \setminus \{\mathbf{x}, \mathbf{y}\} &\rightarrow \mathbb{R} \\ \mathbf{z} &\mapsto \frac{x_i - z_i}{\|\mathbf{x} - \mathbf{z}\|_2^d} - \frac{y_i - z_i}{\|\mathbf{y} - \mathbf{z}\|_2^d}. \end{aligned}$$

Actually, for  $\mathbf{z} \in \mathcal{P} \cap B_n$ ,  $g_i(\mathbf{z})$  is the  $i$ -th component of the random vector  $\mathbf{X}_n$ . As  $g_i$  is a non-constant real harmonic function on  $B_n \setminus \{\mathbf{x}, \mathbf{y}\}$ , by Theorem 1.28 of Axler et al. (2001),  $g_i$  is a non-constant real analytic function on  $B_n \setminus \{\mathbf{x}, \mathbf{y}\}$  (which is connected). By Proposition 1 of Mityagin (2020), the

zero-set  $g_i^{-1}(0)$  of  $g_i$  has Lebesgue measure zero. By translation, we can deduce that any level set of  $g_i$  is negligible. Finally, conditioning on  $\Omega_n$ ,  $\mathbf{z}$  is uniformly distributed on  $B_n$ , so

$$\mathbb{P}(\mathbf{X}_n = \mathbf{c} \mid \Omega_n) \leq |B_n \cap \{g_i^{-1}(c_i)\}| = 0.$$

□

## A.2. Proof of Proposition 1 and Corollary 1 (motion-invariance)

Throughout this section, we fix  $\varepsilon \in \mathbb{R}$  and we prove Proposition 1 and Corollary 1.

First, to prove Proposition 1, we show that for a point process  $\mathcal{X}$ , satisfying the conditions of Proposition 1, any void probability (1) of a translation, respectively rotation, of  $\Pi_\varepsilon \mathcal{X}$  is equal to the void probability for  $\Pi_\varepsilon \mathcal{X}$  of the same Borel set. The key argument is that, for any  $\mathbf{x} \in \mathbb{R}^d$ , the force  $F_{\mathcal{X}}(\mathbf{x})$  defined in (F1) is invariant under both translations and rotations.

**Proof of Proposition 1.** Assume that  $\mathcal{X}$  is motion-invariant. As the law of  $\Pi_\varepsilon \mathcal{X}$  is defined by the void probabilities (1), to show that  $\Pi_\varepsilon \mathcal{X}$  is motion-invariant, it is enough to prove that for any Borel set  $B$ , any  $\mathbf{a} \in \mathbb{R}^d$ , and any rotation  $\mathbf{r}$ ,

$$\mathbb{T}_{\mathbf{a} + \Pi_\varepsilon \mathcal{X}}(B) = \mathbb{T}_{\Pi_\varepsilon \mathcal{X}}(B) \quad \text{and} \quad \mathbb{T}_{\mathbf{r}(\Pi_\varepsilon \mathcal{X})}(B) = \mathbb{T}_{\Pi_\varepsilon \mathcal{X}}(B).$$

First, observe that, by construction,  $\Pi_\varepsilon(\mathbf{a} + \mathcal{X}) = \mathbf{a} + \Pi_\varepsilon \mathcal{X}$ . This yields

$$\begin{aligned} \mathbb{T}_{\mathbf{a} + \Pi_\varepsilon \mathcal{X}}(B) &= 1 - \mathbb{P}((\mathbf{a} + \Pi_\varepsilon \mathcal{X}) \cap B = \emptyset) \\ &= 1 - \mathbb{P}(\Pi_\varepsilon(\mathbf{a} + \mathcal{X}) \cap B = \emptyset), \end{aligned}$$

which, by stationarity of  $\mathcal{X}$ , is  $\mathbb{T}_{\Pi_\varepsilon \mathcal{X}}(B)$ . Thus,  $\Pi_\varepsilon \mathcal{X}$  is stationary.

Second, since  $\|\mathbf{r}(\mathbf{x})\|_2 = \|\mathbf{x}\|_2$  and  $\mathbf{r}$  is linear, we can write

$$\begin{aligned} \Pi_\varepsilon(\mathbf{r}(\mathcal{X})) &= \left\{ \mathbf{r}(\mathbf{x}) + \varepsilon \sum_{\substack{\mathbf{z} \in \mathcal{X} \setminus \{\mathbf{x}\} \\ \|\mathbf{r}(\mathbf{x}) - \mathbf{r}(\mathbf{z})\|_2 \uparrow}} \frac{\mathbf{r}(\mathbf{x}) - \mathbf{r}(\mathbf{z})}{\|\mathbf{r}(\mathbf{x}) - \mathbf{r}(\mathbf{z})\|_2^d} \right\}_{\mathbf{x} \in \mathcal{X}} = \left\{ \mathbf{r}(\mathbf{x}) + \varepsilon \sum_{\substack{\mathbf{z} \in \mathcal{X} \setminus \{\mathbf{x}\} \\ \|\mathbf{r}(\mathbf{x} - \mathbf{z})\|_2 \uparrow}} \frac{\mathbf{r}(\mathbf{x} - \mathbf{z})}{\|\mathbf{r}(\mathbf{x} - \mathbf{z})\|_2^d} \right\}_{\mathbf{x} \in \mathcal{X}} \\ &= \left\{ \mathbf{r}(\mathbf{x}) + \varepsilon \sum_{\substack{\mathbf{z} \in \mathcal{X} \setminus \{\mathbf{x}\} \\ \|\mathbf{x} - \mathbf{z}\|_2 \uparrow}} \frac{\mathbf{r}(\mathbf{x} - \mathbf{z})}{\|\mathbf{x} - \mathbf{z}\|_2^d} \right\}_{\mathbf{x} \in \mathcal{X}} = \{\mathbf{r}(\mathbf{x} + \varepsilon F_{\mathcal{X}}(\mathbf{x}))\}_{\mathbf{x} \in \mathcal{X}}. \end{aligned}$$

Thus, we have  $\Pi_\varepsilon(\mathbf{r}(\mathcal{X})) = \mathbf{r}(\Pi_\varepsilon(\mathcal{X}))$ . This implies that

$$\begin{aligned} \mathbb{T}_{\mathbf{r}(\Pi_\varepsilon(\mathcal{X}))}(B) &= 1 - \mathbb{P}(\mathbf{r}(\Pi_\varepsilon(\mathcal{X})) \cap B = \emptyset) \\ &= 1 - \mathbb{P}(\Pi_\varepsilon(\mathbf{r}(\mathcal{X})) \cap B = \emptyset), \end{aligned}$$

which is  $\mathbb{T}_{\Pi_\varepsilon(\mathcal{X})}(B)$  by the isotropy of  $\mathcal{X}$ . Thus,  $\Pi_\varepsilon(\mathcal{X})$  is isotropic. □

We highlight that the proof remains valid when substituting the repulsion operator  $\Pi_\varepsilon$  with its truncated version  $\Pi_\varepsilon^{(q,p)}$  (14), with  $0 \leq q < p$ .

Now, to prove Corollary 1, we first note that a homogeneous Poisson point process  $\mathcal{P}$  is almost surely a valid configuration, as mentioned in Section 3.4. In view of Proposition 1, it is enough to prove that, almost surely, the images of two distinct points from  $\mathcal{P}$  under  $\Pi_\varepsilon$  remain distinct, and that  $\Pi_\varepsilon \mathcal{P}$  possesses the same intensity as  $\mathcal{P}$ .

**Proof of Corollary 1.** To show that almost surely the images of two distinct points from  $\mathcal{P}$  under  $\Pi_\varepsilon$  remain distinct,<sup>7</sup> we need to show that

$$\mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) = \mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y})\}}(\mathbf{x}, \mathbf{y}) \right] = 0. \quad (25)$$

Using the extended Slivnyak-Mecke theorem (4) we get

$$\begin{aligned} & \mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) = \mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y})\}}(\mathbf{x}, \mathbf{y}) \right] \\ &= \mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}(\mathbf{x}) + \varepsilon \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} = \mathbf{y} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}(\mathbf{y}) + \varepsilon \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|_2^d}\}}(\mathbf{x}, \mathbf{y}) \right] \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{E} \left[ \mathbb{1}_{\{\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) + \varepsilon \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} = \mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y}) + \varepsilon \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|_2^d}\}}(\mathbf{x}, \mathbf{y}) \right] \rho^2 d\mathbf{x} d\mathbf{y} \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{P} \left( F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = (\mathbf{y} - \mathbf{x}) \left( \varepsilon^{-1} + 2\|\mathbf{y} - \mathbf{x}\|_2^{-d} \right) \right) \rho^2 d\mathbf{x} d\mathbf{y}. \end{aligned} \quad (26)$$

By Proposition 5, for  $\mathbf{x} \neq \mathbf{y}$ , the random vector  $F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y})$  is continuous. Thus

$$\mathbb{P} \left( F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = (\mathbf{y} - \mathbf{x}) \left( \varepsilon^{-1} + 2\|\mathbf{y} - \mathbf{x}\|_2^{-d} \right) \right) = 0.$$

Plugging back in (26) yields (25).

It remains to show that the intensity of  $\Pi_\varepsilon \mathcal{P}$  is equal to  $\rho$ . Consider a compact  $K$  of  $\mathbb{R}^d$ . Based on the previous reasoning, almost surely, when the repulsion operator  $\Pi_\varepsilon$  is applied to the points of  $\mathcal{P}$ , no two points will end up at the same location. Thus we have

$$\Pi_\varepsilon \mathcal{P}(K) = \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_K(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})).$$

Applying the extended Slivnyak-Mecke theorem (4) we get

$$\mathbb{E} [\Pi_\varepsilon \mathcal{P}(K)] = \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_K(\mathbf{x} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x})) \right] = \int_{\mathbb{R}^d} \mathbb{E} [\mathbb{1}_K(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}))] \rho d\mathbf{x}.$$

As the distribution of  $F_{\mathcal{P}}(\mathbf{x})$  is translation-invariant (in  $\mathbf{x}$ ) we get

$$\mathbb{E} [\Pi_\varepsilon \mathcal{P}(K)] = \int_{\mathbb{R}^d} \mathbb{E} [\mathbb{1}_K(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{0}))] \rho d\mathbf{x}.$$

<sup>7</sup>In the broader framework of point processes, our objective is to establish that  $\Pi_\varepsilon(\mathcal{P})$  is a *simple* point process, assuming that we define  $\Pi_\varepsilon(\mathcal{P})$  as a multiset rather than a set. To accomplish this, we employ (Last and Penrose, 2017, Proposition 6.7).

