Point processes for numerical integration

Diala Hawat

UNIVERSITÉ DE LILLE

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LILLE

dans la spécialité

AUTOMATIQUE, GÉNIE INFORMATIQUE, TRAITEMENT DU SIGNAL ET DES IMAGES

par

DIALA HAWAT

POINT PROCESSES FOR NUMERICAL INTEGRATION

Processus ponctuels pour l'intégration numérique

Thèse soutenue le 27 Novembre 2023 devant le jury composé de :

Président	David Dereudre	Professeur, Université de Lille
Rapporteurs	Jean-François Coeurjolly	Professeur, Université Grenoble-Alpes
	Marie-Colette van Lieshout	Professeur, University Twente
Examinateurs	Emilie Lebarbier	Professeur, Université Paris-Nanterre
	François Portier	Maître de conférences, ENSAI
Directeurs	Rémi Bardenet	Chargé de Recherche CNRS, Université de Lille
de thèse	Raphaël Lachièze-Rey	Maître de conférences, Université Paris Cité

Thèse préparée dans les laboratoires:
CRIStAL, Villeneuve d'Ascq, France
MAP5, Paris, France

There is a view in some philosophical circles that anything that can be understood by people who have not studied philosophy is not profound enough to be worth saying. To the contrary, I suspect that whatever cannot be said clearly is probably not being thought clearly either.

— *Attributed to* RICHARD P. FEYNMAN, *"Six Easy Pieces: Essentials of Physics Explained by Its Most Brilliant Teacher".*

Acknowledgment

A challenging part of composing this manuscript was crafting this acknowledgments section. I find myself with a multitude of individuals to express gratitude to, and as time passes, my recollections only seem to expand. Nevertheless, I will strive to acknowledge most of you here, and I hope to personally convey my thanks to the others.

First and foremost, I am deeply indebted to my advisors Raphaël and Rémi. Raphaël, your unwavering belief in my abilities, which took root during my M2 internship, has been instrumental. Thank you for your support, your patient explanations down to the finest details, and your steadfast assistance, especially when I felt on the verge of surrendering to a challenging proof. Rémi, your encouragement has been an unwavering driving force, compelling me to continually push my boundaries and deepen my knowledge. Your dedication to precision, clarity, and integrity has been a guiding influence. This work would not have been possible without both of you.

A special acknowledgment goes to Guillaume, a remarkable collaborator and, above all, a good friend. Your assistance in guiding me to gain good programming habits has been invaluable.

To the present and past members of the SIGMA team at the CRIStAL laboratory, especially, Martin, Michaël, Juan, Xiaoyi, and Arnaud to those from the MAP5 laboratory, particularly, Sonia, Zoé, Ivan, Ousmane, and Laurent my thoughts brim with camaraderie and appreciation. I also extend special gratitude to the "processus ponctuels" team at the Painlevé laboratory for warmly welcoming me into their group. Your friendly and intellectually stimulating atmosphere has been essential for my progress.

I also thank the reviewers and jury members who graciously accepted the responsibility of participating in the culmination of this thesis. Additionally, I am indebted to the efficient service missions and secretarial teams, especially Sonia Waskowiak, Marie-Hélène Gbaguid, Christine Yvos, and Sandrine Adiasse.

Across the span of three years devoted to this thesis, numerous individuals and events have directly and indirectly contributed to my work. For instance, the summer school "Statistics for Point Patterns in Space and Beyond", hosted in Aalborg, Denmark, has been a good experience. Many of the participants and instructors have become enduring colleagues, meeting at conferences to share insights. To all the instructors of this summer school, notably Christophe A.N. Biscio and Rasmus Waagepetersen, I extend my appreciation.

I would also like to express my gratitude to all the researchers and teachers who have had a significant impact on my journey, in particular, Aline Hosry, Manjunath Krishnapur, Simon Corste, David Dereudre, Jean-François Coeurjolly, and Michael Andreas Klatt.

To all my friends, both near and far, particularly, my dearest friend Christelle, and my childhood friends Tatiana and Roland thank you for your support despite the distances that separate us.

To my close friends Jezel and Mariem, my cousin Vincent, my brother Challita, and my Mehdi thank you for your steadfast support and encouragement throughout the preparation for the defense and on this special day. Your presence by my side means a lot to me.

A heartfelt thank goes to my precious Mehdi, a constant pillar of support who never ceased to motivate and believe in me. Throughout this journey, you were the rock I leaned on. I'm glad to have you by my side.

Lastly, I reserve my deepest gratitude for those nearest and dearest to my heart, my mother Viviane, my father Kobrianous, my beloved sister Bertha, my siblings and support Challita and Roberto, and my dear uncle Alfred. I dedicate this thesis to you.

و أخص بالشكر الأغلى على قلبي، أمي فيفيان، أبي قبريانوس، أختي العزيزة الغالية بيرتا،

أخواي وسندي، شليطا و روبارتو وعمي العزيز ألفريد. أهدي هذه الأطروحة لكم.

Paris, 27/10/2023,
Diala Hawat.

Abstract

The Monte Carlo method estimates an integral using pointwise evaluations of the integrand at some points called nodes, which can be chosen as the points of a point process. While crude Monte Carlo relies on a homogeneous Poisson point process (PPP), some more regularly spread point processes yield Monte Carlo methods with faster-decaying variance. In this thesis, we study two families of regular point processes that are potential candidate nodes to speed up the convergence of crude Monte Carlo. The first one is the family of hyperuniform point processes (HUPPs). A HUPP is characterized by the variance of the number of points in a large window scaling slower than the volume of that window. In particular, a HUPP yields a Monte Carlo estimator of volumes with a faster decaying variance than under the PPP. Unfortunately, proving that a point process is hyperuniform is usually difficult. Aiming to provide statistical tools for identifying HUPPs we examine a spectral measure called the structure factor whose decay around zero provides a diagnostic of hyperuniformity. We provide a survey and derivation of natural estimators of the structure factor and contribute an asymptotically valid statistical test of hyperuniformity. We further provide a Python toolbox containing all the estimators and tools that we discuss. The second family of point processes under consideration pertains to repelled point processes which we construct using a so-called repulsion operator. The repulsion operator reduces clustering in a configuration of points by slightly pushing the points away from each other. Our main theoretical result is that applying the repulsion operator to a PPP yields an unbiased Monte Carlo method with lower variance than under the original PPP. Moreover, our numerical investigations shed light on the operator's variance reduction ability, even when applied to more regular point processes than the PPP. This suggests that applying the repulsion operator to the nodes of any Monte Carlo method may decrease its variance and thus enhance the method's statistical accuracy.

Résumé

La méthode de Monte Carlo estime une intégrale en utilisant des évaluations de l'intégrande en certains points appelés nœuds, qui peuvent être choisis comme les points d'un processus ponctuel. Alors que la méthode Monte Carlo brute repose sur le processus ponctuel de Poisson homogène (PPP), certains processus de points plus régulièrement répartis produisent des méthodes avec une décroissance plus rapide de la variance. Dans cette thèse, nous étudions deux familles de processus ponctuels réguliers qui sont des candidats potentiels pour accélérer la convergence de la méthode Monte Carlo brute. La première famille est celle des processus ponctuels hyperuniformes (HUPPs). Un HUPP est caractérisé par la variance du nombre de points dans une grande fenêtre qui évolue plus lentement que le volume de cette fenêtre. En particulier, un HUPP fournit un estimateur Monte Carlo des volumes avec une décroissance de la variance plus rapide que la méthode Monte Carlo brute. Malheureusement, prouver qu'un processus ponctuel est hyperuniforme est généralement difficile. Dans le but de fournir des outils statistiques pour identifier les HUPPs, nous examinons une mesure spectrale appelée la fonction de structure, dont la décroissance autour de zéro permet de diagnostiquer l'hyperuniformité. Nous étudions la dérivation des estimateurs existants de la fonction de structure et contribuons un test statistique asymptotiquement valide de l'hyperuniformité. De plus, nous fournissons une librairie Python contenant tous les estimateurs et outils que nous discutons. La deuxième famille de processus ponctuels que nous étudions est constituée des processus ponctuels repoussés, que nous avons construits en utilisant un opérateur de répulsion. L'opérateur de répulsion réduit le regroupement dans une configuration de points en repoussant légèrement les points les uns des autres. Notre principal résultat théorique est que l'application de l'opérateur de répulsion à un PPP fournit une méthode de Monte Carlo non biaisée avec une variance inférieure qu'avec un PPP. De plus, nos simulations numériques mettent en lumière la capacité de l'opérateur de réduire la variance, même lorsqu'il est appliqué à des processus ponctuels plus réguliers que le PPP. Cela suggère que l'application de l'opérateur de répulsion aux nœuds de n'importe quelle méthode de Monte Carlo peut réduire sa variance et ainsi améliorer la précision statistique de la méthode.

Contents

List of notations

0.1 Abbreviations

BPP	binomial point process
DPP	determinantal point process
GPP	Ginibre ensemble
HUPP	hyperuniform point process
KLY	matching algorithm of Klatt, Last, and Yogeshwaran (2020)
OLS	ordinary least squares linear regression
MOP	multivariate OP Ensemble
PPP	homogeneous Poisson point process
RPPP	repelled Poisson point process
RQMC	randomized quasi-Monte Carlo
RMSE	root mean square error
SSS	scrambled Sobol sequence
TPP	modified Thomas point process

0.2 Letters

$A^{(q,p)}(\mathbf{x})$	Annulus of $\mathbb{R}^d$ centered at $\mathbf{x}$ with small radius q and big radius p	p. 88
$A^{(q,p)}$	Annulus of $\mathbb{R}^d$ centered at the origin with small radius q and big radius p	p. 88
$B(\mathbf{x}, R)$	Ball centered at $\mathbf{x} \in \mathbb{R}^d$ of radius R	p. 36
A^c	Complement of A in $\mathbb{R}^d$, i.e., $\mathbb{R}^d \setminus A$	p. 111
$\mathbb{C}$	Complex space	p. 22
$\mathrm{Corr}(.,.)$	Correlation of two random variables	p. 30
$\mathrm{Cov}(.,.)$	Covariance of two random variables	p. 30
det	Determinant of a matrix	p. 21
$\mathrm{div}(.)$	Divergence of a vector	p. 92
$\mathbb{E}[.]$	Expected value	p. 13
$\mathcal{F}(.)$	Fourier transform	p. 149
$\mathcal{F}_s(.)$	Symmetric Fourier transform	p. 149
$F_{\mathsf{X}}(.)$	Coulomb force driven by the configuration of points X	p. 85
$F_{\mathsf{X}}^{(q,p)}(.)$	Truncated version of the Coulomb force $F_{\mathsf{X}}(.)$	p. 88

$\mathcal{H}_\nu(.)$	Hankel transform of order ν	p. 149
$I(.)$	Integral of a function	p. 27
$\widehat{I}_{\mathcal{P}}(.)$	Monte Carlo estimator with PPP	p. 91
$\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(.)$	Monte Carlo estimator with RPPP	p. 89
$\widehat{I}_{\text{MC}}(.)$	Crude Monte Carlo estimator	p. 29
$\widehat{I}_{\text{MCCV}}(.)$	Monte Carlo control variate estimator	p. 30
$\widehat{I}_{\text{MCDPP}}(.)$	Monte Carlo estimator with DPP	p. 31
$\widehat{I}_{\text{MCRB}}(.)$	Monte Carlo estimator with the repelled BPP	p. 100
$\widehat{I}_{\text{RQMC}}(.)$	Randomized Quasi-Monte Carlo estimator	p. 32
$\text{iMSE}(.)$	Integrated mean square error	p. 62
$\text{i}\mathbb{V}\text{ar}(.)$	Integrated variance	p. 62
$J_\nu(.)$	Bessel function of the first kind of order ν	p. 16
$\mathcal{L}^d$	d-dimensional Lebesgue measure	p. 16
$\mathfrak{N}$	Family of configurations of $\mathbb{R}^d$	p. 12
$o(.)$	Little-o notation	p. 115
$O(.)$	Big-O notation	p. 49
$\mathbb{P}(.)$	Probability	p. 13
$\mathbb{R}^d$	d-dimensional Euclidean space	p. 12
$S(\mathbf{x}, R)$	Sphere centered at $\mathbf{x} \in \mathbb{R}^d$ of radius R	p. 48
$\widehat{S}_{\text{SI}}(.)$	Scattering intensity	p. 43
$\widehat{S}_{\text{T}}(.)$	Tapered estimator	p. 44
$\widehat{S}_{\text{MT}}(.)$	Multitapered estimator	p. 46
$\widehat{S}_{\text{BI}}(.)$	Bartlett's isotropic estimator	p. 49
$\widehat{S}_{\text{HO}}(.)$	Hankel-Ogata estimator	p. 51
$\widehat{S}_{\text{HBC}}(.)$	Hankel-Baddour-Chouinard estimator	p. 52
$\mathbb{T}_{\mathcal{X}}(.)$	Void probability of a point process $\mathcal{X}$	p. 13
$\mathbb{V}\text{ar}[.]$	Variance	p. 36
$\mathbf{x}$	Vector of $\mathbb{R}^d$ i.e., $\mathbf{x} = (x_1, \cdots, x_d)$	p. 12
X	Configuration of points	p. 12
$\mathcal{X}$	Point process	p. 12
$\mathcal{X}(B)$	Number of points (cardinality) of $\mathcal{X} \cap B$	p. 12
$\mathcal{X}_N$	Point process of cardinality N	p. 29

0.3 Greek letters

ρ	Intensity of a point process	p. 14
$\rho^{(n)}(.)$	n-th order intensity function of a point process	p. 14
$\mu^{(n)}(.)$	n-th order intensity measure of a point process	p. 14

σ	Standard deviation	p. 20
Π_ε	Repulsion operator of stepsize ε	p. 85
$\Pi_\varepsilon^{(q,p)}$	Truncated version of the repulsion operator Π_ε	p. 88
$\Pi_{\varepsilon,t}$	$t-$th iterate of the repulsion operator Π_ε	p. 98
κ_d	Volume of the unit ball of $\mathbb{R}^d$	p. 87

0.4 Symbols

$\triangleq$	Definition	p. 13
$\|.\|_2$	Euclidean norm	p. 15
$\lfloor . \rfloor$	Floor, also called integer part	p. 119
∇	Gradient	p. 92
Δ	Laplace operator	p. 92
$\|.\|_\infty$	Norm infinity	p. 129
$< ., . >$	Scalar product	p. 16

1 — Introduction

Contents

1.1 Monte Carlo methods and point processes

Numerical integration is the task of approximating the integral

$$I(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \mathrm{d}\mathbf{x}$$

of an integrable function $f : \mathbb{R}^d \to \mathbb{R}$ using pointwise evaluations of f. The most common approximation algorithm involves summing up the evaluations of f at N points $(\mathbf{x}_i)_{i=1}^N$, called *nodes*, each multiplied by an appropriate *weight* $(w_i)_{i=1}^N$, i.e.

$$I(f) \approx \sum_{i=1}^{N} w_i f(\mathbf{x}_i). \tag{1.1}$$

In addition to the analogy of the approximation in (1.1) with a Riemann sum, Bakhvalov (1971) proved that the minimum worst-case error within a class of functions F, can be achieved by an algorithm of this form (1.1), provided that F satisfies certain properties. Specifically, to find the infimum

$$\inf_{\widehat{I} \in \mathcal{A}} \sup_{f \in F} |I(f) - \widehat{I}(f)|,$$

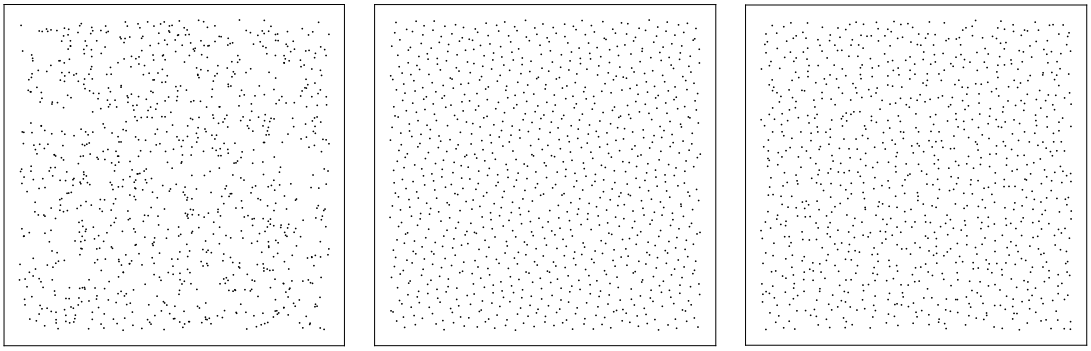
over the collection $\mathcal{A}$ of all the algorithms approximating $I(f)$, it's enough to search within the algorithms of the form (1.1) if F is convex, i.e., if $f_1, f_2 \in F$, then their convex combination $tf_1 + (1-t)f_2 \in F$ for $t \in [0,1]$, and centrally symmetric, i.e., if $f \in F$, then $-f \in F$; see (Novak and Woźniakowski, 2008, Section 4.2.2). Thus, (1.1) seems a reasonable choice of approximation and the remaining focal point lies in determining the configuration of nodes $(\mathbf{x}_i)_{i=1}^N$ (and the corresponding weights) that controls the error, in some sense, for as large class of functions as possible, while also being computationally tractable.