Exchanging the integral and expectation using Tonelli's theorem gives

$$\begin{aligned}\mathbb{E}[\Pi_\varepsilon \mathcal{P}(K)] &= \rho \mathbb{E} \left[ \int_{\mathbb{R}^d} \mathbb{1}_K(\mathbf{x} + \varepsilon F_\mathcal{P}(\mathbf{0})) d\mathbf{x} \right] \\ &= \rho \mathbb{E} \left[ \int_{\mathbb{R}^d} \mathbb{1}_{K - \varepsilon F_\mathcal{P}(\mathbf{0})}(\mathbf{x}) d\mathbf{x} \right] \\ &= \rho \mathbb{E} [|K - \varepsilon F_\mathcal{P}(\mathbf{0})|] = \rho |K|.\end{aligned}$$

Thus the intensity of  $\Pi_\varepsilon \mathcal{P}$  is equal to  $\rho$ , which completes the proof.  $\square$

### A.3. Proof of Proposition 2 (existence of the moments)

In this section, we consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ , and let  $\varepsilon \in (-1, 1)$ . Our objective is to demonstrate the existence of the moments of the repelled Poisson point process  $\Pi_\varepsilon \mathcal{P}$ . To wit, let  $R > 0$  and  $m$  be a positive integer, we need to show that

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x}) \right)^m \right] < \infty.$$

To accomplish this, the key idea is to decompose the Coulomb force  $F_\mathcal{P}(\mathbf{x})$  acting on  $\mathbf{x} \in \mathbb{R}^d$  and defined in (F<sub>2</sub>) into two truncated sums, one that collects the influence of points close to  $\mathbf{x}$  and the other one of those points in  $\mathcal{P}$  far from  $\mathbf{x}$ ; the two terms shall be controlled by different means. Formally, denote  $B(\mathbf{0}, R)^c \triangleq \mathbb{R}^d \setminus B(\mathbf{0}, R)$ . For  $\mathbf{x} \in \mathbb{R}^d$ , we write

$$F_\mathcal{P}(\mathbf{x}) = F_\mathcal{P}^{(0,1)}(\mathbf{x}) + F_\mathcal{P}^{(1,\infty)}(\mathbf{x}), \quad (27)$$

where the truncated forces are defined in (13). We refer to the first term in the right-hand side of (27) as the “internal” force, and to the second term as the “external” force. In words, our proof works by showing that, for  $\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c$  to be pushed inside  $B(\mathbf{0}, R)$ , i.e. for  $\mathbf{x} + \varepsilon F_\mathcal{P}(\mathbf{x}) \in B(\mathbf{0}, R)$ , one of two low-probability events must occur. One of these events involves the internal force, and the other one the external force.

Let  $0 < \gamma < 1/(d-1)$ ,  $0 < \beta < \gamma/(d-1)$ , and  $r : \mathbf{x} \mapsto \|\mathbf{x}\|_2^\beta$ . Let also  $R' = (R + m - 1)^{1/\beta}$ . Now, for  $\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c$ , it holds

$$\{\mathbf{x} + \varepsilon F_\mathcal{P}(\mathbf{x}) \in B(\mathbf{0}, R)\} \subset \left\{ \mathbf{x} + \varepsilon F_\mathcal{P}^{(0,1)}(\mathbf{x}) \in B(\mathbf{0}, r(\mathbf{x})) \right\} \cup \left\{ \|F_\mathcal{P}^{(1,\infty)}(\mathbf{x})\|_2 \geq \frac{r(\mathbf{x}) - R}{|\varepsilon|} \right\}. \quad (28)$$

To see the validity of this inclusion, note that if  $\mathbf{x}$  is not in the right-hand side of (28), then

$$\|\mathbf{x} + \varepsilon F_\mathcal{P}(\mathbf{x})\|_2 \geq \|\mathbf{x} + \varepsilon F_\mathcal{P}^{(0,1)}(\mathbf{x})\|_2 - \|\varepsilon F_\mathcal{P}^{(1,\infty)}(\mathbf{x})\|_2 > r(\mathbf{x}) - (r(\mathbf{x}) - R) = R.$$

Now, using (28), we get

$$\begin{aligned}\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x}) \right)^m \right] &\leq \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x}) + \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_\mathcal{P}^{(0,1)}(\mathbf{x})) \right. \right. \\ &\quad \left. \left. + \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\{\|F_\mathcal{P}^{(1,\infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x}) - R)/|\varepsilon|\}}(\mathbf{x}) \right)^m \right].\end{aligned}$$

By convexity of  $h : x \mapsto x^m$  on  $\mathbb{R}^+ \setminus \{0\}$ , it further comes

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \Pi_\varepsilon \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x}) \right)^m \right] &\leq 3^{m-1} \left( \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x}) \right)^m \right] \right. \\ &\quad + \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^m \right] \\ &\quad \left. + \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x}) - R)/|\varepsilon|\}}(\mathbf{x}) \right)^m \right] \right). \end{aligned} \quad (29)$$

The first term on the right-hand side of (29) is finite since  $\mathcal{P}$  is a PPP. The rest of the proof consists in proving that the remaining two terms are finite, which will be a consequence of Corollaries 2 and 3.

We first focus on the term in (29) involving the external force.

**Lemma 1.** *Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . There exist  $c_1, c_2, c_3 > 0$  such that for all  $p > q > 0$  and  $t > 0$  we have*

$$\mathbb{P} \left( \|F_{\mathcal{P}}^{(q,p)}(\mathbf{0})\|_2 > t \right) \leq c_1 \exp \left( -c_2 q^{d-1} t \log \left( \frac{c_3 t}{q} \right) \right). \quad (30)$$

Note that by translation-invariance, the choice of  $\mathbf{0}$  in (30) is arbitrary. In addition, when  $\rho = 1$ , Equation (30) corresponds to Equation (32) in Theorem 16 of Chatterjee et al. (2010). While the paper does not offer an exhaustive proof of this equation, it does provide a similar and detailed proof for another equation within the same theorem. For completeness, we provide here a simple proof of Lemma 1, valid for any value of  $\rho > 0$ . The proof involves bounding the exponential moment of each component in the random vector  $F_{\mathcal{P}}^{(q,p)}(\mathbf{0})$  and using Markov's inequality to obtain a tail bound. The Poisson assumption then helps simplify the bound.

**Proof of Lemma 1.** In order to streamline the notations used in this proof, we set  $G^{(q,p)} = -F_{\mathcal{P}}^{(q,p)}$ , with  $G_i^{(q,p)}$  its  $i$ -th component.

If we prove that for any  $i \in \{1, \dots, d\}$ , there exist  $C_1, C_2, C_3 > 0$  such that

$$\mathbb{P} \left( G_i^{(q,p)}(\mathbf{0}) > t \right) \leq C_1 \exp \left( -C_2 q^{d-1} t \log \left( \frac{C_3 t}{q} \right) \right). \quad (31)$$

Using that  $-\mathcal{P}$  is also a PPP of intensity  $\rho$ , it follows that

$$\mathbb{P} \left( G_i^{(q,p)}(\mathbf{0}) < -t \right) \leq C_1 \exp \left( -C_2 q^{d-1} t \log \left( \frac{C_3 t}{q} \right) \right).$$

Combining this with (31), we obtain (30), with  $c_1 = 2dC_1$ ,  $c_2 = C_2/d$  and  $c_3 = C_3/d$ . Thus, we only need to verify Equation (31).

Let  $\theta \geq 0$ , using Markov's inequality we have

$$\mathbb{P} \left( G_i^{(q,p)}(\mathbf{0}) > t \right) = \mathbb{P} \left( \exp \left( \theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d} \right) > \exp(\theta t) \right)$$

$$\leq \mathbb{E} \left[ \exp \left( \theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d} \right) \right] \exp(-\theta t). \quad (32)$$

Conditioning on  $\mathcal{P}(A^{(q,p)})$ , the points of  $\mathcal{P} \cap A^{(q,p)}$  are independent and uniformly distributed in  $A^{(q,p)}$ . Let  $\mathbf{u}$  be a uniform r.v. in  $A^{(q,p)}$ , and  $N > 0$ . By symmetry,  $\mathbb{E}[\frac{\mathbf{u}}{\|\mathbf{u}\|_2^d}] = 0$  and we have

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d} \right) \middle| \mathcal{P}(A^{(q,p)}) = N \right] &= \mathbb{E} \left[ \prod_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \exp \left( \theta \frac{z_i}{\|\mathbf{z}\|_2^d} \right) \middle| \mathcal{P}(A^{(q,p)}) = N \right] \\ &= \mathbb{E} \left[ \exp \left( \theta \frac{u_i}{\|\mathbf{u}\|_2^d} \right) \right]^N \\ &= \left( 1 + \mathbb{E} \left[ \sum_{k \geq 2} \frac{1}{k!} \theta^k \frac{u_i^k}{\|\mathbf{u}\|_2^{kd}} \right] \right)^N. \end{aligned}$$

In particular,

$$\mathbb{E} \left[ \exp \left( \theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d} \right) \middle| \mathcal{P}(A^{(q,p)}) = N \right] \leq \left( 1 + \sum_{k \geq 2} \frac{1}{k!} \theta^k \mathbb{E} \left[ \|\mathbf{u}\|_2^{-k(d-1)} \right] \right)^N. \quad (33)$$

Recall that the surface area of the unit ball of  $\mathbb{R}^d$  is equal to  $d\kappa_d$ . As  $\mathbf{u}$  has uniform distribution on  $A^{(q,p)}$  we get

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{\|\mathbf{u}\|_2^{k(d-1)}} \right] &= \frac{1}{|A^{(q,p)}|} \int_{A^{(q,p)}} \frac{1}{\|\mathbf{u}\|_2^{k(d-1)}} d\mathbf{u} \\ &= \frac{d\kappa_d}{|A^{(q,p)}|} \int_q^p r^{d-1-k(d-1)} dr. \end{aligned}$$

As  $k \geq 2$  and  $d \geq 3$ , we have  $k > d/(d-1)$ . Thus

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{\|\mathbf{u}\|_2^{k(d-1)}} \right] &= \frac{d\kappa_d}{|A^{(q,p)}| (k(d-1) - d)} \left( \frac{1}{q^{k(d-1)-d}} - \frac{1}{p^{k(d-1)-d}} \right) \\ &\leq \frac{d\kappa_d}{|A^{(q,p)}| q^{k(d-1)-d}}. \end{aligned}$$

Plugging back into (32), we obtain

$$\mathbb{P} \left( G_i^{(q,p)}(\mathbf{0}) > t \right) \leq \mathbb{E} \left[ \left( 1 + \frac{d\kappa_d q^d}{|A^{(q,p)}|} \sum_{k \geq 2} \frac{1}{k!} \left( \frac{\theta}{q^{d-1}} \right)^k \right)^{\mathcal{P}(A^{(q,p)})} \right] \exp(-\theta t)$$

$$\leq \mathbb{E} \left[ \left( 1 + \frac{d\kappa_d q^d}{|A^{(q,p)}|} \exp \left( \frac{\theta}{q^{d-1}} \right) \right)^{\mathcal{P}(A^{(q,p)})} \right] \exp(-\theta t).$$

Now, remembering that  $\mathcal{P}(A^{(q,p)})$  is a Poisson random variable of parameter  $\rho|A^{(q,p)}|$ , we obtain<sup>8</sup>

$$\begin{aligned} \mathbb{P}(G_i^{(q,p)}(\mathbf{0}) > t) &\leq \exp \left( \rho|A^{(q,p)}| \frac{d\kappa_d q^d}{|A^{(q,p)}|} \exp \left( \frac{\theta}{q^{d-1}} \right) \right) \exp(-\theta t) \\ &= \exp \left( d\kappa_d \rho q^d \exp \left( \frac{\theta}{q^{d-1}} \right) - \theta t \right). \end{aligned}$$

Taking  $\theta = q^{d-1} \log \left( \frac{t}{d\kappa_d \rho q} \right)$ , we get

$$\begin{aligned} \mathbb{P}(G_i^{(q,p)}(\mathbf{0}) > t) &\leq \exp \left( -t q^{d-1} \log \left( \frac{t}{d\kappa_d \rho q} \right) + t q^{d-1} \right) \\ &= \exp \left( -t q^{d-1} \log \left( \frac{t}{ed\kappa_d \rho q} \right) \right), \end{aligned}$$

which ends the proof.  $\square$

Lemma 1 has the following corollary.

**Corollary 2.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $R > 0$ ,  $\varepsilon \in (-1, 1)$ ,  $\beta \in (0, 1)$  and  $r(\mathbf{x}) = \|\mathbf{x}\|_2^\beta$ . Then, for any positive integer  $m$ , there exist positive constants  $(a_k)_{k=1}^m$ ,  $(b_k)_{k=1}^m$ , and  $(c_k)_{k=1}^m$  such that

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\left\{ \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x})-R}{|\varepsilon|} \right\}}(\mathbf{x}) \right)^m \right] &\leq \\ &\sum_{k=1}^m a_k \left( \int_{B(\mathbf{0}, R')^c} \exp(-b_k g_k(\mathbf{x}) \log(c_k g_k(\mathbf{x}))) \rho d\mathbf{x} \right)^k, \end{aligned} \quad (34)$$

where  $R' = (R + m - 1)^{1/\beta}$  and  $g_k(\mathbf{x}) = \frac{r(\mathbf{x}) - (R + k - 1)}{|\varepsilon|^k}$ .

This corollary helps us control the external term in (27). Indeed, as  $\|\mathbf{x}\|_2 \rightarrow \infty$  we have,

$$\exp(-b_k g_k(\mathbf{x}) \log(c_k g_k(\mathbf{x}))) = o \left( \exp \left( -\|\mathbf{x}\|_2^\beta \right) \right).$$

Thus, for any positive integer  $m$ ,

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\left\{ \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x})-R}{|\varepsilon|} \right\}}(\mathbf{x}) \right)^m \right] < \infty. \quad (35)$$

<sup>8</sup>If  $X$  is a Poisson random variable of parameter  $\lambda$ , then for any  $\gamma$  the mean of the random variable  $(1 + \gamma)^X$  is equal to  $\exp(\lambda\gamma)$ .

**Proof of Corollary 2.** Fix a positive integer  $m$ . There exists  $m$  constants  $(d_i)_{i=1}^m$  such that

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x}) - R}{|\varepsilon|}\}}(\mathbf{x}) \right)^m \right] = \sum_{k=1}^m d_k \underbrace{\mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathcal{P} \cap B(\mathbf{0}, R')^c}^{\neq} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}_1)\|_2 > \frac{r(\mathbf{x}_1) - R}{|\varepsilon|}\}}(\mathbf{x}_1) \cdots \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}_k)\|_2 > \frac{r(\mathbf{x}_k) - R}{|\varepsilon|}\}}(\mathbf{x}_k) \right]}_{\triangleq E_k}. \quad (36)$$

Using Lemma 1 we will show that for any  $k \geq 1$  there exists positive constants  $a_k$ ,  $b_k$  and  $c_k$  such that

$$E_k \leq \left( \int_{B(\mathbf{0}, R')^c} a_k \exp[-b_k g_k(\mathbf{x}) \log(c_k g_k(\mathbf{x}))] \rho d\mathbf{x} \right)^k,$$

with  $g_k(\mathbf{x}) = \frac{r(\mathbf{x}) - (R+k-1)}{|\varepsilon|k}$  for  $k \in \{1, \dots, m\}$ .

To simplify the notations, we denote  $\mathcal{P} \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$  by  $\hat{\mathcal{P}}^{(k)}$  for any  $k \in \{1, \dots, m\}$  and sometimes omit to remind that  $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{P} \cap B(\mathbf{0}, R')^c$  when it is clear from the context.

First, remark that for two distinct points  $\mathbf{x}$  and  $\mathbf{y}$  of  $\mathbb{R}^d$ , we have

$$F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}) = \begin{cases} F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}^{(1, \infty)}(\mathbf{x}) & \text{if } \|\mathbf{x} - \mathbf{y}\|_2 < 1 \\ F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}^{(1, \infty)}(\mathbf{x}) + \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} & \text{if } \|\mathbf{x} - \mathbf{y}\|_2 \geq 1 \end{cases}.$$

Thus

$$\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x})\|_2 \leq \|F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}^{(1, \infty)}(\mathbf{x})\|_2 + 1.$$

In particular

$$\mathbb{1}_{\{\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x}) - R)/|\varepsilon|\}}(\mathbf{x}) \leq \mathbb{1}_{\{\|F_{\mathcal{P} \setminus \{\mathbf{x}, \mathbf{y}\}}^{(1, \infty)}(\mathbf{x})\|_2 \geq 2g_2(\mathbf{x})\}}(\mathbf{x}). \quad (37)$$

Generalizing Equation (37) to  $k$  points gives

$$\mathbb{1}_{\{\|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x}) - R)/|\varepsilon|\}}(\mathbf{x}) \leq \mathbb{1}_{\{\|F_{\hat{\mathcal{P}}^{(k)}}^{(1, \infty)}(\mathbf{x})\|_2 \geq k g_k(\mathbf{x})\}}(\mathbf{x}). \quad (38)$$

Using Equation (38) and the extended Slivnyak-Mecke theorem (4) we get

$$\begin{aligned} E_k &\leq \mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_k}^{\neq} \mathbb{1}_{\{\|F_{\hat{\mathcal{P}}^{(k)}}^{(1, \infty)}(\mathbf{x}_1)\|_2 > k g_k(\mathbf{x}_1), \dots, \|F_{\hat{\mathcal{P}}^{(k)}}^{(1, \infty)}(\mathbf{x}_k)\|_2 > k g_k(\mathbf{x}_k)\}}(\mathbf{x}_1, \dots, \mathbf{x}_k) \right] \\ &= \int_{(B(\mathbf{0}, R')^c)^k} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}_1)\|_2 > k g_k(\mathbf{x}_1), \dots, \|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}_k)\|_2 > k g_k(\mathbf{x}_k) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k \\ &\leq \int_{(B(\mathbf{0}, R')^c)^k} \min_{j \in \{1, \dots, k\}} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1, \infty)}(\mathbf{x}_j)\|_2 \geq k g_k(\mathbf{x}_j) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k. \end{aligned}$$

As the distribution of  $F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})$  is translation invariant (w.r.t.  $\mathbf{x}$ ), we get

$$\begin{aligned} E_k &\leq \int_{(B(\mathbf{0}, R')^c)^k} \min_{j \in \{1, \dots, k\}} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq k g_k(\mathbf{x}_j) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k \\ &= \int_{(B(\mathbf{0}, R')^c)^k} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq \max_{j \in \{1, \dots, k\}} k g_k(\mathbf{x}_j) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k \\ &\leq \int_{(B(\mathbf{0}, R')^c)^k} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq \sum_{j=1}^k g_k(\mathbf{x}_j) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k. \end{aligned} \quad (39)$$

For  $\mathbf{x}_1, \dots, \mathbf{x}_k \in B(\mathbf{0}, R')^c$ , Lemma 1 with  $q = 1$  and  $p = \infty$  guarantees the existence of  $C_1, C_2, C_3 > 0$  such that

$$\begin{aligned} \mathbb{P} \left( \|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq \sum_{j=1}^k g_k(\mathbf{x}_j) \right) &\leq C_1 \exp \left( -C_2 \sum_{j=1}^k g_k(\mathbf{x}_j) \log \left( C_3 \sum_{i=1}^k g_k(\mathbf{x}_i) \right) \right) \\ &= C_1 \prod_{j=1}^k \exp \left( -C_2 g_k(\mathbf{x}_j) \log \left( C_3 \sum_{i=1}^k g_k(\mathbf{x}_i) \right) \right) \\ &\leq C_1 \prod_{j=1}^k \exp \left( -C_2 g_k(\mathbf{x}_j) \log (C_3 g_k(\mathbf{x}_j)) \right). \end{aligned}$$

Plugging back into (39), and then in (36), we obtain the existence of positive constants  $\{a_k\}_{k=1}^m$ ,  $\{b_k\}_{k=1}^m$ , and  $\{c_k\}_{k=1}^m$  such that

$$\begin{aligned} &\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x}) - R}{|\varepsilon|}\}}(\mathbf{x}) \right)^m \right] \\ &\leq \sum_{k=1}^m d_k a_k \int_{(B(\mathbf{0}, R')^c)^k} \prod_{j=1}^k \exp \left( -b_k g_k(\mathbf{x}_j) \log (c_k g_k(\mathbf{x}_j)) \right) \rho^k d\mathbf{x}_1 \dots d\mathbf{x}_k \\ &= \sum_{k=1}^m d_k a_k \left( \int_{B(\mathbf{0}, R')^c} \exp \left( -b_k g_k(\mathbf{x}) \log (c_k g_k(\mathbf{x})) \right) \rho d\mathbf{x} \right)^k, \end{aligned}$$

which concludes the proof.  $\square$

Now we switch focus to bounding the contribution from the internal force to (27). Again, we work with a lemma and a corollary.

**Lemma 2.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $R > 0$  and  $\varepsilon \in (-1, 1)$ . Consider a function  $r : \mathbb{R}^d \rightarrow \mathbb{R}^+ \setminus \{0\}$  such that  $r(\mathbf{x}) < \|\mathbf{x}\|_2$  for any  $\mathbf{x} \in B(\mathbf{0}, R)^c$ . Then, for any  $\mathbf{x} \in B(\mathbf{0}, R)^c$ , and  $0 < \gamma < 1/(d-1)$ , there exists  $c_1, \dots, c_5 > 0$  such that

$$\mathbb{P} \left( \mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{0}, r(\mathbf{x})) \right) \leq \frac{c_1}{h(\mathbf{x})g(\mathbf{x})^{1+\gamma}} + c_3 \exp \left( -c_4 g(\mathbf{x}) \log (c_5 g(\mathbf{x})) \right), \quad (40)$$

where  $h(\mathbf{x}) = \max\left(1, c_2 (\|\mathbf{x}\|_2 / r(\mathbf{x}) - 1)^{d-1} - 1\right)$ , and  $g(\mathbf{x}) = (\|\mathbf{x}\|_2 - r(\mathbf{x})) / |\varepsilon|$ .

First, by translation-invariance of the distribution of  $F_{\mathcal{P}}^{(q,p)}(\mathbf{x})$ , the choice of  $\mathbf{0}$  in (40) is arbitrary. Second, by choosing  $r(\mathbf{x}) = \|\mathbf{x}\|_2^\beta$ , with  $0 < \beta < 1/(2(d-1))$ , the upper bound in Equation (40) is  $o(\|\mathbf{x}\|_2^{-d})$  as  $\|\mathbf{x}\|_2$  goes to infinity. Thus,  $\mathbb{P}\left(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{0}, r(\mathbf{x}))\right)$  is integrable over  $B(\mathbf{0}, R)^c$ .

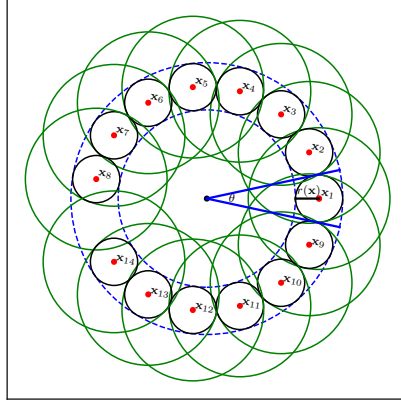


Figure 11: Illustration of the proof idea.