Within the algorithms of the form (1.1), we can distinguish two types: deterministic constructions of nodes, called *quadrature rules*, and random configurations of nodes, called *Monte Carlo methods*. Upon noting that the integral of a smooth function f is the area under the graph of f, the most basic quadrature rules estimate this area using the sum of adjacent rectangular/trapezoidal windows area. For a low space dimension d, the literature is replete with noteworthy quadrature rules. For instance, the Gauss quadrature is exact, i.e., the approximation error is equal to zero when f is a (one-dimensional) polynomial of degree up to $2N-1$ (Gauss, 1815); see also (Owen, 2013, Chapter 7). However, the advantage of quadrature rules rapidly decays as the dimension of the space increases, and Monte Carlo methods become more suitable. This phenomenon shows that deterministic constructions suffer from the curse of dimensionality and was formulated in a theorem of Bakhvalov (1959); see also (Novak, 2016), (Owen, 2013, Chapter 7), and (Dimov, 2008, Chapter 3). Bakhvalov's theorem provides a lower bound of the worst-case error of quadrature rules and the worst-case root mean square error (RMSE) of Monte Carlo methods, for the class of C^k functions supported on a compact with bounded partial derivatives up to order $k > 0$. Specifically, the theorem states that the convergence rate of the worst-case error of a quadrature rule is bounded from below by $N^{-k/d}$ compared to $N^{-1/2-k/d}$ for the worst-case RMSE of a Monte Carlo method. This highlights a notable possible improvement of $N^{-1/2}$ of a Monte Carlo method over a quadrature rule in any space dimension and motivates Monte Carlo methods when estimating high-dimensional integrals.

In its simplest form, crude Monte Carlo amounts to writing the integral of a function supported on a compact K as an expectation under a uniform distribution,

and estimating it by the sample average formed with N i.i.d. uniform samples in K. Figure 1.1a shows 1000 i.i.d. uniform points observed in the cube $[-1,1]^2$. Classical probability results show that the RMSE of the crude Monte Carlo has the form $\sigma(d)N^{-1/2}$ in a d-dimensional space. Interestingly, the convergence rate $N^{-1/2}$ of the RMSE is independent of the dimension d and sampling the nodes is cheap favoring crude Monte Carlo over quadrature rules in high dimensions. However, letting aside the dependence of $\sigma(d)$ on d, $N^{-1/2}$ is a slow convergence rate: dividing the error by 10 requires multiplying the number of nodes N by a factor of 100. This becomes unfavorable in scenarios demanding an estimation with high precision, particularly when evaluating the integrand comes at a substantial cost. Some methods target improving this $N^{-1/2}$ convergence rate by introducing sophisticated dependence through the weights (Delyon and Portier, 2016; Azaïs et al., 2018; Leluc et al., 2023). Other methods, known as variance reduction methods, target reducing the scaling factor $\sigma(d)$, 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. Alternatively, one can consider a more regularly spread configuration of nodes than i.i.d. draws, as in randomized quasi-Monte Carlo (RQMC) which reaches a convergence rate $N^{-3/2}\log(N)^{(d-1)/2}$ under strong smoothness conditions on the integrand (Owen, 2008). Despite the d-dependence of the convergence rate and the strong smoothness assumptions, the RQMC is remarkably efficient in practice for small-to-moderate dimension d. Figure 1.1b illustrates 1000 nodes used in RQMC. This figure visually conveys a more regular arrangement of points compared to the 1000 i.i.d. points displayed in Figure 1.1a.

In spatial statistics, we call a random configuration of points a *point process* (Chiu et al., 2013; Møller and Waagepetersen, 2003; Daley and Vere-Jones, 2003). In particular, the configuration of i.i.d. uniform points corresponds to the binomial point process (BPP). The BPP converges to the famous homogeneous Poisson point process (PPP) as the number of points appropriately approaches infinity (Chiu et al., 2013; Coeurjolly et al., 2017; Last and Penrose, 2017). This has positioned the PPP as the reference point process associated with the crude Monte Carlo. However, when dealing with the PPP, as well as numerous other point



(a) i.i.d. uniform (b) RQMC (c) DPP

Figure 1.1: Samples of 1000 points observed in the cube $[-1,1]^2$.

processes, the number of nodes N is itself a random variable. In such scenarios, we express the convergence rate of the RMSE using the expected number of nodes $\mathbb{E}[N]$ rather than the conventional N. Interestingly, more sophisticated point processes can also be used in numerical integration. Indeed, the first intensity of a point process $\mathcal{X}$ of constant intensity $\rho > 0$ (expected number of points per unit volume) relates the integral of a compactly supported function f to the expectation of a so-called linear statistic

$$\mathbb{E}\left[\frac{1}{\rho}\sum_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})\right] = \int_{\mathbb{R}^d} f(\mathbf{x})\mathrm{d}\mathbf{x},$$

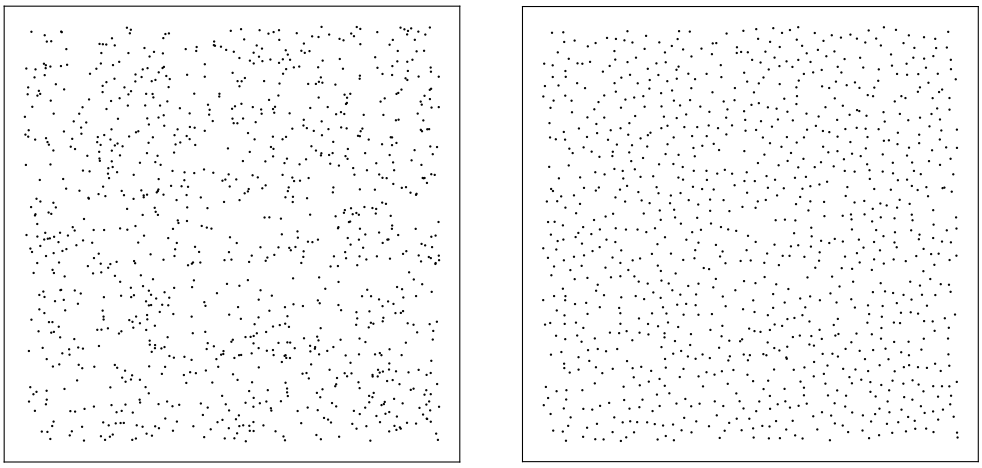
offering a natural *unbiased* Monte Carlo method based on $\mathcal{X}$. Loosely, we call a point process *regular* when it is more regularly spread than the PPP, like in Figure 1.1b. Regular point processes appear more promising than a PPP in numerical integration as they seem more representative of the space. An interesting class of regular point processes is the class of determinantal point processes (DPPs) (Macchi, 1975; Hough et al., 2009). Figure 1.1c displays a realization of 1000 points of a DPP. Monte Carlo methods with some DPPs enhanced the convergence rate of the RMSE beyond $N^{-1/2}$ to reaches $N^{-1/2-1/(2d)}$ for smooth integrands (Ermakov and Zolotukhin, 1960; Bardenet and Hardy, 2020; Coeurjolly et al., 2021; Belhadji et al., 2019).

The purpose of this thesis is to study regular point processes that are potential nodes in a Monte Carlo method. An important class of regular point processes is the class of hyperuniform point processes (HUPPs) (Torquato, 2018; Torquato and Stillinger, 2003; Kim and Torquato, 2018; Klatt et al., 2019, 2022, 2020; Coste,

2021; Dereudre and Flimmel, 2023). A point process is a HUPP if the variance of the number of points in a large window scales slower than the volume of that window. Formally, $\mathcal{X}$ is hyperuniform if

$$\lim_{R\to\infty} \frac{\mathbb{V}\mathrm{ar}\left[\mathcal{X}(B(\mathbf{0},R))\right]}{|B(\mathbf{0},R)|} = 0,$$

where $B(\mathbf{0},R)$ is the ball of radius R centered at the origin. In particular, any HUPP yields a Monte Carlo method with RMSE decaying at a faster rate than crude Monte Carlo, when the integrand f is an indicator function. For example, a HUPP with a number variance scaling in R^{d-1} yields a Monte Carlo method with $\mathbb{E}[N]^{-1/2-1/(2d)}$ decaying rate of the RMSE; this is faster than the $\mathbb{E}[N]^{-1/2}$ convergence rate of the RMSE of crude Monte Carlo. Note that, the DPP displayed in Figure 1.1c is hyperuniform and shows in particular a higher degree of regularity than the i.i.d. sample observed in Figure 1.1a. Outside numerical integration, this statistical property, in turn, implies desirable physical properties for materials; see (Torquato, 2018, Section 14) and references therein. Unfortunately, rigorously proving that a point process is hyperuniform is usually difficult. The common practice is to use a few samples to estimate a spectral measure called the structure factor. The decay of the structure factor around zero provides a diagnostic of hyperuniformity (Coste, 2021; Torquato, 2018). Given a realization of a point process, the standard empirical diagnostic of hyperuniformity thus involves estimating and plotting the structure factor of the underlying point process (Torquato, 2018; Klatt et al., 2019; Kim and Torquato, 2018). This graphical assessment is however not standardized and often not described in full reproducible detail in the literature, with implementation choices and statistical properties often part of each field's folklore. Our first contribution (Hawat et al., 2023b) is a survey and derivation of known estimators of the structure factor. Additionally, we propose some variants of the estimators and new natural estimators. We leverage the consistency of these estimators to contribute an asymptotically valid statistical test of hyperuniformity. Interestingly, another statistical test of hyperuniformity was simultaneously proposed by Klatt, Last, and Henze (2022), tailored to a particular type (Class I in (Torquato, 2018)) of HUPP. We also developed a modular, open-source Python package called structure_factor, that implements the estimators of the structure factor, diagnostics of hyperuniformity, and our statistical test.



(a) PPP (b) Repelled PPP

Figure 1.2: Samples of a PPP and the corresponding repelled sample observed in the cube $[-1,1]^2$.

Back to numerical integration, we mentioned that Monte Carlo with DPPs have demonstrated a faster convergence rate of the RMSE than crude Monte Carlo. The bottleneck of DPP-based methods is the relatively high sampling complexity of DPPs (Gautier et al., 2019a). Aiming to introduce regularity at a smaller computational cost than DPPs, we introduce the repulsion operator, which reduces clustering in a configuration of points X, by slightly pushing the points of X away from each other (Hawat et al., 2023a). The operator's driving gear is the force function

$$\mathbf{x} \mapsto \sum_{\substack{\mathbf{z} \in \mathsf{X} \setminus \{\mathbf{x}\} \\ \|\mathbf{x}-\mathbf{z}\|_2 \uparrow}} \frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|_2^d},$$

frequently used in the gravitational allocations from point processes to the Lebesgue measures (Chatterjee et al., 2010; Nazarov et al., 2007) and the study of Coulomb gases (Serfaty, 2019; Leblé and Serfaty, 2018). Leaving aside further technicality, we observe that each term in the sum defining the force at $\mathbf{x}$ repels the target point $\mathbf{x}$ from a point of the configuration X. Figure 1.2a shows a sample from a PPP observed in $[-1,1]^2$, while Figure 1.2b displays the corresponding repelled sample. The repelled sample shows a higher degree of regularity than the original sample. Our main theoretical result is that applying the repulsion operator to a PPP yields variance reduction for crude Monte Carlo. Moreover, the experiments suggest that the variance reduction also holds when the operator is applied

to already regular point processes. This finding is intriguing on its own because it raises the possibility that the repulsion operator may be universal, in the sense that, it consistently results in a variance reduction of certain linear statistics, regardless of the point process to which it is applied. On the computational side, compared to the (at least) cubic computational complexity for the Monte Carlo estimator with the DPP in (Bardenet and Hardy, 2020), the evaluation of our estimator is only quadratic in the number of nodes and can be easily parallelized without any communication across tasks.

1.2 Outline of the manuscript

The manuscript is organized as follows.

Chapter 2 is an introduction to point processes and their integration within Monte Carlo methods. We review key tools for analyzing point processes, offering an exploration of various point processes, their inherent properties, and distinctive characteristics. Our focus is particularly directed toward regular point processes that are promising candidate nodes of a Monte Carlo method. Additionally, we touch upon various operations that can be applied to manipulate point processes. Finally, we delve into the practical application of point processes in numerical integration and survey some Monte Carlo methods. The knowledge and insights gleaned from this chapter are essential for a comprehensive grasp of the other chapters.

Chapter 3 is based on our journal paper (Hawat et al., 2023b) and focuses on hyperuniform point processes (HUPPs). We first discuss Monte Carlo methods with HUPPs and demonstrate the fast convergence rate of their RMSEs when the integrand is an indicator function. The central part of this chapter is an in-depth examination of a spectral measure called the structure factor, and its application in the study of hyperuniformity. We survey existing estimators of the structure factor, study their main properties, and compare their performances. We investigate the mathematical underpinnings of some prevalent practices within the literature when estimating the structure factor. For example, our results suggest that some of the established conventions on the most common estimator called

scattering intensity, can be relaxed without affecting the accuracy of the estimators. Moreover, we survey commun diagnostics of hyperuniformity, propose an asymptotically valid statistical test of hyperuniformity, and study their performances. Finally, we provide an introduction to the main objects of our Python package structure_factor[1] which implements all known estimators of the structure factor, diagnostics of hyperuniformity, and our statistical test.

Chapter 4 is based on the preprint (Hawat et al., 2023a). In this chapter, we introduce an operator called the repulsion operator acting on a configuration of points to reduce clustering by slightly pushing the points away from each other. We analyze the properties of this operator when applied to a PPP, propose a Monte Carlo method based on it, and describe a sampling procedure that we implement in a Python toolbox MCRPPy[2]. The main theoretical result is a variance reduction of the proposed Monte Carlo method compared to crude Monte Carlo. We illustrate the variance reduction result with numerical experiments. We put the proposed Monte Carlo method in context, by conducting a comparison with the Monte Carlo methods surveyed in Chapter 2. In particular, the experiments validate the variance reduction result and show that our method is competitive with the randomized quasi-Monte Carlo (RQMC), which is currently one of the most powerful methods. Finally, we numerically investigate a few facets of the repulsion operator. Notably, the investigations shed light on the operator's variance reduction ability, even when applied to regular point processes such as a HUPP and the nodes used in RQMC. This suggests that applying the repulsion operator to the nodes of any Monte Carlo method could potentially enhance the method.

Finally, Chapter 5 outlines several potential directions for future research. Additional avenues are presented in the concluding sections of Chapters 3 and 4.

1.3 List of publications

1. **Journal paper:** D. Hawat, G. Gautier, R. Bardenet, and R. Lachieze-Rey. On estimating the structure factor of a point process, with applications to

[1] https://github.com/For-a-few-DPPs-more/structure-factor

[2] https://github.com/dhawat/MCRPPy

hyperuniformity. *Statistics and Computing*, 33, 2023. doi: 10.1007/s11222-023-10219-1.

Companion software: structure_factor. *Python package available on GitHub and PyPI*, 2022.

2. **Preprint:** D. Hawat, R. Bardenet, and R. Lachieze-Rey. Repelled point processes with application to numerical integration, *arXiv*, 2023.
   Companion software: MCRPPy. *Python package available on GitHub*, 2023.

3. **National conference paper:** D. Hawat, G. Gautier, R. Bardenet, and R. Lachieze-Rey. Estimation de la fonction de structure d'un processus ponctuel pour l'étude d'hyperuniformité. *XXVIIIème Colloque Francophone de Traitement du Signal et des Images GRETSI*, 2022.

1.4 Notations

Throughout this manuscript, bold lowercase letters, like $\mathbf{x}$, indicate vectors in $\mathbb{R}^d$, and the corresponding non-bold characters, like x, are scalars. In particular, $\mathbf{x} = (x_1, \cdots, x_d)$. 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 $k = \|\mathbf{k}\|_2$. In general, 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 (a.s.) constant, we sometimes write that point process as $\mathcal{X}_N$, with the cardinality N as the index. Frequently used notations are summarized in the preface to ease the reading. The abbreviations are listed in Section 0.1. The mathfont letters can be found in Section 0.2, organized alphabetically. Lastly, greek letters are presented in Section 0.3, while symbols are covered in Section 0.4.

2 — Point processes

Contents

The primary objects of study in this thesis are *point processes*, which are random configurations of points. Many physical phenomena can be modeled by point processes. For example, the spread of a disease, the arrangement of stars in a galaxy, the positions of trees in a forest, etc. Our main motivation for studying

point processes is numerical integration. Point processes will serve as the integrand evaluation points in a numerical integration scheme. This chapter provides background information on point processes with a special focus on their application in numerical integration. For a broader exploration of point processes, we refer to (Chiu et al., 2013; Møller and Waagepetersen, 2003), and for more insights about numerical integration, we refer to (Owen, 2013; Novak and Woźniakowski, 2008; Dimov, 2008).

Section 2.1 is an introduction to point processes, while Section 2.2 covers basic tools characterizing the distribution of a point process. Section 2.3 focuses on the pair correlation function and the structure factor, which are commonly employed to study the second-order properties of a point process. Moving on, Section 2.4 is dedicated to the renowned homogeneous Poisson point process (PPP), discussing its properties. The PPP is characterized by the independence of the process when restricted to disjoint subsets of the underlying space. In contrast, Section 2.5 focuses on more intricate point processes that exhibit spatial properties such as clustering or regularity with special emphasis on regularity, which proves advantageous in numerical integration. Section 2.6 compiles examples of operations on point processes, including thinning and matching. Finally, in Section 2.7 we explore how point processes find practical use in numerical integration.

2.1 Point processes

Let a configuration of $\mathbb{R}^d$ be a locally finite set $\mathsf{X} \subset \mathbb{R}^d$, that is, for any bounded Borel set B of $\mathbb{R}^d$ the cardinality $\mathsf{X}(B)$ of $\mathsf{X} \cap B$ is finite. Endow the family $\mathfrak{N}$ of such configurations with the σ-algebra generated by the mappings $\mathsf{X} \mapsto \mathsf{X}(B)$, for any bounded Borel set B. Formally, a *(simple)*[1] *spatial point process* (henceforth a *point process*) is a random element $\mathcal{X}$ of $\mathfrak{N}$. Loosely speaking, it is a locally finite random collection of points without duplicated points. The term *point pattern*

[1]The term *simple* indicates that the point process consists of distinct points which is a direct consequence of the configuration definition as a set. However, it's important to note that the condition of simplicity is relaxed in some extensions of the theory, but not in this work; see (Last and Penrose, 2017, Chapter 6).

is commonly used to denote a realization of $\mathcal{X}$, that is, a sample of $\mathcal{X}$. A point process $\mathcal{X}$ is called *stationary* if the law of $\mathcal{X}$ 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, $\mathcal{X}$ is *isotropic* if its law is invariant by rotation. A stationary and isotropic point process is also called *motion-invariant*.

2.2 The distribution of a point process

In this section, we review two essential tools for analyzing and characterizing the distribution of a point process: the system of void probabilities and intensity measures.

2.2.1 Void probabilities

The distribution of a point process $\mathcal{X} \subset \mathbb{R}^d$ is determined by the system of *void probabilities*

$$\mathbb{T}_{\mathcal{X}}(K) \triangleq \mathbb{P}(\mathcal{X}(K) = 0), \tag{2.1}$$

as K ranges through the compact sets of $\mathbb{R}^d$. This follows from interpreting $\mathcal{X}$ as a random closed set with capacity function $1 - \mathbb{T}_{\mathcal{X}}(K)$; see (Chiu et al., 2013, Chapter 4). In words, the void probability $\mathbb{T}_{\mathcal{X}}(K)$ is the probability of observing no points of $\mathcal{X}$ in K.

2.2.2 Intensity measure

The distribution of a point process can also be described by the corresponding intensity measures. The *first intensity measure* $\mu^{(1)}$, is a characteristic analogous to the mean of a real-valued random variable and is defined by

$$\mu^{(1)}(B) \triangleq \mathbb{E}[\mathcal{X}(B)], \tag{2.2}$$

for any Borel set $B \subset \mathbb{R}^d$. When $\mu^{(1)}$ has a density w.r.t. the Lebesgue measure, $\mu^{(1)}(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x})\,\mathrm{d}\mathbf{x}$, we call ρ (denoted also by $\rho^{(1)}$) the *intensity* of $\mathcal{X}$. If $\mathcal{X}$ is

stationary, then $\mu^{(1)}$ is proportional to the Lebesgue measure, and the intensity ρ is a positive **constant**, equal to the mean number of points of $\mathcal{X}$ per unit volume.

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

$$\mathbb{E}\left[\sum_{\mathbf{x}_1,\dots,\mathbf{x}_n\in\mathcal{X}}^{\neq} f(\mathbf{x}_1,\dots,\mathbf{x}_n)\right] = \int_{\mathbb{R}^d\times\cdots\times\mathbb{R}^d} f(\mathbf{x}_1,\dots,\mathbf{x}_n)\mu^{(n)}(\mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_n), \tag{2.3}$$

where f is any non-negative bounded measurable function, and the summation is over all n-tuples of distinct points in $\mathcal{X}$; see (Chiu et al., 2013; Last and Penrose, 2017, Chapter 4). Again, when $\mu^{(n)} = \rho^{(n)}(\mathbf{x}_1,\dots,\mathbf{x}_n)\mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_n$, $\rho^{(n)}$ is called the *n-th order intensity function*. Intuitively, for any pairwise distinct points $\mathbf{x}_1,\dots,\mathbf{x}_n$, we can interpret $\rho^{(n)}(\mathbf{x}_1,\dots,\mathbf{x}_n)\mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_n$ as 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 $\mathrm{d}\mathbf{x}_1,\dots,\mathrm{d}\mathbf{x}_n$.

2.3 Second-order characteristics of a point process

In this section, we recall two essential tools used to analyze the second-order properties of point processes: the pair correlation function and the structure factor. The pair correlation function gives insights into the clustering and regularity of a point process, while the structure factor plays a crucial role in studying the hyperuniformity of a point process, a specific property characterized by a sub-Poissonian number variance. These two functions are instrumental in statistical inference for point processes and can be effectively estimated in practical scenarios. For instance, Chapter 3 will survey various estimators for the structure factor and explore the concept of hyperuniformity. It will also touch upon the estimation of the pair correlation function, acknowledging its extensive study in (Møller and Waagepetersen, 2003, Chapter 4).

2.3.1 Pair correlation function

The second-order intensity measure $\mu^{(2)}$ of $\mathcal{X}$ helps assess pairwise correlations in the points of $\mathcal{X}$. Recall that by (2.3), $\mu^{(2)}$ is formally defined by

$$\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}}^{\neq} f(\mathbf{x},\mathbf{y})\right] = \int_{\mathbb{R}^d\times\mathbb{R}^d} f(\mathbf{x},\mathbf{y})\mu^{(2)}(\mathrm{d}\mathbf{x},\mathrm{d}\mathbf{y}),$$

where f is any non-negative bounded measurable function. When $\mathcal{X}$ is stationary[2] with intensity ρ, the previous expression can be factorized in

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} f(\mathbf{x}+\mathbf{y},\mathbf{y})\rho^2 g_2(\mathrm{d}\mathbf{x})\mathrm{d}\mathbf{y}, \tag{2.4}$$

where g_2 is called the *pair correlation measure*. Intuitively, $g_2(\mathrm{d}\mathbf{x})$ is the probability that $\mathcal{X}$ has a point at location $\mathrm{d}\mathbf{x}$, given that $\mathcal{X}$ already contains the origin. If in addition, g_2 has a density g w.r.t. the Lebesgue measure, i.e., $g_2(\mathrm{d}\mathbf{x}) = g(\mathbf{x})\mathrm{d}\mathbf{x}$, then g is called the *pair correlation function* of $\mathcal{X}$ which can be expressed as

$$g(\mathbf{r}) = \frac{\rho^{(2)}(\mathbf{0},\mathbf{r})}{\rho^2}, \tag{2.5}$$

if the second order intensity function $\rho^{(2)}$ exists; see (Møller and Waagepetersen, 2003, Definition 4.4). Informally, $g(\mathbf{r})$ counts the pairs $(\mathbf{x},\mathbf{y}) \in \mathcal{X}\times\mathcal{X}$ such that $\mathbf{x}-\mathbf{y} = \mathbf{r} \in \mathbb{R}^d$. Moreover, when $\mathcal{X}$ is assumed both stationary and isotropic[3], then the pair correlation function g is radial[4]. In that case, we abusively write $g(\mathbf{r}) = g(r)$, where $r = \|\mathbf{r}\|_2$. Additionally, in this context, the second-order intensity function $\rho^{(2)}$ is also a radial function and we have

$$g(r) = \frac{\rho^{(2)}(r)}{\rho^2}. \tag{2.6}$$

2.3.2 Structure factor

In the previous section, we used the second-order intensity measure to define the pair correlation function g in (2.5). Now, following Coste (2021), we will see

[2]The stationarity condition can be relaxed while still defining g_2 in the same manner; see (Daley and Vesilo, 1997, Proposition 8.1.I) or (Møller and Waagepetersen, 2003, Chapter 4).

[3]The assumptions of stationarity and isotropy can be straightforwardly weakened to assuming that the first and second intensity measures are translations and rotation invariants.

[4]A real-valued function f of $\mathbb{R}^d$ is called a radial function if, for any $\mathbf{x} \in \mathbb{R}^d$, $f(\mathbf{x})$ depends only on $\|\mathbf{x}\|_2$.

that by transitioning to the Fourier domain, we can (often) define the so-called structure factor S which is a main tool in the study of hyperuniformity.

The *structure factor measure*[5] $\mathcal{S}$ of a stationary point process $\mathcal{X} \subset \mathbb{R}^d$ with intensity $\rho > 0$ is the measure on $\mathbb{R}^d$, when it exists, defined by

$$\mathcal{S} = \mathcal{L}^d + \rho\mathcal{F}\left(g_2 - \mathcal{L}^d\right) = \mathcal{F}\left(\delta_0 + \rho(g_2 - \mathcal{L}^d)\right), \tag{2.7}$$

where $\mathcal{L}^d$ is the d-dimensional Lebesgue measure, $\mathcal{F}$ is the Fourier transform (A.1), g_2 is the pair correlation measure (2.4), and δ_0 is the Dirac mass in 0; see (Coste, 2021). We note that in the spectral analysis of stochastic processes $\mathcal{S}$, up to a scale factor, is also called the *Bartlett spectral measure*.

When the measure $\mathcal{S}$ is absolutely continuous w.r.t. the Lebesgue measure, i.e., $\mathcal{S}(\mathrm{d}\mathbf{k}) = S(\mathbf{k})\mathrm{d}\mathbf{k}$, we call S the *structure factor* [6]. If in additionel, g_2 is absolutely continuous w.r.t. the Lebesgue measure and $g - 1$ is integrable, S can be explicitly written as

$$S(\mathbf{k}) = 1 + \rho\mathcal{F}(g - 1)(\mathbf{k}). \tag{2.8}$$

Moreover, when the pair correlation function is symmetric, i.e., $g(\mathbf{r}) = g(-\mathbf{r})$, S is a symmetric real-valued function. In particular, when $\mathcal{X}$ is isotropic, then both the pair correlation function g and the structure factor S are radial functions. Abusively denoting $S(k) = S(\mathbf{k})$, Equation (2.8) reads

$$S(k) = 1 + \rho\mathcal{F}_s(g - 1)(k), \tag{2.9}$$

where $\mathcal{F}_s$ is the symmetric Fourier transform (A.2). Using (A.4) we get

$$S(k) = 1 + \rho\frac{(2\pi)^{d/2}}{k^{d/2-1}}\int_0^\infty (g(r) - 1)r^{d/2}J_{d/2-1}(kr)\,\mathrm{d}r, \tag{2.10}$$

where J_ν is the Bessel function of the first kind of order ν; see Torquato (2018, Section 2). Finally, if the pair correlation function g exists and is smooth, then one can expect that

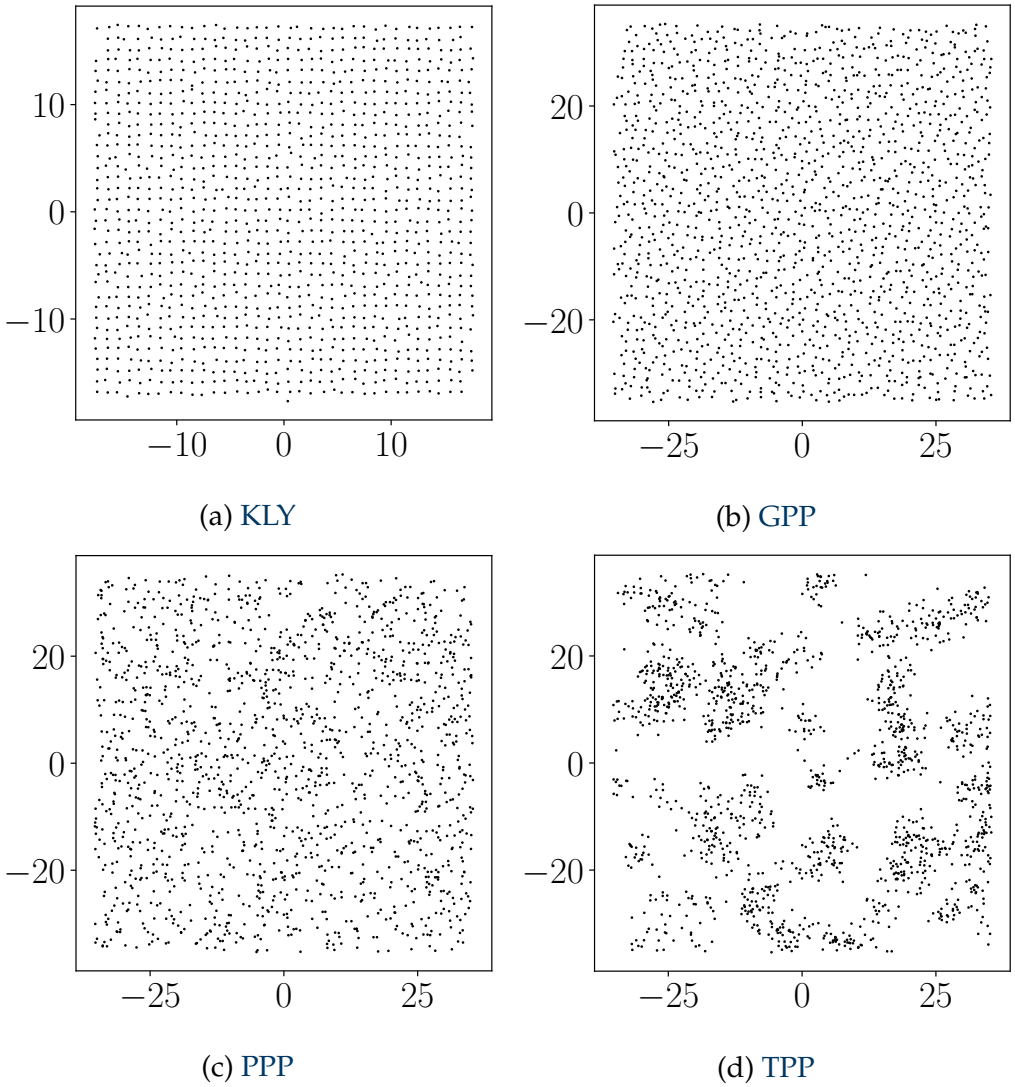
$$S(\mathbf{k}) \xrightarrow[\|\mathbf{k}\|_2 \to \infty]{} 1.$$

On the other hand, the behavior of S in zero measures the fluctuations of g around 1 at large scales $\|\mathbf{r}\|_2 \gg 1$, which can, in turn, help quantify properties like hyperuniformity studied in Chapter 3.

[5] $\mathcal{S}$ is a signed positive-definite measure; see (Coste, 2021, Section 1.3).

[6] The literature is inconsistent as to whether the structure factor is the measure $\mathcal{S}$ or its density S. We choose the density, which is also sometimes known as the *scaled spectral density function*; see (Rajala et al., 2020) and references therein.

2.4 The homogeneous Poisson point process



(a) KLY (b) GPP

(c) PPP (d) TPP

Figure 2.1: Samples of four distinct point processes. Plots obtained using the toolbox structure_factor.

Consider a compact set K of $\mathbb{R}^d$, and pick N i.i.d. points uniformly distributed in K. The point process $\mathcal{B}_N$ formed by these N points is called the binomial point process (BPP) with 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)[7] of constant intensity $\rho > 0$.

[7]In this context, the term "homogeneous" refers to the property of translation invariance, where the process exhibits a constant intensity. However, it is important to note that a Poisson

Figure 2.1c displays a realization from a PPP of intensity $\rho = 1/\pi$. 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 $\mathcal{X}$ 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 $\mathbb{T}_{\mathcal{P}}(B)$ (2.1) of $\mathcal{P}$ is equal to

$$\mathbb{T}_{\mathcal{P}}(B) = \exp(-\rho|B|).$$

Third, $\mathcal{P}(B_1), ..., \mathcal{P}(B_n)$ are independent for any disjoint Borel sets $B_1, ..., B_n$ with $n \geq 2$. 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 (Chiu et al., 2013, Chapter 2). In the same vein, a straightforward calculation reveals that both the pair correlation function (2.5) and the structure factor (2.8) of $\mathcal{P}$ are equal to one; see Figures 2.2a and 2.2b. Fourth, all the moments of $\mathcal{P}$ are determined by ρ, i.e., for any non-negative measurable function f,

$$\mathbb{E}\left[\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_n\in\mathcal{P}}^{\neq} f(\mathbf{x}_1,\cdots,\mathbf{x}_n)\right] = \rho^n \int_{\mathbb{R}^d\times\cdots\times\mathbb{R}^d} f(\mathbf{x}_1,\cdots,\mathbf{x}_n)\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_n. \quad (2.11)$$

With the notation of Section 2.2.2, this reads $\rho^{(n)} = \rho^n$. Moreover, by the extended Slivnyak-Mecke theorem, we have

$$\mathbb{E}\left[\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in\mathcal{P}}^{\neq} h(\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathcal{P}\setminus\{\mathbf{x}_1,\ldots,\mathbf{x}_n\})\right] = \int_{\mathbb{R}^d\times\cdots\times\mathbb{R}^d} \mathbb{E}[h(\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathcal{P})]\,\rho^n\mathrm{d}\mathbf{x}_1\ldots\mathrm{d}\mathbf{x}_n, \quad (2.12)$$

for any non-negative measurable function h on $(\mathbb{R}^d)^n \times \mathfrak{N}$; see (Coeurjolly et al., 2017, Section 5.1). Equation (2.12) 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 $\mathcal{P}$.