**Proof of Lemma 2.** Fix  $\mathbf{x} \in B(\mathbf{0}, R)^c$ , and let  $q = \|\mathbf{x}\|_2 - r(\mathbf{x})$  and  $p = \|\mathbf{x}\|_2 + r(\mathbf{x})$ . Pick points  $\{\mathbf{x}_i\}_{i=1}^m$  iteratively on the sphere  $S(\mathbf{0}, \|\mathbf{x}\|_2)$  with  $\mathbf{x}_1 = -\mathbf{x}$ , such that the balls  $\{B(\mathbf{x}_i, r(\mathbf{x}))\}_{i=1}^m$  are disjoint, and it is not possible to add an additional similar ball not overlapping the previous ones. Next, let  $B_i = B(\mathbf{x}_i, 3r(\mathbf{x}))$  and  $C_1 \triangleq \kappa_d (3r(\mathbf{x}))^d$  be the volume of  $B_1$ . Then,  $\bigcup_{i=1}^m B_i$  is a covering of the annulus  $A^{(q,p)}$ . To see why this holds, suppose that there exists a point  $\mathbf{y} \in A^{(q,p)} \setminus \bigcup_{i=1}^m B_i$ . In particular,  $\min_i \|\mathbf{y} - \mathbf{x}_i\|_2 > 3r(\mathbf{x})$ . Take  $\mathbf{z} = \mathbf{y} \frac{\|\mathbf{x}\|_2}{\|\mathbf{y}\|_2}$ ,  $\mathbf{z}$  is in fact the orthogonal projection of  $\mathbf{y}$  onto  $S(\mathbf{0}, \|\mathbf{x}\|_2)$ . For any  $i \in \{1, \dots, m\}$

$$\begin{aligned} \|\mathbf{z} - \mathbf{x}_i\|_2 &\geq \|\mathbf{y} - \mathbf{x}_i\|_2 - \|\mathbf{z} - \mathbf{y}\| \\ &> 3r(\mathbf{x}) - \|\mathbf{y}\|_2 - \|\mathbf{x}\|_2 \\ &> 2r(\mathbf{x}). \end{aligned}$$

Thus, we have  $\mathbf{z} \in S(\mathbf{0}, \|\mathbf{x}\|_2)$  and  $B(\mathbf{z}, r(\mathbf{x})) \subset A^{(q,p)} \setminus \bigcup_{i=1}^m B(\mathbf{x}_i, r(\mathbf{x}))$  giving a contradiction. While our work pertains to dimensions greater than two, employing a visualization that depicts the two-dimensional case can aid in comprehending the concept. Figure 11 illustrates an example of valid points  $\{\mathbf{x}_i\}_{i=1}^m$  in red and the covering in green of  $A^{(q,p)}$  for  $d=2$ . Note that, for  $d=2$  and any  $i \in \{1, \dots, m\}$ ,  $\mathbf{x}_i$  can be chosen inductively as the rotation of  $\mathbf{x}_{i-1}$  of angle  $\theta = \text{Arcsin}(r(\mathbf{x})/\|\mathbf{x}\|_2)$  around the origin.

As the balls  $\{B(\mathbf{x}_i, r(\mathbf{x}))\}_{i=1}^m$  are disjoint and contained in  $A^{(q,p)}$ , we have

$$\mathbb{P}\left(\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in A^{(q,p)}\right) \geq \mathbb{P}\left(\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in \bigcup_{i=1}^m B(\mathbf{x}_i, r(\mathbf{x}))\right)$$

$$= \sum_{i=1}^m \mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_i, r(\mathbf{x})) \right)$$

By isotropy of the law of  $F_{\mathcal{P}}^{(0,1)}(\mathbf{0})$ , we obtain

$$\mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in A^{(q,p)} \right) \geq m \mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_1, r(\mathbf{x})) \right). \quad (41)$$

We will now proceed to find a suitable lower bound for  $m$ . To accomplish this, we apply the mean value theorem to

$$h: [\|\mathbf{x}\|_2 - r(\mathbf{x}), \|\mathbf{x}\|_2 + r(\mathbf{x})] \rightarrow \mathbb{R} \\ x \mapsto x^d,$$

and we get that there exists  $c \in (-1, 1)$  such that

$$(\|\mathbf{x}\|_2 + r(\mathbf{x}))^d - (\|\mathbf{x}\|_2 - r(\mathbf{x}))^d = 2dr(\mathbf{x})(\|\mathbf{x}\|_2 + cr(\mathbf{x}))^{d-1} \geq 2dr(\mathbf{x})(\|\mathbf{x}\|_2 - r(\mathbf{x}))^{d-1}.$$

Hence

$$\begin{aligned} m &\geq \left\lceil \frac{|A^{(q,p)}|}{C_1} \right\rceil \\ &= \left\lceil \frac{\kappa_d ((\|\mathbf{x}\|_2 + r(\mathbf{x}))^d - (\|\mathbf{x}\|_2 - r(\mathbf{x}))^d)}{\kappa_d (3r(\mathbf{x}))^d} \right\rceil \\ &\geq \left\lceil \frac{\kappa_d (2dr(\mathbf{x})(\|\mathbf{x}\|_2 - r(\mathbf{x}))^{d-1})}{\kappa_d (3r(\mathbf{x}))^d} \right\rceil \geq \max \left( 1, C_2 \left( \frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1 \right)^{d-1} - 1 \right), \end{aligned}$$

where  $C_2 = (2d)/3^d$ . Thus Equation (41) leads to

$$\begin{aligned} \mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_1, r(\mathbf{x})) \right) &\leq \frac{\mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in A^{(q,p)} \right)}{\max \left( 1, C_2 \left( \frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1 \right)^{d-1} - 1 \right)} \\ &\leq \frac{\mathbb{P} \left( \|\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0})\|_2 > \|\mathbf{x}\|_2 - r(\mathbf{x}) \right)}{\max \left( 1, C_2 \left( \frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1 \right)^{d-1} - 1 \right)}. \end{aligned}$$

Since  $F_{\mathcal{P}}^{(0,1)} = F_{\mathcal{P}} - F_{\mathcal{P}}^{(1,\infty)}$ , we get

$$\mathbb{P} \left( \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_1, r(\mathbf{x})) \right) \leq \frac{\mathbb{P} \left( |\varepsilon| \|F_{\mathcal{P}}(\mathbf{0}) - F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 > \|\mathbf{x}\|_2 - r(\mathbf{x}) \right)}{\max \left( 1, C_2 \left( \frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1 \right)^{d-1} - 1 \right)}$$

$$\leq \frac{\mathbb{P}\left(\|F_{\mathcal{P}}(\mathbf{0})\|_2 > \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|}\right) + \mathbb{P}\left(\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 > \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|}\right)}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right)} \quad (42)$$

By Lemma 4, for any  $0 < \gamma < 1/(d-1)$  we have  $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^{1+\gamma}] < \infty$ . Applying Markov's inequality to the first part of Equation (42), and using Lemma 1 with  $t = (\|\mathbf{x}\|_2 - r(\mathbf{x}))/(2|\varepsilon|)$  for the last part, we obtain the existence of positive constants  $c_1, c_2$ , and  $c_3$  such that

$$\begin{aligned} & \mathbb{P}\left(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{0}, r(\mathbf{x}))\right) \\ & \leq \frac{(2|\varepsilon|)^{1+\gamma} \mathbb{E}\left[\|F_{\mathcal{P}}(\mathbf{0})\|_2^{1+\gamma}\right]}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right) (\|\mathbf{x}\|_2 - r(\mathbf{x}))^{1+\gamma}} + \frac{c_1 \exp\left(-c_2 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|} \log\left(c_3 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|}\right)\right)}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right)} \\ & \leq \frac{C_3 |\varepsilon|^{1+\gamma}}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right) (\|\mathbf{x}\|_2 - r(\mathbf{x}))^{1+\gamma}} + C_4 \exp\left(-C_5 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{|\varepsilon|} \log\left(C_6 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{|\varepsilon|}\right)\right) \end{aligned}$$

with  $C_3 = 2^{1+\gamma} \mathbb{E}\left[\|F_{\mathcal{P}}(\mathbf{0})\|_2^{1+\gamma}\right]$ ,  $C_4 = c_1$ ,  $C_5 = c_2/2$  and  $C_6 = c_3/2$ .  $\square$

Note that under the assumptions and definitions of Lemmas 1, 2, for  $\mathbf{x} \in B(\mathbf{0}, R^{1/\beta})^c$ , there exists positive constants  $c_1, \dots, c_5$  such that

$$\mathbb{P}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) \in B(\mathbf{0}, R)) \leq \frac{c_1}{\max\left(1, c_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right) g(\mathbf{x})^{1+\gamma}} + c_3 \exp(-c_4 g(\mathbf{x}) \log(c_5 g(\mathbf{x}))), \quad (43)$$

for any  $0 < \gamma < 1/(d-1)$ . As expected, we observe that

$$\mathbb{P}(\mathbf{x} + \varepsilon F_{\mathcal{P}} \in B(\mathbf{0}, R)) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

We will later see that by selecting an appropriate function  $r$ , typically for  $r(\mathbf{x}) = \|\mathbf{x}\|_2^\beta$  with  $0 < \beta < \frac{\gamma}{d-1}$ , the bound in Equation (43) converges fast enough to zero allowing to bound the moments of  $\sum_{\mathbf{x} \in \Pi_{\varepsilon} \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x})$ .

The next result is a corollary of Lemma 2.

**Corollary 3.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $R > 0$ ,  $\varepsilon \in (-1, 1)$  and  $\beta \in (0, 1)$ . Set  $r(\mathbf{x}) = \|\mathbf{x}\|_2^\beta$ ,  $g(\mathbf{x}) = (\|\mathbf{x}\|_2 - r(\mathbf{x}))/|\varepsilon|$  and denote

$$E_k \triangleq \mathbb{E} \left[ \sum_{\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathcal{P} \cap B(\mathbf{0}, R)^c}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))}(\mathbf{x}_1 + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1)) \dots \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_k))}(\mathbf{x}_k + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_k)) \right]. \quad (44)$$

For any positive integer  $m$  and  $0 < \gamma < 1/(d-1)$  there exists positive constants  $a_1, \dots, a_{m-1}$  and  $c_1, \dots, c_5$  such that

$$E_1 \leq \int_{B(\mathbf{0}, R)^c} \frac{c_1}{\max\left(1, c_2 \left(\|\mathbf{x}\|_2^{1-\beta} - 1\right)^{d-1} - 1\right) g(\mathbf{x})^{1+\gamma}} + c_3 \exp(-c_4 g(\mathbf{x}) \log(c_5 g(\mathbf{x}))) \rho d\mathbf{x}, \quad (45)$$

and

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, f(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^m \right] \leq \sum_{k=1}^{m-1} \left( a_k E_k + b^{k-1} E_1 E_{m-k} \right) + b^{m-1} E_1. \quad (46)$$

where  $b = 2^d \kappa_d \rho$ .

Taking  $0 < \beta < \frac{\gamma}{d-1}$ , the integrand in Equation (45) is  $o(\|\mathbf{x}\|_2^{-d})$  as  $\|\mathbf{x}\|_2 \rightarrow \infty$  implying that

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^m \right] < \infty, \quad (47)$$

for any positive integer  $m$ .

**Proof of Corollary 3.** We will show the validity of Equation (46) by induction on  $m \geq 1$ .

To begin with, for  $m = 1$  using the extended Slivnyak-Mecke theorem (4) and that the law of  $F_{\mathcal{P}}^{(0,1)}(\mathbf{x})$  is translation-invariant (w.r.t.  $\mathbf{x}$ ) we have

$$E_1 = \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}\}}^{(0,1)}(\mathbf{x})) \right] = \int_{B(\mathbf{0}, R)^c} \mathbb{P} \left( \mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{0}, r(\mathbf{x})) \right) \rho d\mathbf{x}.$$

By Lemma 2, for any  $0 < \gamma < 1/(d-1)$  there exists positive constants  $c_1, \dots, c_5$  s.t.

$$E_1 \leq \int_{B(\mathbf{0}, R)^c} \frac{c_1}{\max\left(1, c_2 \left(\|\mathbf{x}\|_2^{1-\beta} - 1\right)^{d-1} - 1\right) g(\mathbf{x})^{1+\gamma}} + c_3 \exp(-c_4 g(\mathbf{x}) \log(c_5 g(\mathbf{x}))) \rho d\mathbf{x}.$$

Next, suppose that Equation (46) is valid until  $m$ , and let's verify that it holds for  $m+1$ . There exists a sequence of constants  $(d_k)_{k=1}^m$  such that

$$\mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^{m+1} \right] = \sum_{k=1}^m d_k E_k + E_{m+1}. \quad (48)$$

We only need to focus on finding an upper bound of  $E_{m+1}$ .