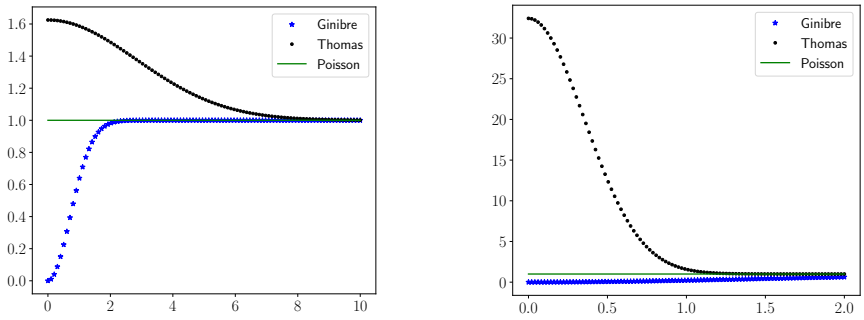
The Poisson point process is a reference point process. First, it plays a central role as a null hypothesis in statistical tests of interaction. For example, by comparing the pair correlation function of a given point process $\mathcal{X}$ with that of a PPP

point process can also have a non-constant intensity, indicating a departure from translation invariance. In such cases, the process is called an "inhomogeneous Poisson point process".

$\mathcal{P}$, which is equal to one, we can investigate the spatial clustering or regularity of points. Intuitively, the pair correlation function of $\mathcal{P}$, reflects the complete independence between the pair of points at all scales. In other words, the probability of finding a pair of points of $\mathcal{P}$ separated by any distance remains constant. Roughly, when the pair correlation function $g(r)$ of $\mathcal{X}$ exceeds 1, it indicates more clustering in the points of $\mathcal{X}$ compared to $\mathcal{P}$ at distances r, while $g(r)$ below 1 suggests more regularity i.e., less clustering than $\mathcal{P}$ at distances r (Møller and Waagepetersen, 2003, Section 4.2.1).

Another role of the Poisson point process is its usage as a fundamental building block for more complex models, including the Thomas point process (Example 2.1), the KLY point process (Example 2.4), and the repelled Poisson point process that will be introduced in Chapter 4.

Finally, note that the Python toolbox structure_factor provides a method for sampling a PPP; see Section 3.4.



(a) $g(r)$ (b) $S(k)$

Figure 2.2: Pair correlation functions $g(r)$ and structure factors $S(k)$ of some point processes.

2.5 Clustering and regularity

While the PPP is characterized by complete randomness, a positive correlation or attraction between the points of a point process $\mathcal{X}$ leads to the formation of clumps and voids in its point patterns. As a consequence, $\mathcal{X}$ will exhibit more clustering of points than a PPP. Apart from the visual inspection, this clustering behavior can be detected from the pair correlation function of $\mathcal{X}$ that exceeds one

(within some range of distances). We denote such a point process, a *clustering point process*.

Conversely, negative correlation or repulsion between the points of a point process $\mathcal{Y}$ entails a reduced tendency for clustering in its point patterns. This results in a more uniform or regular arrangement of points compared to a PPP. In contrast to clustering, regularity is typically characterized by a pair correlation function that is less than one within a range of distances. We denote this type of point process, a *regular point process*.

This section collects examples of each type: Section 2.5.1 contains examples of clustering point processes, while Section 2.5.2 provides examples of regular point processes.

2.5.1 Clustering point processes

Although the PPP is a simplistic model, it can serve as a basis for constructing clustering models. One example is the family of Neyman-Scott processes introduced by Neyman and Scott (1958) as a model for the locations of galaxies in space. A Neyman-Scott process is obtained by artificially creating clustering of points around the point of a PPP $\mathcal{P}$. First, a random number of i.i.d. points are placed around each point of $\mathcal{P}$ according to a probability distribution ϕ. Then, the original points of $\mathcal{P}$ are removed. The resulting point process $\mathcal{X}$ is a clustering point process. $\mathcal{X}$ is stationary, and if ϕ is radially symmetric, then $\mathcal{X}$ will also be isotropic (Chiu et al., 2013, Section 5.3).

Example 2.1 (Thomas point process). The modified Thomas point process (TPP) [8] is a Neyman-Scott point process of $\mathbb{R}^d$ that is reminiscent of a mixture of Gaussians. To obtain a TPP, fix λ, σ and ρ_{parent} positive, and let $\mathcal{P}_{\text{parent}}$ be a PPP of intensity ρ_{parent}. Conditionally to $\mathcal{P}_{\text{parent}}$, let $(N_{\mathbf{x}})_{\mathbf{x}\in\mathcal{X}_{\text{parent}}}$ be i.i.d. Poisson random variables with mean λ. For any $\mathbf{x} \in \mathcal{P}_{\text{parent}}$, conditionally to $N_{\mathbf{x}}$, let $(\phi_{\mathbf{x},i})_{i=1}^{N_{\mathbf{x}}}$

[8] In the initial formulation of the Thomas point process by Thomas (1949), clusters were constrained to be non-empty, meaning that for every $\mathbf{x} \in \mathcal{X}_{\text{parent}}$, $N_{\mathbf{x}}$ was conditioned to be non-zero. Diggle (1978) proposed removing this constraint to make the algebra simpler, resulting in the so-called "modified" Thomas point process; see (Baddeley et al., 2015, Section 12.3.3).

be i.i.d. d-dimensional centered isotropic Gaussian vectors with variance σ^2. The resulting point process

$$\mathcal{X}_{\text{Thomas}} = \bigcup_{\mathbf{x}\in\mathcal{P}_{\text{parent}}} \{\mathbf{x} + \phi_{\mathbf{x},i},\, i = 1, \ldots, N_{\mathbf{x}}\},$$

is a TPP of intensity $\rho_{\text{Thomas}} = \rho_{\text{parent}} \times \lambda$ (Chiu et al., 2013, Section 5.3). Figure 2.1d shows a realization of a TPP of intensity $1/\pi$, generated from a parent PPP of intensity $\rho_{\text{parent}} = 1/(20\pi)$, and a standard deviation $\sigma = 2$. Since σ is small enough compared to $\lambda = 20$, clusters naturally appear. The pair correlation function and the structure factor of the TPP, shown in Figures 2.2a and 2.2b, are both radial functions (Møller and Waagepetersen, 2003, Example 5.3), given by

$$g_{\text{Thomas}}(r) = 1 + \frac{1}{\rho_{\text{parent}}(\sqrt{4\pi\sigma^2})^d} \exp(-\frac{r^2}{4\sigma^2}), \tag{2.13}$$

and

$$S_{\text{Thomas}}(k) = 1 + \lambda \exp\left(-k^2\sigma^2\right). \tag{2.14}$$

The fact that the pair correlation function g_{Thomas} consistently remains above one indicates the attractive nature of the TPP. This aspect is visually evident from the clustering observed in Figure 2.1d. Note that the Python toolbox structure_factor provides a method for sampling a TPP; see Section 3.4.

2.5.2 Regular point processes

A typical example of regular point processes is the family of determinantal point processes (DPPs), formalized by Macchi (1975)[9]. DPPs are point processes parameterized by a kernel $\mathtt{K}$. The distinctive characteristic of a determinantal point process (DPP) lies in the determinantal form of the intensity functions. Formally, for any integer $n \geq 1$, and $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^d$ the n-th order intensity function $\rho^{(n)}$ (2.3) of a DPP $\mathcal{X}$ of kernel $\mathtt{K}$ is given by

$$\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det\left(\mathtt{K}(\mathbf{x}_i, \mathbf{x}_j)\right)_{1\leq i,j\leq n}. \tag{2.15}$$

[9]Before the general notion of determinantal point processes was formalized by Macchi in 1975, several examples of DPPs were already known, including the Ginibre ensemble introduced by Ginibre (1965), and the DPP used by Ermakov and Zolotukhin (1960) to sample quadrature nodes.

This is at the heart of the virtues of a DPP, which simplifies the theoretical analysis, making DPPs practical and facilitating their study. For more theoretical detail see (Hough et al., 2009, Chapter 4) and for sampling from the DPPs see the thesis of Gautier (2020).

When $\mathcal{X}$ is stationary of intensity ρ and $\mathtt{K}$ is symmetric, Equations (2.15) and (2.5) give

$$g(\mathbf{r}) = 1 - \frac{\mathtt{K}^2(\mathbf{0}, \mathbf{r})}{\rho^2}.$$

Thus, the pair correlation function g of $\mathcal{X}$ is less than one. This observation highlights that stationary DPPs with symmetric kernels inherently encompass the repulsion aspect as part of their definition.

Example 2.2 (Ginibre ensemble). The Ginibre ensemble (GPP) introduced by Ginibre (1965) is a stationary and isotropic DPP of $\mathbb{C} \approx \mathbb{R}^2$ with kernel

$$\mathtt{K}_{\text{Ginibre}}(\mathbf{z}, \mathbf{w}) = \sum_{k \geq 0} \frac{(\mathbf{z}\overline{\mathbf{w}})^k}{k!},$$

where $\mathbf{z}, \mathbf{w} \in \mathbb{C}$ (Hough et al., 2009, Example 4.5.8). It can be defined (and approximately sampled) as the limit of the set of eigenvalues of a matrix 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).

Its intensity is equal to $\rho_{\text{Ginibre}} = 1/\pi$, and its pair correlation function (Figure 2.2a) is

$$g_{\text{Ginibre}}(r) = 1 - \exp(-r^2). \tag{2.16}$$

The fact that g_{Ginibre} consistently remains below one indicates the GPP is a regular point process. This behavior is visually evident in Figure 2.1b. Finally, its structure factor (Figure 2.2b) can be computed exactly,

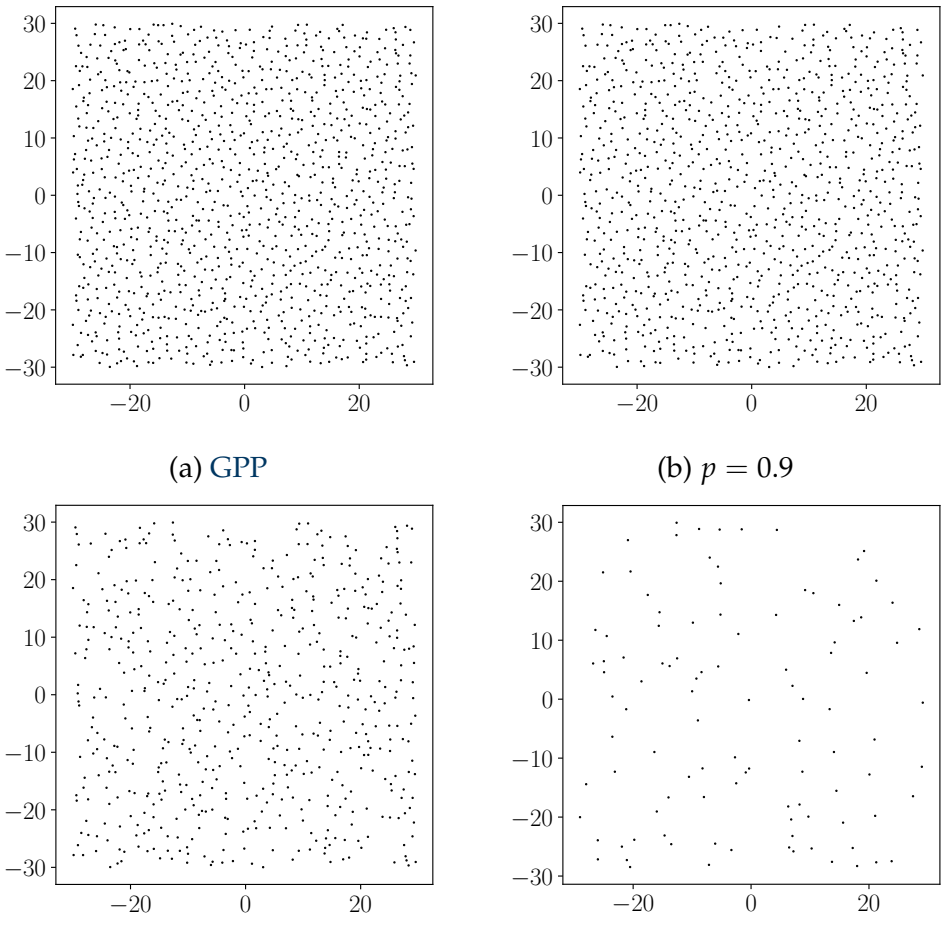
$$S_{\text{Ginibre}}(k) = 1 - \exp(-k^2/4). \tag{2.17}$$

Note that the Python toolbox structure_factor provides a method for sampling a GPP; see Section 3.4.

2.6 Operations on point processes

This section presents some operations on point processes. Section 2.6.1 is concerned with independent thinning and Section 2.6.2 focuses on matching.

2.6.1 Independent thinning



(a) GPP (b) $p = 0.9$

(c) $p = 0.5$ (d) $p = 0.1$

Figure 2.3: A sample from the GPP, and the samples obtained after applying an independent thinning with different retaining probabilities p. Plots obtained using the toolbox structure_factor.

In the context of point processes, a thinning is an operation on a point process, where points are selectively removed based on a specific rule, resulting a new point process called the *thinned point process*.

For a given point process $\mathcal{X}_{\text{parent}}$, an (independent) p- thinning of $\mathcal{X}_{\text{parent}}$ consists on deleting each point of $\mathcal{X}_{\text{parent}}$ with probability $1 - p$. The deletion of each point is independent of the locations and possible deletions of any other points of $\mathcal{X}_{\text{parent}}$. The remaining points constitute the thinned point process (Chiu et al., 2013, Section 5.1). The parameter p is called the *retaining probability*, being the probability of keeping a point. Figure 2.3 shows a p-thinning of a sample from the GPP, with retaining probabilities $p \in \{0.1, 0.5, 0.9\}$.

If the parent point process $\mathcal{X}_{\text{parent}}$ is motion-invariant of intensity ρ_{parent}, the resulting thinned point process $\mathcal{X}_p$ is also motion-invariant of intensity

$$\rho_p = p \times \rho_{\text{parent}}.$$

The pair correlation function of $\mathcal{X}_p$ is independent of p and actually equal to the parent point process's counterpart $g_p(r) = g_{\text{parent}}(r)$. However, the structure factor of $\mathcal{X}_p$ is equal to

$$S_p(k) = pS_{\text{parent}}(k) + 1 - p, \tag{2.18}$$

where S_{parent} is the structure factor of $\mathcal{X}_{\text{parent}}$. The structure factors of thinned GPP are shown in Figure 2.4 for a few values of p.

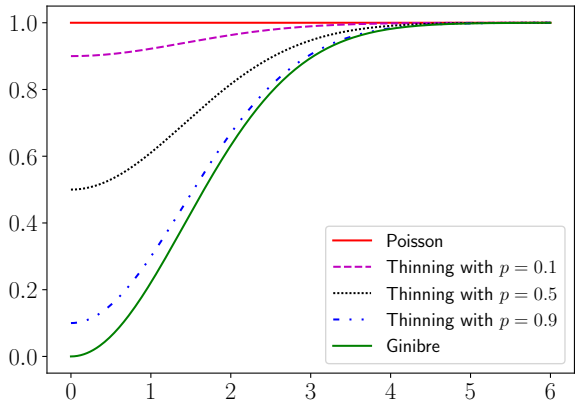


Figure 2.4: Structure factor of the GPP (Ginibre) before and after applying independent thinning with different retaining probabilities p. The red line (Poisson) corresponds to the structure factor of the PPP for reference.

Example 2.3. A p-thinning of a PPP of intensity ρ results again a PPP of intensity $p\rho$. This property can be easily verified by computing the void-probabilities (2.1) of the thinned point process.

2.6.2 Matching

This section focuses on the operation of matching for point processes. We start by a matching of a point process with the grid $\mathbb{Z}^d$.

Example 2.4 (Matching a point process with $\mathbb{Z}^d$). This matching was designed by Klatt, Last, and Yogeshwaran (2020) to generate a particular type of point processes exhibiting sub-Poisson number variance known as a hyperuniform point process; see Chapter 3. Loosely speaking, each point of $\mathbb{Z}^d$ is matched with a close-by point of a user-defined point process $\mathcal{X}$ in $\mathbb{R}^d$, like a PPP. The matched points from $\mathcal{X}$ form a new point process $\mathcal{X}_{\text{child}}$. When $\mathcal{X}$ is a PPP, we denote $\mathcal{X}_{\text{child}}$ by the matching algorithm of Klatt, Last, and Yogeshwaran (2020) (KLY). The KLY is an example of a point process that is known to be stationary but not isotropic. However, the corresponding pair correlation function and structure factor are unknown. Figure 2.1a shows a sample from the KLY generated by matching a subset of $\mathbb{Z}^2$ with a realization of a PPP of intensity $\rho = 11$. The intensity of the resulting point process is equal to 1. Figure 2.1a strongly suggests that the KLY is a regular point process. Finally, note that the Python toolbox structure_factor provides a method for sampling from the KLY; see Section 3.4. Additionally, it's worth mentioning that Klatt et al. (2020) have supplied a corresponding code written in R.

The last example of this section is a matching of a point process with the Lebesgue measure known as *gravitational allocation* (or *transportation by gradient flow*). This matching will inspire the construction of a new point process in Chapter 4.

Example 2.5 (Matching a point process with the Lebesgue measure). A gravitational allocation is a fair matching between a point process $\mathcal{X}$ and the Lebesgue measure. This matching aims to partition equitably the space -up to a negligible set- between the points of $\mathcal{X}$ via a map called an allocation rule. That is, every point in $\mathcal{X}$ is assigned with a subset of the space called *basin*, and all basins have equal Lebesgue measure. The allocation rule is commonly established through a

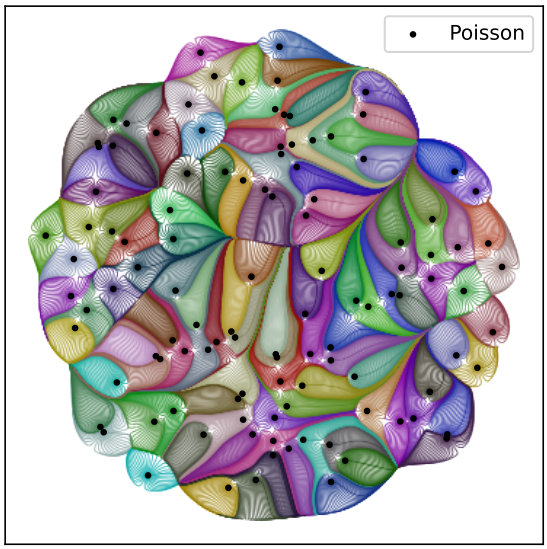


Figure 2.5: Illustration of the gravitational allocation from Lebesgue to a sample (black dots) of a PPP of intensity $1/\pi$ in a disk. The provided image is purely for illustrative purposes and no claims are made regarding the existence, in a strict sense, of a gravitational allocation from the Lebesgue measure to a PPP in dimension $d = 2$. Plot obtained using the toolbox MCRPPy.

differential equation that governs the movement of the points of $\mathbb{R}^d \setminus \mathcal{X}$ driven by the Newtonian gravitation potential U (Nazarov et al., 2007)

$$\frac{\partial}{\partial t} Y_{\mathbf{x}}(t) = -\nabla U(Y_{\mathbf{x}}(t)), \tag{2.19}$$

or force F (Chatterjee et al., 2010)

$$\frac{\partial}{\partial t} Y_{\mathbf{x}}(t) = F(Y_{\mathbf{x}}(t)). \tag{2.20}$$

The solution $Y_{\mathbf{x}}(t)$ of the corresponding differential equation with starting point $\mathbf{x} \in \mathbb{R}^d \setminus \mathcal{X}$, i.e. $Y_{\mathbf{x}}(0) = \mathbf{x}$, is called the *gravitational flow curve* of $\mathbf{x}$. Roughly, for almost any $\mathbf{x} \in \mathbb{R}^d \setminus \mathcal{X}$, the gravitational flow curve of $\mathbf{x}$ collects the trajectory of $\mathbf{x}$ as it moves towards a point $\mathbf{z}$ of $\mathcal{X}$, where it terminates ($\lim_{t\to\infty} Y_{\mathbf{x}}(t) = \mathbf{z}$). In this case, i.e., when $Y_{\mathbf{x}}(t)$ terminates at $\mathbf{z} \in \mathcal{X}$, with $Y_{\mathbf{x}}(0) = \mathbf{x}$, we say that $\mathbf{x}$ is allocated to $\mathbf{z}$. The points of $\mathbb{R}^d \setminus \mathcal{X}$ that are not allocated to any points of $\mathcal{X}$ form the boundaries of the basins.

Figure 2.5 illustrates a gravitation allocation[10] from the Lebesgue measure to a PPP in the disk. The simulation procedure is detailed in Appendix B. The region underlying the curves sharing the same color illustrates a basin, that is the points of the space allocated to the point of the PPP that belongs to that particular colored region; see also Figure B.1 and Section 4.3.5.

2.7 Numerical integration using point processes

Numerical integration is the task of approximating function integrals. Given an integrable function f supported on a compact set $K \subset \mathbb{R}^d$, the goal is to approximate the integral of f

$$I(f) \triangleq \int_K f(\mathbf{x}) \mathrm{d}\mathbf{x}. \tag{2.21}$$

The problem of finding good algorithms for approximating $I(f)$ has been extensively explored, resulting in wealth types of algorithms. Our focus is on linear non-adaptive estimators of the form

$$\widehat{I}_{\mathcal{X}_N}(f) = \sum_{\mathbf{x} \in \mathcal{X}_N} w_i f(\mathbf{x}), \tag{2.22}$$

where $\mathcal{X}_N = \{\mathbf{x}_i\}_{i=1}^N$ is a random non-adaptive[11] configuration of points belonging to $W \supseteq K$, typically $W = K$, and $\{w_i\}_{i=1}^N$ are some weights, typically fixed. We call an estimator of the form (2.22) a *Monte Carlo method* and we use the root mean square error (RMSE)

$$\mathbb{E}\left[\left(I(f) - \widehat{I}_{\mathcal{X}_N}(f)\right)^2\right]^{1/2} \tag{2.23}$$

to measure the error of $\widehat{I}_{\mathcal{X}_N}(f)$. The complexity of $\widehat{I}_{\mathcal{X}_N}(f)$ is quantified by its convergence rate i.e., the convergence rate of the error RMSE (2.23) to zero as N increases, and the time needed to compute $\widehat{I}_{\mathcal{X}_N}(f)$ on machines.

[10] The provided image serves a purely illustrative purpose, and no assertions are made regarding the existence of a gravitational allocation from the Lebesgue measure to a PPP in dimension $d = 2$.

[11] An adaptive method uses *a priori* and/or *a posteriori* information obtained during the calculation. In this framework, non-adaptive means that for any i, the choice of the point $\mathbf{x}_i$ doesn't depend on the values of $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_{i-1})$.

In this section, we explore how point processes find practical use in numerical integration. In Section 2.7.1, we outline how we can build an unbiased Monte Carlo method using a point process. Finally, Section 2.7.2 provides an overview of some Monte Carlo methods. While the landscape of Monte Carlo methods is extensive, our focus here is to highlight the standard and relevant methods within the scope of our study and demonstrate the advantages of regular point processes in this context.

2.7.1 Monte Carlo methods and point processes

Recall that from the definition of the first intensity (2.2) of a point process $\mathcal{X}$, we have for any integrable function f of compact support K

$$\int_{\mathbb{R}^d} f(\mathbf{x})\rho(\mathbf{x})\mathrm{d}\mathbf{x} = \mathbb{E}\left[\sum_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})\right],$$

where ρ is the intensity function of $\mathcal{X}$. Hence, for any compact W such that $K \subseteq W$

$$\widehat{I}_{\mathcal{X}\cap W}(f) = \sum_{\mathbf{x}\in\mathcal{X}\cap W} \frac{f(\mathbf{x})}{\rho(\mathbf{x})} \tag{2.24}$$

is a natural unbiased Monte Carlo estimator of $I(f)$ and its RMSE (2.23) reduces to the standard deviation

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\mathcal{X}\cap W}(f)\right]^{1/2}.$$

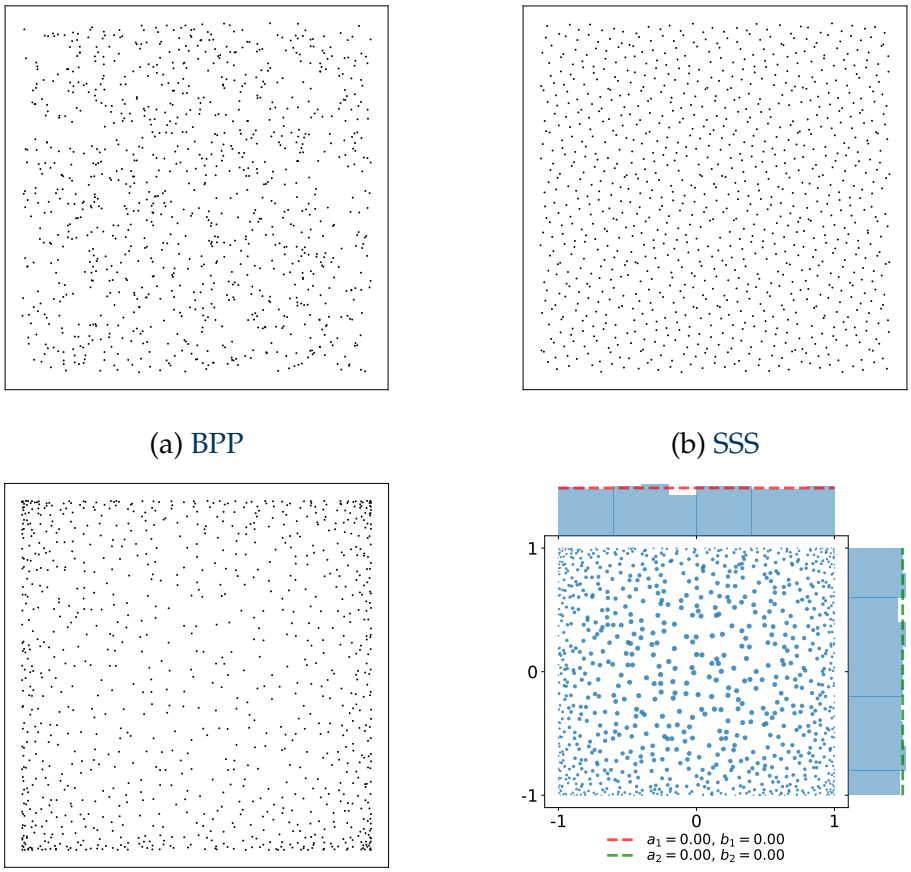
A common choice for W is the smallest ball or box window that encompasses K. We stress that the number of evaluation points of $\widehat{I}_{\mathcal{X}\cap W}$ is a random variable and its expected value is equal to $\mathbb{E}[\mathcal{X}(W)] = \int_W \rho(\mathbf{x})\mathrm{d}\mathbf{x}$.

Example 2.6. For $\mathcal{P}$ a PPP of intensity $\rho > 0$, we have

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\mathcal{P}\cap W}(f)\right] = \rho^{-1}\int_K f^2(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{2.25}$$

2.7.2 A few Monte Carlo methods

We highlighted in Section 2.7.1 how a point process offers a natural Monte Carlo method (2.24). In this section, we review some unbiased Monte Carlo methods and discuss the corresponding convergence rates.



(a) BPP (b) SSS

(c) MOP (d) Weighted MOP

Figure 2.6: A sample of 1000 points from the BPP in (a), SSS in (b), and MOP in (c) observed in the cube $[-1,1]^2$. Figure (d) displays the weighted sample corresponding to the MOP observed in (a). Each disk in Figure (d) is centered at a point $\mathbf{x}$ of the MOP, and the area of the disk is proportional to the weight $1/\mathrm{K}_N(\mathbf{x},\mathbf{x})$ corresponding to $\mathbf{x}$. The marginal plots on each axis in Figure (d) depict the marginal histograms of the weighted sample, with a green and a red curve indicating the density of the marginal Jacobi measures which we have set the Lebesgue measure on $[-1,1]^d$. The plots (c) and (d) are obtained using the toolbox DPPy (Gautier et al., 2019c)

Let f be a continuous function supported on $K = [-1/2,1/2]^d$. The crude (or simple) Monte Carlo employs a binomial point process (BPP) $\mathcal{B}_N$ (Figure 2.6a) supported on K, to estimate $I(f)$ (2.21) as follows

$$\widehat{I}_{\mathrm{MC}}(f) = \frac{1}{N}\sum_{\mathbf{x}\in\mathcal{B}_N} f(\mathbf{x}). \tag{2.26}$$

The number of points used is fixed to N. $\widehat{I}_{\mathrm{MC}}(f)$ is an unbiased estimator of $I(f)$ with a variance equal to $N^{-1}\mathbb{V}\mathrm{ar}[f(\mathbf{u})]$ where $\mathbf{u}$ is a uniformly drawn point of K. Hence, the convergence rate of $\widehat{I}_{\mathrm{MC}}$ is $N^{-1/2}$.

Remark 2.1. As mentioned in Section 2.4, when the number of points of a BPP increases appropriately with the size of the observation window, the BPP converges to a PPP. Consequently, one can think of estimator $\widehat{I}_{\mathrm{MC}}(f)$ as a self-normalized version of the estimator $\widehat{I}_{\mathcal{P}\cap K}(f)$ (2.24), where $\mathcal{P}$ is a PPP with suitable intensity. Thereby, the PPP is also associated with the crude Monte Carlo. However, setting the intensity of $\mathcal{P}$ equal to N, we get

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

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

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

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 $\widehat{I}_{\mathrm{MC}}$; see e.g. (Owen, 2013) and references therein. 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

$$\widehat{I}_{\mathrm{MCCV}}(f) = \widehat{I}_{\mathrm{MC}}(f) - c(\widehat{I}_{\mathrm{MC}}(h) - I(h)), \tag{2.27}$$

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

$$\frac{\mathrm{Cov}(f(\mathbf{u}), h(\mathbf{u}))}{\mathbb{V}\mathrm{ar}[h(\mathbf{u})]},$$

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

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

Hence, any function h that correlates with f is helpful to reduce the variance, even one that correlates negatively. However, the convergence rate remains $N^{-1/2}$

as the variance reduction is governed by a factor independent of N. Note that finding an appropriate h is usually challenging. We opt in Section 4.3.6 for a polynomial regression of f as our choice for h, and we estimate the coefficient c by

$$\hat{c} = \frac{\sum_{\mathbf{x}\in\mathcal{B}'} f(\mathbf{x})(h(\mathbf{x}) - \bar{h})}{\sum_{\mathbf{x}\in\mathcal{B}'}(h(\mathbf{x}) - \bar{h})^2}, \tag{2.28}$$

where $\mathcal{B}'$ is a BPP on K, which is independent of the point process used to compute $\widehat{I}_{\mathrm{MC}}$ in (2.27), 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 $\widehat{I}_{\mathrm{MCCV}}(f)$ (Owen, 2013, Example 8.3). Finally, note that the evaluation of $\hat{c}$ requires extra integrand evaluations.

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 some DPPs has been shown to enhance the convergence rate beyond $N^{-1/2}$ (Bardenet and Hardy, 2020; Coeurjolly et al., 2021)[12]. For instance, the unbiased estimator given in (Bardenet and Hardy, 2020, Section 2.1) and implemented in DPPy (Gautier et al., 2019b) is defined by[13]

$$\widehat{I}_{\mathrm{MCDPP}}(f) = \sum_{\mathbf{x}\in\mathcal{D}_N} \frac{f(\mathbf{x}/2)}{2^d \mathtt{K}_N(\mathbf{x},\mathbf{x})}, \tag{2.29}$$

where $\mathtt{K}_N$ is the N-th Christoffel-Darboux kernel associated with the Lebesgue measure on $[-1,1]^d$ and $\mathcal{D}_N$ is the multivariate OP Ensemble (MOP) of kernel $\mathtt{K}_N$; see Figure 2.6c. According to Bardenet and Hardy (2020), if f is C^1, under some mild conditions, $\widehat{I}_{\mathrm{MCDPP}}(f)$ is an unbiased estimator of $I(f)$ and its convergence rate is $N^{-1/2-1/(2d)}$, which is faster than the usual $N^{-1/2}$. One of the limitations of DPP-based methods lies in the computational complexity associated with sampling from DPPs. For instance, the computational complexity for sampling from the MOP is at least cubic in the number of points N (Gautier et al., 2019a).

[12] DPPs were previously employed in Monte Carlo integration, tracing back to Ermakov and Zolotukhin (1960). However, the formal introduction and comprehensive studies of DPPs didn't occur until the 1970s with Macchi (1975).

[13] $\widehat{I}_{\mathrm{MCDPP}}$ incorporates additional normalization factors compared to the original estimator defined in (Bardenet and Hardy, 2020, Equation 2.11). This normalization accounts for the variation in the integration support. While we set $K = [-1/2, 1/2]^d$, the original work used $K = [-1,1]^d$.

We conclude with an instance of the randomized quasi-Monte Carlo (RQMC), which is an attempt at getting both the convenience of variance statements and the error reduction of stratified deterministic quadrature. The estimator used is

$$\widehat{I}_{\mathrm{RQMC}}(f) = \frac{1}{N} \sum_{\mathbf{x} \in \mathcal{S}_N} f(\mathbf{x}), \tag{2.30}$$

where $\mathcal{S}_N$ is a randomized low-discrepancy sequence; see Section 5.1. In particular, each point of $\mathcal{S}_N$ is uniformly distributed in K, so (2.30) is an unbiased estimator of $I(f)$ (Owen, 2013, Chapter 17). 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 convergence rate of (2.30) is $\log(N)^{(d-1)/2} N^{-3/2}$ (Owen, 2008). Despite the d-dependence of the rate and the strong smoothness assumptions, $\widehat{I}_{\mathrm{RQMC}}$ is remarkably efficient in small-to-moderate dimensions, and sampling is computationally cheap compared to, e.g., DPPs. A commun choice of $\mathcal{S}_N$ is a scrambled Sobol sequence (SSS)[14] (Sobol, 1967); see Figure 2.6b.