To simplify the notations, we denote  $\mathcal{P} \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$  by  $\hat{\mathcal{P}}^{(k)}$  for any  $k \in \{1, \dots, m\}$  and sometimes omit to remind that  $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{P} \cap B(\mathbf{0}, R)^c$  when it is clear from the context.

To begin, we break down the sum that defines  $E_{m+1}$  according to the count  $k$  of the points of  $\mathcal{P} \cap B(\mathbf{0}, R)^c$  that fall within a ball of radius 2 centered at  $\mathbf{x}_1$  as follows

$$E_{m+1} = \mathbb{E} \left[ \underbrace{\sum_{\substack{\mathbf{x}_1, \dots, \mathbf{x}_{m+1} \\ \max_{1 \leq i} \|\mathbf{x}_i - \mathbf{x}_1\|_2 \leq 2}}^{\neq} \mathbb{1}_{B(\mathbf{x}_1, r(\mathbf{x}_1))}(-\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1)) \cdots \mathbb{1}_{B(\mathbf{x}_{m+1}, r(\mathbf{x}_{m+1}))}(-\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_{m+1}))}_{A_{m+1}} \right. \\ \left. + \sum_{k=1}^m \mathbb{E} \left[ \underbrace{\sum_{\substack{\mathbf{x}_1, \dots, \mathbf{x}_{m+1} \\ \max_{1 \leq i \leq k} \|\mathbf{x}_i - \mathbf{x}_1\|_2 \leq 2 \\ \min_{k < i \leq m+1} \|\mathbf{x}_i - \mathbf{x}_1\|_2 > 2}}^{\neq} \mathbb{1}_{B(\mathbf{x}_1, r(\mathbf{x}_1))}(-\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1)) \cdots \mathbb{1}_{B(\mathbf{x}_{m+1}, r(\mathbf{x}_{m+1}))}(-\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_{m+1}))}_{A_k} \right] \right].$$

First, for  $A_{m+1}$  the extended Slivnyak-Mecke theorem (4) followed by a change of variables gives

$$\begin{aligned} A_{m+1} &\leq \mathbb{E} \left[ \sum_{\substack{\mathbf{x}_1, \dots, \mathbf{x}_{m+1} \\ \max_{1 \leq i} \|\mathbf{x}_i - \mathbf{x}_1\|_2 \leq 2}}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))} \left( \mathbf{x}_1 + \varepsilon F_{\mathcal{P}^{(m+1)}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^{m+1} \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d} \right) \right] \\ &= \int_{B(\mathbf{0}, R)^c} \int_{B(\mathbf{x}_1, 2)^m} \mathbb{P} \left( \mathbf{x}_1 + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1) \in B \left( \varepsilon \sum_{j=2}^{m+1} \frac{\mathbf{x}_j - \mathbf{x}_1}{\|\mathbf{x}_j - \mathbf{x}_1\|_2^d}, r(\mathbf{x}_1) \right) \right) \rho^{m+1} d\mathbf{x}_{m+1} \dots d\mathbf{x}_1 \\ &= \int_{B(\mathbf{0}, 2)^m \times B(\mathbf{0}, R)^c} \mathbb{P} \left( \mathbf{x}_1 + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1) \in B \left( \varepsilon \sum_{j=1}^m \frac{\mathbf{t}_j}{\|\mathbf{t}_j\|_2^d}, r(\mathbf{x}_1) \right) \right) \rho^{m+1} d\mathbf{x}_1 d\mathbf{t}_1 \dots d\mathbf{t}_m. \end{aligned}$$

By the definition of the first intensity measure (2) the last equation is equal to

$$\int_{B(\mathbf{0}, 2)^m} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B \left( \varepsilon \sum_{j=1}^m \frac{\mathbf{t}_j}{\|\mathbf{t}_j\|_2^d}, r(\mathbf{x}) \right)}(\mathbf{x} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}\}}^{(0,1)}(\mathbf{x})) \right] \rho^m d\mathbf{t}_1 \dots d\mathbf{t}_m.$$

Employing further the stationarity of  $\Pi_{\varepsilon}^{(0,1)} \mathcal{P}$ , we get

$$\begin{aligned} A_{m+1} &\leq \int_{B(\mathbf{0}, 2)^m} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right] \rho^m d\mathbf{t}_1 \dots d\mathbf{t}_m \\ &= (2^d \kappa_d \rho)^m E_1. \end{aligned} \tag{49}$$

Second, for  $k \in \{2, \dots, m\}$ , we have

$$A_k \leq \mathbb{E} \left[ \sum_{\substack{\mathbf{x}_1 \in B(\mathbf{0}, R)^c \\ \mathbf{x}_2, \dots, \mathbf{x}_k \in B(\mathbf{x}_1, 2)}}^{\neq} \left( \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))}(\mathbf{x}_1 + \varepsilon F_{\hat{\phi}^{(k)}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^k \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d}) \right. \right. \\ \left. \sum_{\substack{\mathbf{x}_{k+1}, \dots, \mathbf{x}_{m+1} \in B(\mathbf{x}_1, 2)^c}}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{k+1}))}(\mathbf{x}_{k+1} + \varepsilon F_{\hat{\phi}^{(k+1)}}^{(0,1)}(\mathbf{x}_{k+1})) \right. \\ \left. \dots \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{m+1}))}(\mathbf{x}_{m+1} + \varepsilon F_{\hat{\phi}^{(k)} \setminus \{\mathbf{x}_{m+1}\}}^{(0,1)}(\mathbf{x}_{m+1})) \right) \Bigg].$$

Applying the extended Slivnyak-Mecke theorem (4) then, using the independence of  $F_{\hat{\phi}}^{(0,1)}(\mathbf{x})$  and  $F_{\hat{\phi}}^{(0,1)}(\mathbf{y})$  whenever  $\|\mathbf{x} - \mathbf{y}\|_2 > 2$  yield

$$A_k \leq \int_{B(\mathbf{0}, R)^c \times B(\mathbf{x}_1, 2)^{k-1}} \mathbb{E} \left[ \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))} \left( \mathbf{x}_1 + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^k \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d} \right) \right. \\ \sum_{\substack{\mathbf{x}_{k+1}, \dots, \mathbf{x}_{m+1} \in B(\mathbf{x}_1, 2)^c}}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{k+1}))}(\mathbf{x}_{k+1} + \varepsilon F_{\hat{\phi} \setminus \{\mathbf{x}_{k+1}\}}^{(0,1)}(\mathbf{x}_{k+1})) \\ \dots \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{m+1}))}(\mathbf{x}_{m+1} + \varepsilon F_{\hat{\phi} \setminus \{\mathbf{x}_{m+1}\}}^{(0,1)}(\mathbf{x}_{m+1})) \Bigg] \rho^k d\mathbf{x}_k \dots d\mathbf{x}_1 \\ = \int_{B(\mathbf{0}, R)^c \times B(\mathbf{x}_1, 2)^{k-1}} \mathbb{E} \left[ \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))} \left( \mathbf{x}_1 + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^k \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d} \right) \right] \\ \mathbb{E} \left[ \sum_{\substack{\mathbf{x}_{k+1}, \dots, \mathbf{x}_{m+1} \in B(\mathbf{x}_1, R)^c}}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{k+1}))}(\mathbf{x}_{k+1} + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x}_{k+1})) \right. \\ \dots \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_{m+1}))}(\mathbf{x}_{m+1} + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x}_{m+1})) \Bigg] \rho^k d\mathbf{x}_k \dots d\mathbf{x}_1 \\ = E_{m+1-k} \times \\ \int_{B(\mathbf{0}, R)^c \times B(\mathbf{x}_1, 2)^{k-1}} \mathbb{E} \left[ \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))} \left( \mathbf{x}_1 + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^k \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d} \right) \right] \rho^k d\mathbf{x}_k \dots d\mathbf{x}_1.$$

Following the same technique used to bound  $A_{m+1}$  we get

$$A_k \leq E_{m+1-k} \int_{B(\mathbf{0}, 2)^{k-1}} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\varepsilon \sum_{j=1}^{k-1} \frac{\mathbf{t}_j}{\|\mathbf{t}_j\|_2^d}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x})) \right] \rho^{k-1} d\mathbf{t}_1 \dots d\mathbf{t}_{k-1} \\ = E_{m+1-k} \int_{B(\mathbf{0}, 2)^{k-1}} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\hat{\phi}}^{(0,1)}(\mathbf{x})) \right] \rho^{k-1} d\mathbf{t}_1 \dots d\mathbf{t}_{k-1}$$

$$= (2^d \kappa_d \rho)^{k-1} E_1 E_{m+1-k}. \quad (50)$$

Inserting (49), and (50) in (48) we get

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))} (\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^{m+1} \right] &= A_{m+1} + \sum_{k=1}^m d_k E_k + A_k \\ &\leq c_1^m E_1 + \sum_{k=1}^m \left( d_k E_k + c_1^{k-1} E_1 E_{m+1-k} \right), \end{aligned}$$

with  $c_1 = 2^d \kappa_d \rho$ , which concludes the proof.  $\square$

#### A.4. Proof of Theorem 3 (variance reduction)

In this section we prove the main result of the paper, Theorem 3. Our starting point is a lemma from the theory of harmonic functions, which is key in subsequent proofs. Then we state and prove a few technical lemmas on the behavior as  $\varepsilon \rightarrow 0$  of various quantities relevant to the variance of linear statistics under  $\Pi_\varepsilon \mathcal{P}$ . Finally, we put these lemmas in action in the proof of Theorem 3.

**Lemma 3.** *Let  $d > 2$  and  $g \in C^2(\mathbb{R}^d)$  have compact support. For all  $\mathbf{x} \in \mathbb{R}^d$ , we have*

$$g(\mathbf{x}) = \frac{1}{(2-d)d\kappa_d} \int_{\mathbb{R}^d} \Delta g(\mathbf{y}) \frac{1}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}} d\mathbf{y} = \frac{1}{d\kappa_d} \int_{\mathbb{R}^d} \nabla g(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} d\mathbf{y}. \quad (51)$$

The first equality in (51) can be found in (Axler et al., 2001, Chapter 9). We thus only prove the second equality, which results from an appropriate integration by parts.