Remark 2.2. The SSS observed in Figure 2.6b visually conveys a more regular arrangement of points compared to the BPP displayed in Figure 2.6a. This observation motivates favoring regular point processes as nodes in a Monte Carlo method. Although the MOP depicted in Figure 2.6c may not appear significantly more regular than the BPP, it is important to note that the MOP is not stationary. However, the weights in the estimator (2.29) take into account the non-stationarity inherent in the MOP. This appears when examining the weighted MOP depicted in Figure 2.6d, where the relatively more isolated points are assigned greater weight, as indicated by their larger blue disks.

[14] The Python package scipy provide the method Sobol to sample from the SSS (Virtanen et al., 2020).

3 — Diagnosing hyperuniform point processes

Contents

Condensed matter physicists have observed that some random particle systems exhibit small density fluctuations, that is, the variance of the number of points in a large window scales slower than the volume of that window, a phenomenon called *hyperuniformity* (Torquato, 2018). This statistical property, in turn, implies desirable physical properties for materials (Torquato, 2018, Section 14). Outside physics, hyperuniform point processes (HUPPs) have generated a broad interest in statistics, machine learning, and probability. In the context of our search for an efficient configuration of nodes for a Monte Carlo estimator of integrals, our focus lies on HUPPs as they inherently lead to a fast-decaying Monte Carlo error when estimating volumes. Hyperunifomity also appears in probability across many fields. The Ginibre ensemble (GPP), arguably one of the most famous determinantal point processes (DPPs) arising from random matrix theory, is a typical example of a HUPP. In stochastic geometry, and beyond DPPs, HUPPs appear

e.g., in the zeros of Gaussian analytic functions (Hough et al., 2009), and matching constructions (Klatt et al., 2020). Intriguingly, there is also empirical evidence that repeatedly applying an algorithm involving local repulsion of points to an initial point process results a HUPP (Klatt et al., 2019).

Nevertheless, it is important to note that hyperuniformity is not a standard property. The Poisson point process, for instance, is clearly not HUPP as the variance of the number of points in a window scales like the volume of the window. Consequently, HUPPs are occasionally denoted as sub-Poisson point processes. In the same vein, clustering point processes like the modified Thomas point process are typically not hyperuniform. Moreover, an independent thinning of a motion-invariant point process is not hyperuniform. Less evident, short-range Gibbs point processes are not hyperuniform (Dereudre and Flimmel, 2023).

Although the concept of hyperuniformity seems important across many scientific fields, the theoretical understanding of such systems is still limited, and proving that a point process is hyperuniform is in general a hard task. In theory, under mild assumptions, a point process in $\mathbb{R}^d$ is hyperuniform if and only if its structure factor is equal to zero at the origin (Coste, 2021).

This chapter focuses on detecting the hyperuniformity of a stationary point process by estimating its structure factor. We investigate the mathematical foundations that underlie common practices used in the literature for estimating the structure factor and testing the hyperuniformity. The primary goal is to provide statistical tools to detect HUPPs. The investigation described in this chapter was carried out in our paper (Hawat et al., 2023b), accompanied by the Python toolbox structure_factor[1]. Section 3.1 provides the needed theoretical background about HUPPs and a discussion about the usage of a HUPP in a Monte Carlo method. In Section 3.2, we rederive two families of estimators of the structure factor. The first family assumes stationarity of the underlying point process, while the second family additionally assumes isotropy. We mainly focus on studying the bias of the estimators near the origin. We showcase these estimators and compare them using the toolbox structure_factor. In Section 3.3 we review existing hyperuniformity diagnostics and contribute a test of hyperuniformity based on the debiasing techniques introduced by Rhee and Glynn (2015). Additionally, we demonstrate

[1] https://github.com/For-a-few-DPPs-more/structure-factor

the hyperuniformity diagnostics, the newly proposed statistical test, and the estimation of the hyperuniformity class using the toolbox structure_factor showing their performances and limitations. Section 3.4 presents a brief introduction to the Python toolbox structure_factor. We conclude this chapter with a few research leads in Section 3.5.

3.1 Hyperuniformity

Now, we introduce the theoretical groundwork for hyperuniformity. For more details, we refer to (Coste, 2021). In Section 3.1.1, two standard definitions of hyperuniformity are introduced. The first definition pertains to the number variance and is widely used in the literature. On the other hand, the second definition hinges on the structure factor and has been recently employed in the numerical investigations of hyperuniformity. Section 3.1.2 delves into the classification of HUPPs. Finally, Section 3.1.3 investigates the convergence rate of Monte Carlo methods with HUPPs for the class of indicator functions.

3.1.1 Hyperuniformity and the structure factor

A stationary point process $\mathcal{X}$ of $\mathbb{R}^d$ is said to be a hyperuniform point process (HUPP) (Torquato and Stillinger, 2003) (or super-homogeneous (Gabrielli et al., 2002)) if the variance of the number of points that fall in a Euclidean ball scales slower than the volume of that ball, i.e.

$$\lim_{R\to\infty} \frac{\mathbb{V}\mathrm{ar}\left[\mathcal{X}(B(\mathbf{0},R))\right]}{|B(\mathbf{0},R)|} = 0. \tag{3.1}$$

Some comments are in order. Although hyperuniformity *a priori* depends on the shape of the window, e.g., a ball in (3.1), mild technical assumptions allow to show that the definition is robust to the choice of the growing window (Coste, 2021, Section 2). Second, being hyperuniform is not a standard feature of point processes; a homogeneous Poisson process (PPP) (Section 2.4), for instance, is not hyperuniform as the ratio in (3.1) is a positive constant. Third, the most general definition of hyperuniformity goes through the structure factor S (Equation (2.8))

of a point process; under the mild assumption that $g_2 - 1$ (Equation (2.4)) is a signed measure, (3.1) is equivalent to

$$S(\mathbf{0}) = 0, \tag{3.2}$$

see (Coste, 2021, Proposition 2.2), and (Torquato, 2018, Section 5.3.1). This last definition through the structure factor, is commonly used in practical applications to detect the hyperuniformity numerically. The numerical study involves estimating the structure factor and analyzing its behavior near zero to diagnose whether the given point process is a HUPP. In Section 3.2, we will explore a variety of estimators for the structure factor, which will then be used to construct a statistical test of hyperuniformity in Section 3.3.2.

Example 3.1. The GPP (Example 2.2) and the KLY (Example 2.4) are HUPPs. The hyperuniformity of the GPP is evident from its structure factor (2.17), which is zero at the origin. On the other hand, the structure factor of the KLY is unknown, and its hyperuniformity was proven using the number variance convergence rate (3.1); see (Klatt et al., 2020).

Example 3.2. Consider a stationary point process of intensity ρ with well-defined pair correlation function g in (2.5) and structure factor S. Assuming that $g - 1$ is integrable we get

$$S(\mathbf{0}) = 1 + \rho \int_{\mathbb{R}^d} g(\mathbf{r}) - 1 \mathrm{d}\mathbf{r}.$$

If $g \geq 1$, then $S(\mathbf{0}) \geq 1$ as well, and the corresponding point process is clearly not hyperuniform. Therefore, point processes that exhibit attraction at all scales, such as the TPP (Example 2.1), are not hyperuniform. This also suggests that a HUPP exhibits repulsion at some scales. With the notations of Chapter 2, we can think of a HUPP as a regular point process with strong regularity conditions at large scales.

Example 3.3. Consider a motion-invariant point process $\mathcal{X}$ of intensity ρ. Apply a p-thinning to $\mathcal{X}$ with a retaining probability p, as described in Section 2.6.1. The structure factor of the thinned point process is given by

$$S(k) = pS_{\text{parent}}(k) + 1 - p,$$

where S_{parent} is the structure factor of $\mathcal{X}$. Consequently, $S(0) \geq 1 - p > 0$. Hence, applying a p-thinning to a motion-invariant point process will never yield

a HUPP. Even thinning a HUPP with a *large* retaining probability p will not preserve the hyperuniformity. This shows the very particular structure of HUPPs, as removing a few points from a HUPP is sufficient to break its structure. Nevertheless, thinning a HUPP is a way to generate a parametrized non-HUPP with $S(0)$ controlled by p. In particular, large values of p result $S(0)$ arbitrarily close to zero, which is very useful for assessing the accuracy of a hyperuniformity test (Kim and Torquato, 2018; Coste, 2021; Klatt et al., 2022).

3.1.2 Hyperuniformity classes

For a HUPP, when the structure factor undergoes a power decay $|S(\mathbf{k})| \sim c\|\mathbf{k}\|_2^{\alpha}$ in the neighborhood of zero, the point process can be classified into three categories depending on how α compares to 1, as summarized in Table 3.1 (Torquato, 2018, Section 5.3.2).

Table 3.1: Classes of hyperuniformity

Class	α	$\mathbb{V}\mathrm{ar}\left[\mathcal{X}(B(0,R))\right]$
I	> 1	$O(R^{d-1})$
II	$= 1$	$O(R^{d-1}\log(R))$
III	$\in (0,1)$	$O(R^{d-\alpha})$

Since the variance cannot grow more slowly than R^{d-1} for a spherical window (Beck, 1987), Class I represents the *strongest* form of hyperuniformity. It includes, for instance, the GPP (Example 2.2), with a power decay $\alpha = 2$, as observed from its structure factor expression (2.17). Additionally, Klatt et al. (2020) demonstrated that KLY (Example 2.4) also falls in Class I. Class II includes, for instance, the Sine process, a central object in random matrix theory (Anderson et al., 2009). By contrast, systems that fall in Class III present the *weakest* form of hyperuniformity. Many stationary point processes are believed to be Class III, although few rigorous proofs exist; one can mention the asymptotic result of (Boursier, 2021) for the one-dimensional Riesz gaz.

3.1.3 Monte Carlo with a hyperuniform point process

We mentioned earlier that our main interest in HUPPs is that by definition (3.1) a HUPP yields a Monte Carlo method with a fast convergence rate when estimating volumes. Although this may seem straightforward, it deserves to be articulated given the that concept of hyperuniformity involves enlarging the observation window while the convergence rate of a Monte Carlo method relies on increasing the number of evaluation points.

Consider a Class I HUPP $\mathcal{X} \subset \mathbb{R}^d$ of intensity $\rho_{\mathcal{X}}$. Let f be the function defined by $f(\mathbf{x}) = \mathbb{1}_K(\mathbf{x})$, where K is a compact of $\mathbb{R}^d$ of which we want to estimate the volume $I(f)$ in (2.21) using *approximately* N evaluations of f. We will use a scaled version of $\mathcal{X}$

$$\mathcal{Y} \triangleq (\rho_{\mathcal{X}}|W|)^{1/d} N^{-1/d} \mathcal{X}, \tag{3.3}$$

where W is another compact of $\mathbb{R}^d$ of known volume $|W|$ such that $K \subset W$. As $\mathcal{Y}$ is a stationary point process of intensity $\rho_{\mathcal{Y}} = N/|W|$, in particular,

$$\widehat{I}_{\mathcal{Y}\cap W}(f) \triangleq \frac{|W|}{N} \sum_{\mathbf{x}\in\mathcal{Y}\cap W} f(\mathbf{x}) \tag{3.4}$$

is an unbiased Monte Carlo estimator of $I(f)$ with an *expected* number of evaluation points equal to $\mathbb{E}[\mathcal{Y}(W)] = N$. Moreover, as $\mathcal{X}$ is a Class I HUPP for any $R > 0$ we have

$$\mathbb{V}\mathrm{ar}\left[\sum_{\mathbf{x}\in\mathcal{X}} \mathbb{1}_{B(\mathbf{0},R)}(\mathbf{x})\right] = O\left(R^{d-1}\right), \tag{3.5}$$

and under mild technical assumptions on $\mathcal{X}$, (3.5) is also valid for diverse window shapes; see Coste (2021). We will study the convergence rate of (3.4) for $K = B(\mathbf{0}, R)$ keeping in mind that this can be generalized to other window-shapes. We have

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\mathcal{Y}\cap W}(f)\right] = \frac{|W|^2}{N^2} \mathbb{V}\mathrm{ar}\left[\sum_{\mathbf{x}\in\mathcal{X}} \mathbb{1}_{B\left(\mathbf{0}, \frac{N^{1/d}}{(|W|\rho_{\mathcal{X}})^{1/d}} R\right)}(\mathbf{x})\right] \tag{3.6}$$

Inserting (3.5) in (3.6) we get

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\mathcal{Y}\cap W}(f)\right] = O\left(N^{-1-1/d}\right). \tag{3.7}$$

Some comments are in order. First, we stress that the number of evaluation points here is random and N is its expected value. Second, by Beck (1987) the

variance in (3.5) cannot grow more slowly than R^{d-1}; see also (Coste, 2021). Thus, $N^{-1/2-1/(2d)}$ is the optimal convergence rate that a Monte Carlo method based on a stationary point process, like (3.4), can achieve for the class of indicator functions. Third, by replacing $\mathcal{X}$ with a HUPP of Class II and employing the same methodology as before, the convergence rate of $\widehat{I}_{\mathcal{Y}\cap W}(f)$ becomes $N^{-1/2-1/(2d)}\log(N)^{1/2}$. Similarly for a Class III with a power decay $\alpha \in (0,1)$, the convergence rate becomes $N^{-1/2-\alpha/(2d)}$. Now it comes to light that a Monte Carlo method with any HUPP outperforms the $N^{-1/2}$ convergence rate of the crude Monte Carlo. Finally, although it might seem intuitive to anticipate a similar fast convergence rate for classes of functions beyond indicator functions, this inference is not straightforward.

3.2 Estimators of the structure factor

Starting from the theoretical definition of the structure factor S (2.8), in Sections 3.2.1 and 3.2.2, we derive all known estimators of S and add a few other natural candidates based on numerical quadratures of the symmetric Fourier transform. We pay particular attention to the sources of asymptotic (in the size of the observation window) bias for each estimator. The estimators are categorized based on whether they assume stationarity or both stationarity and isotropy of the point process. Then, Section 3.2.3 demonstrates the estimators using the Python toolbox structure_factor, while Section 3.2.4 provides a more quantitative comparison of their costs and accuracies.

The closest work to Sections 3.2.1 and 3.2.2 is the paper by Rajala, Olhede, and Murrell (2020). They introduce new estimators of the structure factor based on the idea of tapering in time series analysis (Percival and Walden, 2020), and investigate central limit theorems for their estimators. On our end, we limit ourselves to a survey – including the estimators of Rajala et al. (2020) – and simpler properties like asymptotic unbiasedness and its relation to implicit implementation choices in statistical physics papers. One reason for this is our motivation for the study of hyperuniformity: HUPPs are unlikely to satisfy the assumptions[2] behind the

[2]E.g., Hypothesis (H4) of Biscio and Waagepetersen (2019), when the linear statistic is the number of points, contradicts hyperuniformity.

central limit theorems referenced by Rajala et al. (2020). Moreover, our survey includes a broader choice of estimators, including numerical quadratures of Hankel transforms. Overall, Sections 3.2.1, 3.2.2 and the paper of Rajala et al. (2020) may be complementary [3].

3.2.1 Estimators assuming stationarity

The most common estimator of the structure factor is the so-called *scattering intensity* (Torquato, 2018; Klatt et al., 2019, 2020; Coste, 2021). Its name stems from its origins in the physics of diffraction, and it corresponds to a scaled version of Bartlett's periodogram in time series analysis. After introducing the scattering intensity, we follow Rajala et al. (2020), who generalize it to so-called *tapered* estimators. We group our assumptions and notations in Assumption A1.

Assumption A1. *$\mathcal{X}$ is a stationary simple point process of $\mathbb{R}^d$ with intensity $\rho > 0$. Its pair correlation function g exists, and $\mathbf{r} \mapsto g(\mathbf{r}) - 1$ is integrable on $\mathbb{R}^d$. Moreover, we only observe a realization of*

$$\mathcal{X} \cap W = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$$

in a centered, rectangular window $W = \prod_{j=1}^{d} [-L_j/2, L_j/2]$ of volume $|W|$. We write $\mathbf{L} = (L_1, \dots, L_d)$.

The scattering intensity

In the physics literature, the following derivation is often assumed to be known to the reader, and it seemed worthwhile to us to make it explicit. This allows, in particular, to understand the role played by the so-called *allowed wavevectors*. Note that Rajala et al. (2020) provide a similar derivation.

The basic idea is to introduce the *scaled intersection volume* α_0

$$\alpha_0(\mathbf{r}, W) = \frac{1}{|W|} \int_{\mathbb{R}^d} \mathbb{1}_W(\mathbf{r} + \mathbf{y}) \mathbb{1}_W(\mathbf{y}) d\mathbf{y}, \tag{3.8}$$

[3] We note that during the reviewing process of our paper (Hawat et al., 2023b), which forms the basis of this chapter, another version of the preprint (Rajala et al., 2020) has been published (Rajala et al., 2023). The changes in the new version do not seem to impact the research we present in this chapter.

in the definition (2.8) of the structure factor. We obtain

$$\begin{aligned} S(\mathbf{k}) &= 1 + \rho \int_{\mathbb{R}^d} (g(\mathbf{r}) - 1) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{r} \rangle} \, \mathrm{d}\mathbf{r} \\ &= 1 + \rho \int_{\mathbb{R}^d} \lim_{W \uparrow \mathbb{R}^d} (g(\mathbf{r}) - 1) \alpha_0(\mathbf{r}, W) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{r} \rangle} \, \mathrm{d}\mathbf{r} \\ &= 1 + \lim_{W \uparrow \mathbb{R}^d} \rho \int_{\mathbb{R}^d} (g(\mathbf{r}) - 1) \alpha_0(\mathbf{r}, W) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{r} \rangle} \, \mathrm{d}\mathbf{r}, \end{aligned} \tag{3.9}$$

where we used dominated convergence and the limit $\lim_{W \uparrow \mathbb{R}^d} \alpha_0(\mathbf{r}, W) = 1$. In the notation $W \uparrow \mathbb{R}^d$, the limit is taken so that the window progressively covers the whole space at roughly equal speed in all directions, i.e., $\min_j L_j \to \infty$ in Assumption A1. Splitting the integral in (3.9), we shall recognize an expectation under our censored point process $\mathcal{X} \cap W$ and a bias term,

$$S(\mathbf{k}) - 1 = \lim_{W \uparrow \mathbb{R}^d} \left[\frac{\rho}{|W|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \underbrace{e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{r} \rangle} \mathbb{1}_{W \times W}(\mathbf{r} + \mathbf{y}, \mathbf{y})}_{f(\mathbf{r}+\mathbf{y}, \mathbf{y})} g(\mathbf{r}) \mathrm{d}\mathbf{y} \mathrm{d}\mathbf{r} - \underbrace{\rho \mathcal{F}(\alpha_0)(\mathbf{k}, W)}_{= \epsilon_0(\mathbf{k}, \mathbf{L})} \right].$$

Now, by definition (2.4) of the pair correlation measure, and still for any $\mathbf{k} \in \mathbb{R}^d$, $S(\mathbf{k}) - 1$ is the limit as $W \uparrow \mathbb{R}^d$ of

$$\frac{\rho}{\rho^2 |W|} \mathbb{E} \left[\sum_{\mathbf{x}, \mathbf{y} \in \mathcal{X}}^{\mathbf{x} \neq \mathbf{y}} \mathbb{1}_W(\mathbf{x}) \mathbb{1}_W(\mathbf{y}) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x} - \mathbf{y} \rangle} \right] - \epsilon_0(\mathbf{k}, \mathbf{L}),$$

so that

$$\begin{aligned} S(\mathbf{k}) &= \lim_{W \uparrow \mathbb{R}^d} \frac{1}{\rho |W|} \mathbb{E} \left[\sum_{\mathbf{x}, \mathbf{y} \in \mathcal{X} \cap W} e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x} - \mathbf{y} \rangle} \right] - \epsilon_0(\mathbf{k}, \mathbf{L}) \\ &= \lim_{W \uparrow \mathbb{R}^d} \frac{1}{\rho |W|} \mathbb{E} \left[\left| \sum_{\mathbf{x} \in \mathcal{X} \cap W} e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x} \rangle} \right|^2 \right] - \epsilon_0(\mathbf{k}, \mathbf{L}). \end{aligned} \tag{3.10}$$

Note that, by (A.7), the bias term satisfies

$$|\epsilon_0(\mathbf{k}, \mathbf{L})| = \rho \left(\prod_{j=1}^d \frac{\sin(k_j L_j / 2)}{\sqrt{L_j} k_j / 2} \right)^2 \leq \begin{cases} 0 & \text{if } \mathbf{k} \in \mathbb{A}_{\mathbf{L}}, \\ \rho \prod_{j=1}^d L_j & \text{as } \|\mathbf{k}\|_2 \to 0, \\ 2^{2d} \prod_{j=1}^d \frac{1}{L_j k_j^2} & \text{as } \|\mathbf{k}\|_2 \to \infty, \end{cases}$$

where we defined

$$\mathbb{A}_{\mathbf{L}} = \left\{ \mathbf{k} \in (\mathbb{R}^*)^d \text{ such that } k_j = \frac{2\pi n}{L_j} \text{ for some } j \in \{1, \dots, d\} \text{ and some } n \in \mathbb{Z}^* \right\}. \tag{3.11}$$

We have thus proved the following proposition.

Proposition 3.1. *Under Assumption A1, for* $\mathbf{k} \in \mathbb{A}_{\mathbf{L}}$*, the scattering intensity estimator*

$$\widehat{S}_{\mathrm{SI}}(\mathbf{k}) \triangleq \frac{1}{\rho|W|}\left|\sum_{j=1}^{N} e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x}_j\rangle}\right|^2 \tag{3.12}$$

is asymptotically unbiased, i.e.,

$$\sup_{\mathbf{k}\in\mathbb{A}_{\mathbf{L}}} \left|\mathbb{E}[\widehat{S}_{\mathrm{SI}}(\mathbf{k})] - S(\mathbf{k})\right| \xrightarrow[W\uparrow\mathbb{R}^d]{} 0.$$

This motivates restricting to $\mathbf{k} \in \mathbb{A}_{\mathbf{L}}$ if one is interested in estimating the behavior of S in zero. In the literature, the scattering intensity is actually often evaluated on a subset of $\mathbb{A}_{\mathbf{L}}$, namely

$$\mathbb{A}_{\mathbf{L}}^{\mathrm{res}} = \left\{\left(\frac{2\pi n_1}{L_1}, \ldots, \frac{2\pi n_d}{L_d}\right); \mathbf{n} \in (\mathbb{Z}^*)^d\right\}. \tag{3.13}$$

The set $\mathbb{A}_{\mathbf{L}}^{\mathrm{res}}$ is called the set of *allowed wavevectors* in physics (Klatt et al., 2020, Section 10), or part of the *dual lattice* of fundamental cell W in sampling theory (Osgood, 2014, Chapter 5), or *Fourier grid* in time series analysis (Rajala et al., 2020). We add the superscript *res* to underline that it is actually a restriction of the set $\mathbb{A}_{\mathbf{L}}$ of wavevectors justified by the cancellation of the asymptotic bias.

It is unclear to us why one should consider $\mathbb{A}_{\mathbf{L}}^{\mathrm{res}}$ instead of $\mathbb{A}_{\mathbf{L}}$. In particular, the minimal *wavenumber* (norm of a wavevector) in $\mathbb{A}_{\mathbf{L}}^{\mathrm{res}}$ is

$$k_{\min}^{\mathrm{res}} = 2\pi\sqrt{\sum_{j=1}^{d} L_j^{-2}},$$

while working with $\mathbb{A}_{\mathbf{L}}$ in (3.11) relaxes this threshold to a $k_{\min}$ satisfying

$$\frac{2\pi}{\max_j L_j} < k_{\min} < k_{\min}^{\mathrm{res}}.$$

To see how far we can hope going down in $k > 0$, we give evidence in Remark 3.1 as to why the scattering intensity should never be evaluated for $\|\mathbf{k}\|_2 \leq \frac{\pi}{\sqrt{d}\max_j L_j}$.

Finally, when ρ is unknown, and $\mathcal{X}$ is further assumed to be ergodic, it is common to replace the denominator $\rho|W|$ by N in (3.12), leading to the self-normalized scattering intensity estimator

$$\widehat{S}_{\mathrm{SI,s}}(\mathbf{k}) \triangleq \frac{1}{N}\left|\sum_{j=1}^{N} e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x}_j\rangle}\right|^2. \tag{3.14}$$

Indeed, by ergodicity,

$$\frac{\widehat{S}_{\mathrm{SI,s}}}{\widehat{S}_{\mathrm{SI}}} \xrightarrow[W\uparrow\mathbb{R}^d]{\text{a.s.}} 1.$$

Many authors define the scattering intensity as its self-normalized version (Torquato, 2018; Klatt et al., 2019; Coste, 2021), but we argue that when ρ is known, it is not clear whether this estimator has a smaller mean squared error than (3.12); see also Remark 4.1 about self-normalized estimators.

For readers interested in the second-order properties of this estimator, its second moment for a PPP input is derived in Section 3.3.2.

Remark 3.1 (A lower bound for $k_{\min}$). It is clear that when W is fixed, the scattering intensity (3.14) is not relevant when $\mathbf{k}$ is too close to 0, since for a fixed sample $\mathcal{X}_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$,

$$\lim_{\mathbf{k}\to 0} \widehat{S}_{\mathrm{SI,s}}(\mathbf{k}) = N.$$

The theoretical convergence to 0 for HUPPs is dictated by the compensation occurring between exponential terms $e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x}_j\rangle}$ on large portions of the space. One can safely infer that if a large portion of the terms gives a positive contribution, then the compensation does not occur and the global result will not be accurate. Still with the notation in Assumption A1, if $\|\mathbf{k}\|_2 < \frac{\pi}{\sqrt{d}\max_j L_j}$, then for $\mathbf{x} \in W$, $|\langle \mathbf{x}, \mathbf{k}\rangle| \leq \pi/2$, so that there exists $\epsilon > 0$ such that

$$\frac{1}{N}\left|\sum_{j=1}^{N} e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x}_j\rangle}\right|^2 \geq \frac{(N\epsilon)^2}{N}.$$

Thus for any $\|\mathbf{k}\|_2 < \frac{\pi}{\sqrt{d}\max_j L_j}$, we have a lower bound of $\widehat{S}_{\mathrm{SI,s}}(\mathbf{k})$ that is independent of the point process, and which diverges as the number of points goes to infinity. Consequently, we argue that $\pi/\sqrt{d}\max_j L_j$ is a lower bound for accessible wavenumbers, which one might improve with a finer study of the estimator bias.

Tapered variants of the scattering intensity

The derivation of the scattering intensity can be generalized to the *tapered estimator*

$$\widehat{S}_{\mathrm{T}}(t, \mathbf{k}) \triangleq \frac{1}{\rho}\left|\sum_{j=1}^{N} t(\mathbf{x}_j, W) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x}_j\rangle}\right|^2, \quad \mathbf{k} \in \mathbb{R}^d, \tag{3.15}$$

where $t(\cdot, W)$ is a uniformly (in W) square-integrable function supported on the observation window W, called a *taper*. The tapered estimator (3.15) corresponds to a scaled version of what is called a tapered periodogram in the signal processing literature (Rajala et al., 2020, Section 3); see also (Daley and Vere-Jones, 2003). The vocabulary is adapted from the spectral analysis of time series, where tapers are now well established (Percival and Walden, 2020).

In particular, one recovers the scattering intensity (3.12) from the tapered formulation (3.15) by plugging the taper

$$t_0(\mathbf{x}, W) \triangleq \frac{1}{\sqrt{|W|}} \mathbb{1}_W(\mathbf{x}). \tag{3.16}$$

To follow the derivation of (3.10), we further let

$$\alpha_t(\mathbf{r}, W) \triangleq \int_{\mathbb{R}^d} t(\mathbf{r}+\mathbf{y}, W) t(\mathbf{y}, W) \mathrm{d}\mathbf{y}, \tag{3.17}$$

and require that

$$\lim_{W \uparrow \mathbb{R}^d} \alpha_t(\mathbf{r}, W) = 1 \quad \text{and} \quad \alpha_t(\mathbf{0}, W) = 1, \tag{3.18}$$

where the limit is again taken as $\min_j L_j \to \infty$ under Assumption A1. Note first that our requirement that (3.18) holds differs from the treatment of Rajala et al. (2020). We find (3.18) to be a more natural generalization of the scattering intensity arguments of Torquato (2018). Note also that, for simplicity, we denoted α_{t_0} by α_0 in Equation (3.8), and we shall stick to this simplified notation.

Now, Cauchy-Schwarz inequality and the uniform integrability of t guarantee that $\alpha_t(\cdot, W)$ is uniformly (in W) bounded, so that by dominated convergence,

$$S(\mathbf{k}) = 1 + \rho \lim_{W \uparrow \mathbb{R}^d} \int_{\mathbb{R}^d} (g(\mathbf{r}) - 1) \alpha_t(\mathbf{r}, W) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{r} \rangle} \, \mathrm{d}\mathbf{r}.$$

Following the lines of (3.10), we obtain

$$S(\mathbf{k}) = \lim_{W \uparrow \mathbb{R}^d} \left(\mathbb{E}\left[\widehat{S}_{\mathrm{T}}(t, \mathbf{k}) \right] - \underbrace{\rho \, |\mathcal{F}(t)(\mathbf{k}, W)|^2}_{\epsilon_t(\mathbf{k}, \mathbf{L})} \right), \tag{3.19}$$

where the tapered estimator $\widehat{S}_{\mathrm{T}}$ is defined in (3.15). To eliminate the asymptotic bias $\epsilon_t(\mathbf{k}, \mathbf{L})$, one can restrict oneself again to a set of allowed wavevectors as we did in (3.11), i.e. the zeros of $\epsilon_t(\cdot, \mathbf{L})$. For general tapers, however, finding these

zeros is not straightforward, and an alternative way to escape the bias is to correct it, as in

$$\widehat{S}_{\mathrm{UDT}}(t,\mathbf{k}) \triangleq \frac{1}{\rho}\left|\sum_{j=1}^{N} t(\mathbf{x}_j,W)e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}_j\rangle}\right|^2 - \rho\,|\mathcal{F}(t)(\mathbf{k},W)|^2\,. \tag{3.20}$$

We refer to $\widehat{S}_{\mathrm{UDT}}$ as the *undirectly debiased tapered estimator*, which is a scaled version of what Rajala et al. (2020) define. The major issue of Estimator (3.20) is that it may give negative values. To remedy this, Rajala et al. (2020) propose to remove the bias inside the summation before taking the squared modulus, namely, to define the *directly debiased tapered estimator*

$$\widehat{S}_{\mathrm{DDT}}(t,\mathbf{k}) \triangleq \frac{1}{\rho}\left|\sum_{j=1}^{N} t(\mathbf{x}_j,W)e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}_j\rangle} - \rho\mathcal{F}(t)(\mathbf{k},W)\right|^2. \tag{3.21}$$

A straightforward computation gives that the estimator $\widehat{S}_{\mathrm{DDT}}$ is also asymptotically unbiased. We only need to observe that using Equation (2.11) we get

$$\begin{aligned}\mathbb{E}\left[\sum_{j=1}^{N} t(\mathbf{x}_j,W)e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}_j\rangle}\overline{\mathcal{F}(t)(\mathbf{k},W)}\right] &= \mathbb{E}\left[\overline{\sum_{j=1}^{N} t(\mathbf{x}_j,W)e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}_j\rangle}}\,\mathcal{F}(t)(\mathbf{k},W)\right]\\ &= \rho|\mathcal{F}(t)(\mathbf{k},W)|^2.\end{aligned}$$

Finally, note that these debiasing techniques naturally apply to the special case of the scattering intensity, and thus offer an alternative to using allowed values (3.11).

The multitapered estimator

In the spectral analysis of time series, *multitapering* was first introduced by Thomson (1982); see also Percival and Walden (2020) for a modern reference. The idea is to average a periodogram over many tapers, in the hope to reduce the variance of the resulting estimator. Rajala et al. (2020) propose to adapt the method to point processes, and we follow their lines.

For any $\mathbf{k} \in \mathbb{R}^d$ and $P \in \mathbb{N}^*$, and under Assumption A1, Rajala et al. (2020) define the *multitapered estimator* $\widehat{S}_{\mathrm{MT}}$ by

$$\widehat{S}_{\mathrm{MT}}((t_q)_{q=1}^{P},\mathbf{k}) = \frac{1}{P}\sum_{q=1}^{P}\widehat{S}(t_q,\mathbf{k}), \tag{3.22}$$

where the P tapers $(t_q)_{q=1}^P$ are typically taken to be pairwise orthogonal square-integrable functions, and $\widehat{S}(t_q, \cdot)$ is any of the tapered estimators, whether undebiased (3.15), undirectly debiased (3.20), or directly debiased (3.21). The directly and undirectly debiased versions of $\widehat{S}_{\mathrm{MT}}$ will be respectively denoted by $\widehat{S}_{\mathrm{DDMT}}$ and $\widehat{S}_{\mathrm{UDMT}}$.