**Proof of Lemma 3.** First, recall Green's "integration by parts" formula (Axler et al., 2001, Chapter 1)

$$\int_A f(\mathbf{x}) \Delta g(\mathbf{x}) d\mathbf{x} = \int_{\partial A} (f \nabla g \cdot \mathbf{n} + g \nabla f \cdot \mathbf{n}) dS - \int_A \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) d\mathbf{x}, \quad (52)$$

where  $A$  is a bounded subset of  $\mathbb{R}^d$  with smooth boundary,  $f$  and  $g$  are  $C^2$  on a neighborhood of  $\bar{A}$ ,  $\mathbf{n}$  is the outward unit normal vector,  $S$  is the surface-area measure on  $\partial A$ . Let the support of  $g$  be a compact  $K \subset A$ , (52) simplifies and we have

$$\int_A f(\mathbf{x}) \Delta g(\mathbf{x}) d\mathbf{x} = \int_{\partial A} f \nabla g \cdot \mathbf{n} dS - \int_A \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) d\mathbf{x}. \quad (53)$$

Let  $\mathbf{x} \in \mathbb{R}^d$ , and let  $R > 0$  be large enough such that  $\mathbf{x}$  and  $K$  are contained in  $B(\mathbf{0}, R)$ . Then using the dominated convergence we get

$$\frac{1}{(2-d)d\kappa_d} \int_{\mathbb{R}^d} \frac{\Delta g(\mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}} d\mathbf{y} = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2-d)d\kappa_d} \int_{B(\mathbf{0}, R) \setminus B(\mathbf{x}, \varepsilon)} \frac{\Delta g(\mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}} d\mathbf{y}. \quad (54)$$

Now, using (53) with  $A = B(\mathbf{0}, R) \setminus B(\mathbf{x}, \varepsilon)$  and  $f : \mathbf{y} \mapsto \frac{1}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}}$ , it comes

$$\int_A \frac{\Delta g(\mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}} d\mathbf{y} = \int_{\partial B(\mathbf{x}, \varepsilon)} \frac{1}{\|\mathbf{x} - \mathbf{y}\|_2^{d-2}} \nabla g(\mathbf{y}) \cdot \mathbf{n} dS + (2-d) \int_A \nabla g(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} d\mathbf{y}$$

$$\begin{aligned}
&= \frac{1}{\varepsilon^{d-2}} \int_{\partial B(\mathbf{x}, \varepsilon)} \nabla g(\mathbf{y}) \cdot \mathbf{n} \, dS + (2-d) \int_A \nabla g(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} \, d\mathbf{y} \\
&= \varepsilon \int_{\partial B(\mathbf{x}, 1)} \nabla g(\varepsilon \mathbf{y}) \cdot \mathbf{n} \, dS + (2-d) \int_A \nabla g(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} \, d\mathbf{y}.
\end{aligned}$$

Plugging into (54) evaluating the limit, we obtain the desired limit.  $\square$

We use Lemma 3 to prove the following result, which takes us closer to controlling the variance of linear statistics under the repelled Poisson point process.

**Lemma 4.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $f \in C^2(\mathbb{R}^d)$  of compact support  $K \subset B(\mathbf{0}, R)$  with  $R > 0$ . For any  $R' \geq R$  we have

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right] = -2d\kappa_d \rho^2 I_K(f^2), \quad (55)$$

where  $I_K$  is defined in (7).

The proof of this Lemma relies on the dominated convergence theorem, a first-order Taylor expansion of  $f$ , and Lemma 3.

Note that, the two sums in equation (55) are actually over the same subsets of  $\mathcal{P}$  since  $f$  has support  $K \subset B(\mathbf{0}, R')$  and this equation represents a first-order derivative. As highlighted in Remark 3, pursuing the calculation of the second-order derivative through a second-order Taylor expansion poses difficulties since  $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{x})\|_2^2] = \infty$ .

**Proof of Lemma 4.** Let  $R' \geq R$ . For  $\varepsilon \in (-1, 1)$ , define

$$X_{\varepsilon} \triangleq \varepsilon^{-1} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right].$$

We need to show that  $\mathbb{E}[X_{\varepsilon}] \rightarrow -2d\kappa_d \rho^2 I_K(f^2)$  as  $\varepsilon \rightarrow 0$ . The first step is to find a random variable  $Y \in L^1$  such that  $|X_{\varepsilon}| \leq Y$  for all  $\varepsilon \in (-1, 1)$  so that we can later apply the dominated convergence theorem to switch the limit and the expectation.

Recall that every  $C^1$  function compactly supported is Lipschitz, so  $f$  is Lipschitz. Let  $L > 0$  be the Lipschitz constant of  $f$ , it comes

$$\begin{aligned}
|X_{\varepsilon}| &\leq \frac{1}{|\varepsilon|} \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} |f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - f(\mathbf{x})| \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} |f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) + f(\mathbf{x})| \\
&\leq \frac{1}{|\varepsilon|} \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} L \|\varepsilon F_{\mathcal{P}}(\mathbf{x})\|_2 \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} 2\|f\|_{\infty} \\
&= 2L\|f\|_{\infty} \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} \|F_{\mathcal{P}}(\mathbf{x})\|_2 \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x}).
\end{aligned}$$

Let  $Y = \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} \|F_{\mathcal{P}}(\mathbf{x})\|_2 \sum_{\mathbf{x} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x})$ . We claim that the nonnegative random variable  $Y$  is in  $L^1$ . To see why, note that the extended Slivnyak-Mecke theorem (4) yields

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} \|F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x})\|_2 + \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} \|F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x})\|_2 \sum_{\mathbf{y} \in \mathcal{P} \setminus \{\mathbf{x}\}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right] \\ &= \int_{B(\mathbf{0}, R')} \mathbb{E} [\|F_{\mathcal{P}}(\mathbf{x})\|_2] \rho d\mathbf{x} + \int_{B(\mathbf{0}, R')} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2 \sum_{\mathbf{y} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right] \rho d\mathbf{x} \\ &\leq \rho \kappa_d(R')^d \left( \mathbb{E} [\|F_{\mathcal{P}}(\mathbf{0})\|_2] + \mathbb{E} [\|F_{\mathcal{P}}(\mathbf{0})\|_2^a]^{1/a} \left[ \mathbb{E} \left( \sum_{\mathbf{y} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right)^b \right]^{1/b} \right), \end{aligned}$$

where the last inequality results from Hölder's inequality with  $1 < a < d/(d-1)$  and  $b = a/(a-1)$ . As indicated in Section 3.4,  $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^\gamma] < \infty$  for any  $\gamma < d/(d-1)$ . Therefore  $\mathbb{E}[Y] < \infty$  and our claim is true.

Applying the dominated convergence theorem, and a first-order Taylor expansion, we obtain

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \mathbb{E}[X_\varepsilon] &= \mathbb{E} \left[ \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left( \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right) \right] \\ &= \mathbb{E} \left[ \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left( \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x}) + \varepsilon \nabla f(\mathbf{x}) \cdot F_{\mathcal{P}}(\mathbf{x}) + o_{\varepsilon \rightarrow 0}(\varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right) \right] \\ &= 2\mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)} f(\mathbf{x}) \sum_{\mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R)} \nabla f(\mathbf{y}) \cdot F_{\mathcal{P}}(\mathbf{y}) \right]. \end{aligned}$$

Expanding the sum, it comes

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \mathbb{E}[X_\varepsilon] &= 2\mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)} f(\mathbf{x}) \nabla f(\mathbf{x}) \cdot F_{\mathcal{P}}(\mathbf{x}) + \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R)}^{\neq} f(\mathbf{x}) \nabla f(\mathbf{y}) \cdot F_{\mathcal{P}}(\mathbf{y}) \right] \\ &= 2\mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R)} f(\mathbf{x}) \nabla f(\mathbf{x}) \cdot F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x}) \right] - 2\mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R)}^{\neq} f(\mathbf{x}) \nabla f(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} \right] \\ &\quad + 2\mathbb{E} \left[ \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R)}^{\neq} f(\mathbf{x}) \nabla f(\mathbf{y}) \cdot F_{\mathcal{P} \setminus \{\mathbf{y}, \mathbf{x}\}}(\mathbf{y}) \right]. \end{aligned} \tag{56}$$

Using Slivnyak-Mecke (4) and Equation (3), we obtain

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}[X_\varepsilon] = 2 \int_{B(\mathbf{0}, R)} \mathbb{E} [f(\mathbf{x}) \nabla f(\mathbf{x}) \cdot F_{\mathcal{P}}(\mathbf{x})] \rho d\mathbf{x} - 2 \int_{B(\mathbf{0}, R) \times B(\mathbf{0}, R)} f(\mathbf{x}) \nabla f(\mathbf{y}) \cdot \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|_2^d} \rho^2 d\mathbf{x} d\mathbf{y}$$

$$+ 2 \int_{B(\mathbf{0}, R) \times B(\mathbf{0}, R)} \mathbb{E} [f(\mathbf{x}) \nabla f(\mathbf{y}) \cdot F_{\mathcal{P}}(\mathbf{y})] \rho^2 d\mathbf{x} d\mathbf{y}. \quad (57)$$

But the first and last terms in (57) are zero due to  $F_{\mathcal{P}}$  being a centered process, and the second term is  $-2d\kappa_d\rho^2 I_K(f^2)$  by Lemma 3.  $\square$

Remarkably, the proof of Lemma 4 remains valid even if  $F_{\mathcal{P}}$  is replaced by its truncated version  $F_{\mathcal{P}}^{(0,p)}$ , where  $p \geq 2R$ . In particular, the choice of  $p \geq 2R$  is crucial to ensure that Equation (56) remains valid.

**Lemma 5.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $f \in C^2(\mathbb{R}^d)$  of compact support  $K \subset B(\mathbf{0}, R)$  with  $R > 0$ . Let further  $0 < \beta < \frac{1}{2(d-1)^2}$ . For any  $R' > (2R+2)^{1/\beta}$

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 \right] = 0.$$

The proof is based on Corollary 2 and 3.

**Proof of Lemma 5.** Let  $R' > (2R+2)^{1/\beta}$ ,  $\varepsilon \in (-1, 1)$ , and

$$Y_{\varepsilon} \triangleq \frac{1}{\varepsilon} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 \right].$$

Setting  $r(\mathbf{x}) = \|\mathbf{x}\|_2^{\beta}$  with  $0 < \beta < \frac{1}{2(d-1)^2}$ , we use the same splitting technique as in (28), and write

$$\begin{aligned} |Y_{\varepsilon}| &\leq \frac{\|f\|_{\infty}}{|\varepsilon|} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 \right] \\ &\leq \frac{\|f\|_{\infty}}{|\varepsilon|} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) + \mathbb{1}_{\{\|\varepsilon F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > r(\mathbf{x}) - R\}}(\mathbf{x}) \right)^2 \right] \\ &\leq \frac{2\|f\|_{\infty}}{|\varepsilon|} \mathbb{E} \left[ \underbrace{\left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x})) \right)^2}_{\triangleq E_{\text{int}}(\varepsilon)} \right] \\ &\quad + \frac{2\|f\|_{\infty}}{|\varepsilon|} \mathbb{E} \left[ \underbrace{\left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \mathbb{1}_{\{\|\varepsilon F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > r(\mathbf{x}) - R\}}(\mathbf{x}) \right)^2}_{\triangleq E_{\text{ext}}(\varepsilon)} \right]. \end{aligned}$$

By Corollary 3 with  $m = 2$ , there exists  $a_1 > 0$  such that for any  $0 < \gamma < 1/(d-1)$

$$E_{\text{int}}(\varepsilon) \leq a_1 E_1 + E_1^2 \leq a_1 |\varepsilon|^{1+\gamma} + 2|\varepsilon|^{2+2\gamma} + h_1(\varepsilon) + 2h_1(\varepsilon)^2,$$

where  $E_1$  is defined by Equation (44), for  $k = 1$  and

$$\begin{aligned} h_1(\varepsilon) &= \int_{B(\mathbf{0}, R')^c} a_2 \exp \left( -a_3 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{|\varepsilon|} \log \left( a_4 \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{|\varepsilon|} \right) \right) \rho \, d\mathbf{x} \\ &= \int_{B(\mathbf{0}, R')^c} a_2 \exp \left( -a_3 \frac{\|\mathbf{x}\|_2^\beta}{|\varepsilon|} \left( \|\mathbf{x}\|_2^{1-\beta} - 1 \right) \log \left( a_4 \frac{\|\mathbf{x}\|_2^\beta}{|\varepsilon|} \left( \|\mathbf{x}\|_2^{1-\beta} - 1 \right) \right) \right) \rho \, d\mathbf{x} \end{aligned}$$

for some positive constants  $a_2, \dots, a_4$ . But

$$\begin{aligned} h_1(\varepsilon) &\leq \int_{B(\mathbf{0}, R')^c} a_2 \exp \left( -a_3 \frac{\|\mathbf{x}\|_2^\beta}{|\varepsilon|} \left( R'^{1-\beta} - 1 \right) \log \left( a_4 \frac{\|\mathbf{x}\|_2^\beta}{|\varepsilon|} \left( R'^{1-\beta} - 1 \right) \right) \right) \rho \, d\mathbf{x} \\ &= |\varepsilon|^{d/\beta} \int_{B(\mathbf{0}, R'|\varepsilon|^{-1/\beta})^c} a_2 \exp \left( -a_3 \|\mathbf{y}\|_2^\beta \left( R'^{1-\beta} - 1 \right) \log \left( a_4 \|\mathbf{y}\|_2^\beta \left( R'^{1-\beta} - 1 \right) \right) \right) \rho \, d\mathbf{y} \\ &\leq |\varepsilon|^{d/\beta} \underbrace{\int_{B(\mathbf{0}, 1)^c} a_2 \exp \left( -a_3 \|\mathbf{y}\|_2^\beta \left( R'^{1-\beta} - 1 \right) \log \left( a_4 \|\mathbf{y}\|_2^\beta \left( R'^{1-\beta} - 1 \right) \right) \right) \rho \, d\mathbf{y}}_{\triangleq c} \end{aligned}$$

where in the second line we used the change of variable  $\mathbf{y} = |\varepsilon|^{-1/\beta} \mathbf{x}$ , and in the last line we used that  $R' > |\varepsilon|^{1/\beta}$ . Thus

$$0 \leq h_1(\varepsilon) \leq c|\varepsilon|^{d/\beta}.$$

As  $d/\beta > 1$ , we have  $h_1(\varepsilon)/|\varepsilon|$  going to zero as  $\varepsilon$  approaches zero, so that  $E_{\text{int}}(\varepsilon) = o(\varepsilon)$ .

It remains to show that  $E_{\text{ext}}(\varepsilon) = o(\varepsilon)$ . By Corollary 2 with  $m = 2$  there exists  $a_5, \dots, a_8 > 0$  such that

$$E_{\text{ext}}(\varepsilon) \leq a_5 \left( h_2(\varepsilon) + h_2(\varepsilon)^2 \right),$$

with

$$h_2(\varepsilon) = \int_{B(\mathbf{0}, R')^c} a_6 \exp \left( -a_7 \frac{r(\mathbf{x}) - (R+1)}{|\varepsilon|} \log \left( a_8 \frac{r(\mathbf{x}) - (R+1)}{|\varepsilon|} \right) \right) \rho \, d\mathbf{x}.$$

In particular

$$\begin{aligned} h_2(\varepsilon) &= \int_{B(\mathbf{0}, R')^c} a_6 \exp \left( -a_7 \frac{\|\mathbf{x}\|^\beta - (R+1)}{|\varepsilon|} \log \left( a_8 \frac{\|\mathbf{x}\|^\beta - (R+1)}{|\varepsilon|} \right) \right) \rho \, d\mathbf{x} \\ &\leq \int_{B(\mathbf{0}, R')^c} a_6 \exp \left( -a_7 \frac{\|\mathbf{x}\|^\beta}{2|\varepsilon|} \log \left( a_8 \frac{\|\mathbf{x}\|^\beta}{2|\varepsilon|} \right) \right) \rho \, d\mathbf{x}, \end{aligned}$$

where in the last line we used that  $R' > (2R+2)^{1/\beta}$ . Following the same method used earlier to bound  $h_1(\varepsilon)$ , we can show that  $h_2(\varepsilon) \leq C|\varepsilon|^{d/\beta}$  for some constant  $C$ . This implies that  $h_2(\varepsilon) = o(\varepsilon)$  and  $E_{\text{ext}}(\varepsilon) = o(\varepsilon)$ . We conclude that  $Y_\varepsilon = o_{\varepsilon \rightarrow 0}(1)$ , which ends the proof.  $\square$

Again, the proof's validity is unaffected by replacing  $F_\rho$  with its truncated counterpart  $F_\rho^{(0,p)}$ , where  $p > 0$ . The upcoming lemma is the final tool required to demonstrate Theorem 3.

**Lemma 6.** Consider a homogeneous Poisson point process  $\mathcal{P} \subset \mathbb{R}^d$  of intensity  $\rho$ , with  $d \geq 3$ . Let  $f$  be a  $C^2(\mathbb{R}^d)$  function of compact support  $K \subset B(\mathbf{0}, R)$  with  $R > 0$ . For any  $R' \geq R$  we have

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - \sum_{\mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y})) \right] = 0.$$

**Proof of Lemma 6.** Let  $R' \geq R$ ,  $\varepsilon \in (-1, 1)$ , and denote

$$Z_{\varepsilon} \triangleq \frac{1}{\varepsilon} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - \sum_{\mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{y} + \varepsilon F_{\mathcal{P}}(\mathbf{y})) \right].$$

We show that  $Z_{\varepsilon} = o_{\varepsilon \rightarrow 0}(1)$ . Since  $f$  is  $C^2$  with compact support, it is Lipschitz and we denote its Lipschitz constant by  $L \geq 0$ . For  $\mathbf{x} \in B^c(\mathbf{0}, R')$ ,  $f(\mathbf{x}) = 0$  so

$$|f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}))| = |f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - f(\mathbf{x})| \leq L \|\varepsilon F_{\mathcal{P}}(\mathbf{x})\|_2 \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})).$$

Hence

$$\begin{aligned} |Z_{\varepsilon}| &\leq \frac{1}{|\varepsilon|} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} L \|\varepsilon F_{\mathcal{P}}(\mathbf{x})\|_2 \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - \sum_{\mathbf{y} \in \mathcal{P} \cap B(\mathbf{0}, R')} \|f\|_{\infty} \right] \\ &\leq L \|f\|_{\infty} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} \left( \|F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x})\|_2 \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P} \setminus \{\mathbf{x}\}}(\mathbf{x})) - \sum_{\mathbf{y} \in \mathcal{P} \setminus \{\mathbf{x}\}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right) \right] \\ &= L \|f\|_{\infty} \int_{B(\mathbf{0}, R')^c} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2 \mathbb{1}_{B(\mathbf{0}, R)}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) - \sum_{\mathbf{y} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right] \rho d\mathbf{x} \end{aligned}$$

The last equality was obtained using the extended Slivnyak-Mecke theorem (4). Next, we employ Hölder's inequality with  $1 < p < \frac{d}{d-1}$  and  $q = p/(p-1)$  to obtain

$$\begin{aligned} |Z_{\varepsilon}| &\leq L \|f\|_{\infty} \int_{B(\mathbf{0}, R')^c} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right]^{1/p} \mathbb{E} \left[ \left( \sum_{\mathbf{y} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right)^q \right]^{1/q} \rho d\mathbf{x} \\ &= \underbrace{\rho L \|f\|_{\infty} \mathbb{E} \left[ \left( \sum_{\mathbf{y} \in \mathcal{P}} \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{y}) \right)^q \right]^{1/q}}_C \int_{B(\mathbf{0}, R')^c} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right]^{1/p} d\mathbf{x} \end{aligned}$$

When attempting to compute the limit of the last equation as  $\varepsilon$  approaches zero, it becomes challenging to interchange the limit and integral using the dominated convergence theorem. This difficulty arises because  $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{x})\|_2^p]^{1/p}$  is not integrable over  $B(\mathbf{0}, R')^c$ . To address this, we will handle the cases of  $\|F_{\mathcal{P}}(\mathbf{x})\|_2 > 1$  and  $\|F_{\mathcal{P}}(\mathbf{x})\|_2 < 1$  separately. By doing so, we can manage the computation differently for each case, allowing us to remove the exponent  $\frac{1}{p}$  in the first case and proceed with the calculations

accordingly. For  $\mathbf{x} \in B(\mathbf{0}, R')^c$ , we have

$$\begin{aligned}
& \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right]^{1/p} \\
&= \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \left( \mathbb{1}_{\{\|F_{\mathcal{P}}(\mathbf{x})\|_2 < 1\}}(\mathbf{x}) + \mathbb{1}_{\{\|F_{\mathcal{P}}(\mathbf{x})\|_2 \geq 1\}}(\mathbf{x}) \right) \right]^{1/p} \\
&\leq \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \mathbb{1}_{\{\|F_{\mathcal{P}}(\mathbf{x})\|_2 < 1\}}(\mathbf{x}) \right]^{1/p} \\
&\quad + \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \mathbb{1}_{\{\|F_{\mathcal{P}}(\mathbf{x})\|_2 \geq 1\}}(\mathbf{x}) \right]^{1/p} \\
&\leq \mathbb{E} \left[ \mathbb{1}_{B(\mathbf{0}, R' + \varepsilon)}(\mathbf{x}) \right]^{1/p} + \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right] \\
&= \mathbb{1}_{B(\mathbf{0}, R' + \varepsilon)}(\mathbf{x}) + \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right].
\end{aligned}$$

Thus

$$\begin{aligned}
|Z_{\varepsilon}| &\leq C \int_{B(\mathbf{0}, R')^c} \mathbb{1}_{B(\mathbf{0}, R' + \varepsilon)}(\mathbf{x}) + \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{x})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right] d\mathbf{x} \\
&= C |B(\mathbf{0}, R')^c \cap B(\mathbf{0}, R' + \varepsilon)| + C \int_{B(\mathbf{0}, R')^c} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{0})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{0})) \right] d\mathbf{x}.
\end{aligned}$$

The first term in the last inequality goes to zero as  $\varepsilon \rightarrow 0$ . It remains to show that the second term also goes to zero as  $\varepsilon \rightarrow 0$ . Denote by  $f_{F_{\mathcal{P}}(\mathbf{0})}$  the density function of  $F_{\mathcal{P}}(\mathbf{0})$  and observe that

$$\begin{aligned}
\lim_{\varepsilon \rightarrow 0} |Z_{\varepsilon}| &= \lim_{\varepsilon \rightarrow 0} C \int_{B(\mathbf{0}, R')^c} \mathbb{E} \left[ \|F_{\mathcal{P}}(\mathbf{0})\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{0})) \right] d\mathbf{x} \\
&= \lim_{\varepsilon \rightarrow 0} C \int_{B(\mathbf{0}, R')^c} \int_{\mathbb{R}^d} \|\mathbf{u}\|_2^p \mathbb{1}_{B(\mathbf{0}, R')}(\mathbf{x} + \varepsilon \mathbf{u}) f_{F_{\mathcal{P}}(\mathbf{0})}(\mathbf{u}) d\mathbf{u} d\mathbf{x} \\
&= \lim_{\varepsilon \rightarrow 0} C \int_{\mathbb{R}^d} \underbrace{\|\mathbf{u}\|_2^p f_{F_{\mathcal{P}}(\mathbf{0})}(\mathbf{u}) |B(\mathbf{0}, R')^c \cap B(\varepsilon \mathbf{u}, R')|}_{A_{\varepsilon}(\mathbf{u})} d\mathbf{u}.
\end{aligned}$$

Now, for any  $\mathbf{u} \in \mathbb{R}^d$ ,  $A_{\varepsilon}(\mathbf{u}) \xrightarrow{\varepsilon \rightarrow 0} 0$ . It is enough to show that we can use the dominated convergence theorem to conclude. Actually

$$|A_{\varepsilon}(\mathbf{u})| \leq \kappa_d R'^d \|\mathbf{u}\|_2^p f_{F_{\mathcal{P}}(\mathbf{0})}(\mathbf{u}) \triangleq Y(\mathbf{u})$$

and  $\int_{\mathbb{R}^d} Y(\mathbf{u}) d\mathbf{u} = \kappa_d R'^d \mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^p]$ , which is finite since  $p < d/(d-1)$ ; see Section 3.4. The dominated convergence theorem thus concludes the proof.  $\square$

The proof's validity is unaffected by replacing  $F_{\mathcal{P}}$  with its truncated counterpart  $F_{\mathcal{P}}^{(0,p)}$ , where  $p > 0$ .

We end this section with the proof of Theorem 3. We will use Lemmas 4, 5 and 6.

**Proof of Theorem 3.** Consider a function  $f \in C^2(\mathbb{R}^d)$  of compact support  $K \subset B(0, R)$  for some  $R > 0$ . We start by proving that

$$\left[ \frac{\partial \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f) \right]}{\partial \varepsilon} \right]_{\varepsilon=0} = -2d\kappa_d I_K(f^2).$$

First, Proposition 2 implies the existence of  $\mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right]$  for any  $\varepsilon \in (-1, 1)$ . Now, fix

$$0 < \beta < \frac{1}{2(d-1)^2},$$

and let  $R' \geq (2R+2)^{1/\beta}$ . As  $\mathbb{E}[\widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f)] = \mathbb{E}[\widehat{I}_{\mathcal{P} \cap K}(f)]$  we have

$$\begin{aligned} & \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right] - \mathbb{V}\text{ar} \left[ \widehat{I}_{\mathcal{P} \cap K}(f) \right] \\ &= \mathbb{E} \left[ \left( \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right)^2 \right] - \mathbb{E} \left[ \left( \widehat{I}_{\mathcal{P} \cap K}(f) \right)^2 \right] \\ &= \rho^{-2} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right] \\ &= \rho^{-2} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right] + \rho^{-2} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 \right] \\ &\quad + 2\rho^{-2} \mathbb{E} \left[ \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')^c} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right] \end{aligned}$$

Using Lemmas 4, 5, and 6 we get

$$\begin{aligned} \left[ \frac{\partial \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right]}{\partial \varepsilon} \right]_{\varepsilon=0} &= \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \rho^{-2} \mathbb{E} \left[ \left( \sum_{\mathbf{x} \in \mathcal{P} \cap B(\mathbf{0}, R')} f(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x})) \right)^2 - \left( \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right)^2 \right] \\ &= -2d\kappa_d I_K(f^2) = -2d\kappa_d \rho \mathbb{V}\text{ar}[\widehat{I}_{\mathcal{P}}(f)]. \end{aligned}$$

Finally, the Taylor expansion of  $\mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right]$  at  $\varepsilon = 0$  gives

$$\begin{aligned} \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right] &= \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_0 \mathcal{P} \cap K}(f) \right] + \varepsilon \left[ \frac{\partial \mathbb{V}\text{ar} \left[ \widehat{I}_{\Pi_\varepsilon \mathcal{P} \cap K}(f) \right]}{\partial \varepsilon} \right]_{\varepsilon=0} + O(\varepsilon^2) \\ &= \mathbb{V}\text{ar} \left[ \widehat{I}_{\mathcal{P}}(f) \right] (1 - 2d\kappa_d \rho \varepsilon) + O(\varepsilon^2) \end{aligned}$$

which ends the proof.  $\square$

**Remark 10.** Substituting  $F_{\mathcal{P}}$  with its truncated version  $F_{\mathcal{P}}^{(0,p)}$ , where  $p \geq \text{diam}(K)$ , preserves the validity of the proof of Theorem 3. However, the same cannot be said for  $F_{\mathcal{P}}^{(q,p)}$  when  $q > 0$ , as it leads to the breakdown of the proof of Lemma 4. This highlights the possibility that the variance reduction may be attributed to the singularity of  $F_{\mathcal{P}}$  at the points of  $\mathcal{P}$ .

## Appendix B: Additional experiments

### B.1. Further analysis of Section 4.2

	SW	CI	SW	CI	SW	CI	SW	CI	SW	CI
$f_1$	stat=0.9, p=0.04	[-0.58, -0.46]	stat=0.96, p=0.48	[-0.69, -0.51]	stat=0.95, p=0.34	[-0.66, -0.54]	stat=0.95, p=0.39	[-0.65, -0.53]	stat=0.91, p=0.06	[-0.52, -0.28]
$f_2$	stat=0.98, p=0.88	[-0.63, -0.51]	stat=0.94, p=0.26	[-0.65, -0.53]	stat=0.93, p=0.17	[-0.61, -0.49]	stat=0.96, p=0.61	[-0.63, -0.51]	stat=0.94, p=0.30	[-0.60, -0.48]
$f_3$	stat=0.97, p=0.65	[-0.57, -0.45]	stat=0.98, p=0.92	[-0.67, -0.55]	stat=0.97, p=0.84	[-0.69, -0.57]	stat=0.93, p=0.15	[-0.64, -0.52]	stat=0.99, p=0.99	[-0.60, -0.48]
	d=2		d=3		d=4		d=5		d=7	

**Table 1.** Shapiro Wilk (SW) test for the residual of the OLS of the estimated log-standard deviation of  $\hat{I}_{\text{MCRB}}$  over  $\log(N)$  and the confidence interval (CI) with 99.7% confidence level of the slopes for  $f_1$ ,  $f_2$  and  $f_3$  when  $d \in \{2, 3, 4, 5, 7\}$ .

We used the Shapiro-Wilk test (Shapiro and Wilk, 1965) to assess whether the residuals of each OLS of  $\log(\sigma)$  of  $\hat{I}_{\text{MCRB}}$  over  $\log(N)$  shown in Figure 4 are Normally distributed, and computed confidence intervals for the slope corresponding to  $\hat{I}_{\text{MCRB}}$ . The  $p$ -values of the Shapiro-Wilk test corresponding to  $\hat{I}_{\text{MCRB}}$  are summarized in Table 1. The null hypothesis of the test is that the residuals are normally distributed. As the statistics are close to 1 and the  $p$ -values are large (typically larger than 0.01), the distributions of the residuals are not significantly different from a normal distribution at a 99% significance level. We have also verified that the Quantile-Quantile plots of the residuals are compatible with the result of Shapiro-Wilk’s test; for a comparison between Normality tests see (Mohd Razali and Yap, 2011). Hence, we can construct confidence intervals for the slopes using their estimated standard deviations. Table 1 shows the 99.7% confidence intervals (corresponding to three standard deviations) of the slopes. For  $d \in \{3, 4, 5\}$  the confidence intervals suggest that the variance of  $\hat{I}_{\text{MCRB}}$  may decrease slightly faster than the usual Monte Carlo rate.

Finally, to account for the slight possible bias of the estimator  $\hat{I}_{\text{MCRB}}$ , we conclude this section by presenting the errors obtained in the experiment of 4. Figure 12 displays the box plots of the error of the estimators  $\hat{I}_{\text{MC}}$ ,  $\hat{I}_{\text{MCRB}}$ ,  $\hat{I}_{\text{MCCV}}$ , and  $\hat{I}_{\text{RQMC}}$ , for the functions  $f_2$  and  $f_3$  for  $d$  in  $\{2, 3, 4, 5, 7\}$  and for  $\hat{I}_{\text{MCDPP}}$  when  $d$  in  $\{2, 3\}$ . There is no clear evidence in Figure 12 to suggest that  $\hat{I}_{\text{MCRB}}$  exhibits any notable bias.

### B.2. Further analysis of Section 6

Figure 13 presents an extension of the iterative repulsion experiments discussed in Section 5.1. In this extended analysis, we explore a broader range of stopping times, specifically considering  $t \in \{30, 80, 130, 180, 200\}$ . To set up the experiment, we consider a sample from a PPP  $\mathcal{P}$  of unit intensity. Subsequently, we perform an iterative application of the repulsion operator  $\Pi_{\varepsilon} \mathcal{P}$  to the original point process  $\mathcal{P}$ , yielding the resulting point processes  $\Pi_{\varepsilon,t} \mathcal{P}$  (22) for  $t \in \{30, 80, 130, 180, 200\}$ . In

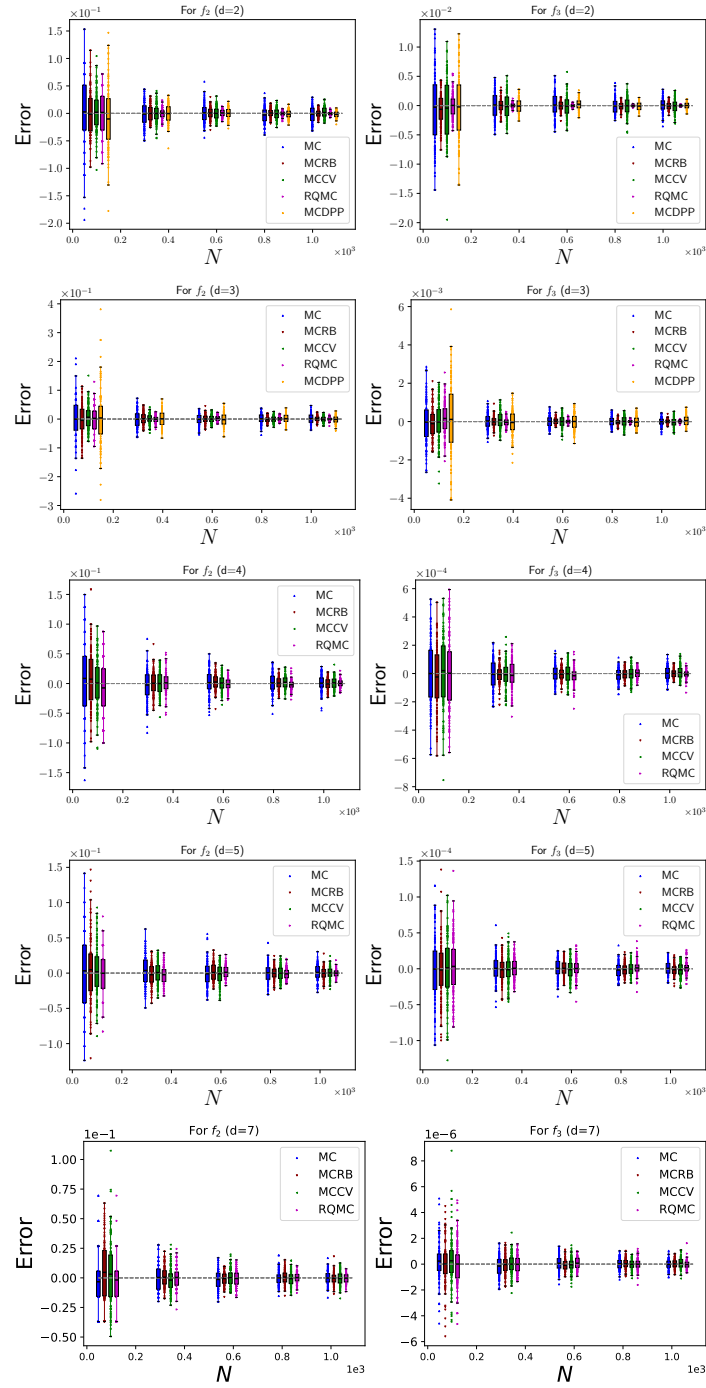
Figure 12: Estimated error of various Monte Carlo methods for  $f_2$ , and  $f_3$  and  $d \in \{2, 3, 4, 5, 7\}$ .

Figure 13, we present the outcomes of this iterative repulsion process. The left panels of the figure showcase the  $t$ -th iterates  $\Pi_{\varepsilon,t}\mathcal{P}$ , specifically constrained to the observation window, where we set  $\varepsilon = -\varepsilon_0$  (12). On the other hand, the right panels display the results obtained with  $\varepsilon = \varepsilon_0$ .