On the choice of tapers

Common taper choices in time series analysis are Slepian tapers, sinusoidal tapers, and minimum bias tapers (Riedel and Sidorenko, 1995). For instance, still assuming a centered rectangular window $W = \prod_{j=1}^d [-L_j/2, L_j/2]$, the family of sinusoidal tapers $(t_q)_{q\geq 1}$ supported on W is defined by

$$t_q(\mathbf{x}, W) = t(\mathbf{x}, \mathbf{p}^q, W) \triangleq \frac{\mathbb{1}_W(\mathbf{x})}{\sqrt{|W|}} \prod_{j=1}^d \sqrt{2} \sin\left(\frac{\pi p_j^q}{L_j}\left(x_j + \frac{L_j}{2}\right)\right), \tag{3.23}$$

where $\mathbf{p}^q = (p_1^q, \ldots, p_d^q) \in (\mathbb{N}^d)^*$ and $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$. The sinusoidal tapers are pairwise orthogonal, and an easy direct computation shows that they also satisfy (3.18). Moreover, for $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{R}^d$ the Fourier transform $\mathcal{F}(t_q)(\mathbf{k}, W)$ of t_q for any q is

$$\frac{1}{\sqrt{|W|}} \prod_{j=1}^d \sqrt{2} i^{(p_j^q+1)} \left[\frac{\sin\left(\left(k_j - \frac{\pi p_j^q}{L_j}\right)\frac{L_j}{2}\right)}{k_j - \frac{\pi p_j^q}{L_j}} - (-1)^{p_j^q} \frac{\sin\left(\left(k_j + \frac{\pi p_j^q}{L_j}\right)\frac{L_j}{2}\right)}{k_j + \frac{\pi p_j^q}{L_j}} \right].$$

This closed-form expression can thus be used in any debiasing scheme. Note that this analytical tractability, along with the absence of a sensitive parameter like a lengthscale, lead us to choose the sine taper over, say, a multidimensional generalization of Slepian tapers (Percival and Walden, 2020; Rajala et al., 2020).

3.2.2 Estimators assuming stationarity and isotropy

For isotropic point processes, a common approach is simply to numerically rotation average the structure factor estimators presented in Section 3.2.1. Alternatively, one could start from the analytical expression (2.9) of the structure factor as symmetric Fourier transform – a univariate integral – involving the pair correlation function. Then again, two approaches have been identified. First, identifying

an expectation under the point process as in Section 3.2.1 leads to a natural estimator originally derived by Bartlett (1964). Second, estimation of the pair correlation followed by numerical quadrature leads to at least two natural estimators, depending on whether the quadrature is that of Ogata (2005), or that of Baddour and Chouinard (2015). We review all these estimators in turn. All point processes in this section satisfy Assumption A2.

Assumption A2. *$\mathcal{X}$ is a simple stationary isotropic point process of $\mathbb{R}^d$, of intensity ρ. Its pair correlation function g exists, and $r \mapsto g(r) - 1$ is integrable. Moreover, we only observe a realization of $\mathcal{X} \cap W = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ in the centered ball $W = B^d(\mathbf{0}, R)$.*

Bartlett's isotropic estimator

From the observation that the scaled intersection volume α_0 (3.8) is a radial function, $\alpha_0(\mathbf{r}, W) = \alpha_0(r, W)$, and following the lines of Section 3.2.1, dominated convergence theorem yields

$$S(k) - 1 = \lim_{R\to\infty} \left[\rho \frac{(2\pi)^{\frac{d}{2}}}{k^{\frac{d}{2}-1}} \int_0^\infty \alpha_0(r, W)\, g(r) r^{\frac{d}{2}} J_{d/2-1}(kr) \mathrm{d}r - \underbrace{\rho \mathcal{F}_s(\alpha_0)(k, W)}_{\triangleq \epsilon_1(k,R)} \right]. \tag{3.24}$$

Now, precisely because α_0 is radial, we have

$$\alpha_0(r, W) = \frac{1}{\omega_{d-1}} \int_{S(\mathbf{0},1)} \alpha_0(r\mathbf{u}, W) \mathrm{d}\mathbf{u}, \tag{3.25}$$

where $\mathrm{d}\mathbf{u}$ is the $(d-1)$-dimensional Hausdorff measure and $S(\mathbf{0}, 1)$ is the unit sphere of $\mathbb{R}^d$, with surface area ω_{d-1}. Plugging (3.25) into (3.24) yields

$$\begin{aligned}
S(k) - 1 = & \\
& \lim_{R\to\infty} \frac{\rho (2\pi)^{\frac{d}{2}} k^{1-\frac{d}{2}}}{|W| \omega_{d-1}} \int_0^\infty \int_{S(\mathbf{0},1)} \int_{\mathbb{R}^d} r^{\frac{d}{2}} J_{d/2-1}(kr) g(r) \mathbb{1}_{W\times W}(r\mathbf{u} + \mathbf{y}, \mathbf{y}) \, \mathrm{d}\mathbf{y}\mathrm{d}\mathbf{u}\mathrm{d}r - \epsilon_1(k, R) = \\
& \lim_{R\to\infty} \frac{\rho (2\pi)^{\frac{d}{2}}}{|W| \omega_{d-1}} \int_0^\infty \int_{S(\mathbf{0},1)} \int_{\mathbb{R}^d} \frac{J_{d/2-1}(kr)}{(kr)^{d/2-1}} g(r) \mathbb{1}_{W\times W}(r\mathbf{u} + \mathbf{y}, \mathbf{y}) \, \mathrm{d}\mathbf{y}\mathrm{d}\mathbf{u}\, r^{d-1} \mathrm{d}r - \epsilon_1(k, R) = \\
& \lim_{R\to\infty} \frac{\rho (2\pi)^{\frac{d}{2}}}{|W| \omega_{d-1}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{W\times W}(\mathbf{r} + \mathbf{y}, \mathbf{y}) \frac{J_{d/2-1}(k\|\mathbf{r}\|_2)}{(k\|\mathbf{r}\|_2)^{\frac{d}{2}-1}} g(\|\mathbf{r}\|_2) \, \mathrm{d}\mathbf{y}\mathrm{d}\mathbf{r} - \epsilon_1(k, R).
\end{aligned}$$

We now recognize an expectation using (2.4), so that $S(k) - 1$ rewrites as

$$\lim_{R\to\infty} \frac{(2\pi)^{\frac{d}{2}}}{\rho |W| \omega_{d-1}} \mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}\cap W}^{\mathbf{x}\neq\mathbf{y}} \frac{J_{d/2-1}(k\|\mathbf{x} - \mathbf{y}\|_2)}{(k\|\mathbf{x} - \mathbf{y}\|_2)^{d/2-1}} \right] - \epsilon_1(k, R). \tag{3.26}$$

We thus define a new estimator

$$\widehat{S}_{\mathrm{BI}}(k) = 1 + \frac{(2\pi)^{\frac{d}{2}}}{\rho|W|\omega_{d-1}} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{J_{d/2-1}(k\|\mathbf{x}_i - \mathbf{x}_j\|_2)}{(k\|\mathbf{x}_i - \mathbf{x}_j\|_2)^{d/2-1}}, \tag{3.27}$$

along with its self-normalized version

$$\widehat{S}_{\mathrm{BI,s}}(k) = 1 + \frac{(2\pi)^{\frac{d}{2}}}{N\omega_{d-1}} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{J_{d/2-1}(k\|\mathbf{x}_i - \mathbf{x}_j\|_2)}{(k\|\mathbf{x}_i - \mathbf{x}_j\|_2)^{d/2-1}}, \tag{3.28}$$

as in the case of the self-normalized scattering intensity (3.14). When $d = 2$, $\widehat{S}_{\mathrm{BI,s}}$ corresponds to Bartlett's isotropic estimator (Bartlett, 1964).

Here also, there are two sources of bias in the estimator (3.27). The first one is due to the restriction of the point process to a bounded observation window, which shall disappear as the window grows. The second source of bias $\epsilon_1(k, R)$ is again related to the Fourier transform of the scaled intersection volume $\alpha_0(\cdot, W)$. Diggle et al. (1987) observed that $|\epsilon_1(k, R)|$ is larger when $k > 0$ is small, and proposed to artificially set the value of the estimator to some constant when k is smaller than a certain threshold (Diggle et al., 1987, Equation 3.4). Obviously, this correction is inadequate to study hyperuniformity, i.e., the behavior of S near zero.

An alternative to the clipping procedure is to proceed as done for the scattering intensity (3.12) and estimate the structure factor only at a set of allowed wavenumbers, defined as the zeros of $\epsilon_1(\cdot, R)$. Using (A.6), it comes, for fixed d,

$$\begin{aligned}
\epsilon_1(k, R) &= \mathcal{F}_s(\alpha_0)(k) \\
&= 2^d \pi^{d/2} \frac{\Gamma(d/2+1)}{k^d} J^2_{\frac{d}{2}}(kR) \\
&= \begin{cases} 0 & \text{if } k \in \{\frac{x}{R}; J_{\frac{d}{2}}(x) = 0\}, \\ O(R^d) & \text{as } k \to 0, \\ O\left(\frac{1}{k^d (Rk)^{2/3}}\right) & \text{as } k \to \infty. \end{cases}
\end{aligned}$$

The two bounds respectively come from the fact that $J_\nu(x) \sim \frac{1}{\Gamma(\nu+1)}(\frac{x}{2})^\nu$ in the neighborhood of zero, and that for all $\nu > 0$ and $x \in \mathbb{R}$, $|J_\nu(x)| \leq c|x|^{-1/3}$ (with $c \approx 0.8$) (Landau, 2000). Thus, for the estimator (3.27), we let the set of allowed wavenumbers associated with the window $W = B^d(\mathbf{0}, R)$ be

$$\mathbb{A}_R = \left\{ \frac{x}{R} \in \mathbb{R};\ J_{d/2}(x) = 0 \right\}. \tag{3.29}$$

Proposition 3.2. *Under Assumption A2, for $k \in \mathbb{A}_R$, the estimator $\widehat{S}_{\mathrm{BI}}$ is asymptotically unbiased, i.e.,*

$$\sup_{k \in \mathbb{A}_R} \left| \mathbb{E}\left[\widehat{S}_{\mathrm{BI}}(k)\right] - S(k) \right| \xrightarrow[R\to\infty]{} 0.$$

Note that one can also define debiased tapered and multitapered versions of Bartlett's estimator, as done in Section 3.2.1, but the choice of the taper(s) requires more attention, as they must be separable and radial.

Using Ogata's quadrature

Still working under Assumption A2, we can define alternative estimators of the structure factor (2.9) by first approximating the pair correlation function from a realization of $\mathcal{X}$, and then approximating the Hankel transform (A.4).

Estimators of the pair correlation function have been thoroughly investigated; see (Baddeley et al., 2015). In a nutshell, they divide in, on one side, numerical derivatives of estimates of Ripley's K function, and on the other side, direct kernel density estimators based on the collection of pairwise distances in the sample. Both families come with sophisticated edge correction techniques, and, at least for small sample sizes, it seems reasonable to build on this previous work. Henceforth, we assume that an estimator of the pair correlation function is available, and defer the discussion of which estimator to use to Section 3.2.3.

It remains to perform a numerical quadrature on a Hankel transform. Ogata (2005) approximates integrals of the form

$$\mathcal{I}_\nu(f) = \int_0^\infty f(x) J_\nu(x) \mathrm{d}x$$

as

$$\pi \sum_{j=1}^{\infty} w_{\nu,j} f\left(\frac{\pi}{h}\psi(h\xi_{\nu,j})\right) J_\nu\left(\frac{\pi}{h}\psi(h\xi_{\nu,j})\right)\psi'(h\xi_{\nu,j}), \tag{3.30}$$

with $w_{\nu,j} = \frac{Y_\nu(\pi\xi_{\nu,j})}{J_{\nu+1}(\pi\xi_{\nu,j})}$ and $\psi(t) = t \times \tanh(\frac{\pi}{2}\sinh(t))$. Y_ν is the Bessel function of the second kind of order ν, h is a positive constant called the *stepsize*, and $(\xi_{\nu,j})_{j\geq 1}$ are the positive zeros of the Bessel function $J_\nu(\pi x)$ of the first kind of order ν, arranged in increasing order. In practice, the infinite sum on the right-hand side

of (3.30) can be truncated at a small number of function evaluations since the quadrature nodes approach the zeros of $J_\nu(x)$, that is $\frac{\pi}{h}\psi(h\xi_{\nu,j}) \sim \pi\xi_{\nu,j}$, very fast as $j \to \infty$.

Ogata's quadrature applies to the Hankel transform (A.3) of an integrable function f, since

$$\mathcal{H}_\nu(f)(k) = \mathcal{I}_\nu\left(x \mapsto \frac{x}{k^2} f\left(x/k\right)\right). \tag{3.31}$$

In particular, it applies to the computation of structure factors since, by (2.9) and (A.4),

$$\begin{aligned} S(k) &= 1 + \rho\mathcal{F}_s(g-1)(k) \\ &= 1 + \rho\frac{(2\pi)^{d/2}}{k^{d/2-1}}\mathcal{I}_{d/2-1}\left(x \mapsto \frac{x^{d/2}}{k^{d/2+1}} g(x/k) - 1\right). \end{aligned}$$

We thus define the Hankel-Ogata estimator of the structure factor as

$$\widehat{S}_{\mathrm{HO}}(k) = 1 + \rho\frac{(2\pi)^{d/2}\pi}{k^\nu}\sum_{j=1}^{N} w_{\nu,j}\tilde{h}_k\left(\frac{\pi}{h}\psi(h\xi_{\nu,j})\right) J_\nu\left(\frac{\pi}{h}\psi(h\xi_{\nu,j})\right)\psi'(h\xi_{\nu,j}), \tag{3.32}$$

with $\nu = d/2-1$, $N \in \mathbb{N}$, $\tilde{h}_k(x) = \frac{x^{d/2}}{k^{d/2+1}}(\hat{g}(x/k)-1)$, and $\hat{g}$ an estimator of the pair correlation function. Finally, note that Ogata's quadrature is also implemented in the Python package hankel of Murray and Poulin (2019).

Remark 3.2 (Relation between $k_{\min}$ and $r_{\max}$). There exists a hidden inverse proportionality relation in Equation (3.32), between the minimal wavenumber $k_{\min}$ for which we can hope the estimator to be accurate and the maximal radius $r_{\max}$ at which the pair correlation function has been estimated. Truncating the infinite sum after N terms in Equation (3.30) has been informally justified by

$$\psi(h\xi_{d/2-1,N}) \approx h\xi_{\frac{d}{2}-1,N}. \tag{3.33}$$

The maximum radius $r_{\max}$ at which $\hat{g}$ is available should in turn satisfy

$$r_{\max} = \max_j\left\{\frac{\pi}{hk}\psi(h\xi_{\frac{d}{2}-1,j}); k \in \mathbb{R}^*\right\}. \tag{3.34}$$

Together, (3.33) and (3.34) entail that

$$k_{\min} \approx \frac{\pi\xi_{d/2-1,N}}{r_{\max}}. \tag{3.35}$$

Thus $k_{\min}$ is not only proportional to $1/r_{\max}$ but also to the largest considered zero of the Bessel function $J_{d/2-1}(x)$.

Using the quadrature of Baddour and Chouinard

Instead of using the quadrature of Ogata (2005), one can estimate Hankel transforms more directly, similarly to how the discrete Fourier transform is used to approximate Fourier transforms. Intuitively, assuming that either f or its Hankel transform $\mathcal{H}_\nu(f)$ (A.3) has bounded support allows rewriting it as a Fourier-Bessel series, with coefficients involving evaluations of either $\mathcal{H}_\nu(f)$ or f, respectively. Truncating the resulting Fourier-Bessel series yields approximate direct and inverse Hankel transforms. This discrete Hankel transform (DHT) was derived by Baddour and Chouinard (2015) and originally implemented in Matlab. Moreover, Guizar-Sicairos and Gutiérrez-Vega (2004) developed a Python package, pyhank, based on the same idea.

In detail, let $N > 0$ and $f : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function, Baddour and Chouinard (2015) approximate

$$\mathcal{H}_\nu(f)(k_m) \approx \frac{r_{\max}^2}{\eta_{\nu,N}} \sum_{j=1}^{N-1} \frac{2J_\nu\left(\frac{\eta_{\nu,m}\eta_{\nu,j}}{\eta_{\nu,N}}\right) f(r_j)}{\eta_{\nu,N} J_{\nu+1}^2(\eta_{\nu,j})}, \tag{3.36}$$

where $\eta_{\nu,m} = \pi\xi_{\nu,m}$ is the m^{th} positive zero of the Bessel function $J_\nu(x)$ of the first kind, $1 \leq j, m \leq N-1$ and

$$r_j = \frac{\eta_{\nu,j}}{\eta_{\nu,N}} r_{\max}, \quad k_m = \frac{\eta_{\nu,m}}{\eta_{\nu,N}} k_{\max}, \quad k_{\max} = \frac{\eta_{\nu,N}}{r_{\max}}, \quad \text{with} \quad r_{\max} > 0. \tag{3.37}$$

The user thus needs to specify both N and $r_{\max}$. Intuitively, the choice of $r_{\max}$ is governed by how far on the positive axis one has been able to evaluate f. Once $r_{\max}$ is fixed, N decides how large $k_{\max}$ is, that is, how high in frequency one wishes to estimate the Hankel transform.

To conclude, given an estimator $\hat{g}$ of the pair correlation function, we define yet another estimator of the structure factor, called the Hankel-Baddour-Chouinard estimator,

$$\widehat{S}_{\mathrm{HBC}}(k_m) = 1 + \rho(2\pi)^{\frac{d}{2}} \frac{r_{\max}^2}{\eta_{\nu,N}} \sum_{j=1}^{N-1} \frac{2J_\nu\left(\frac{\eta_{\nu,m}\eta_{\nu,j}}{\eta_{\nu,N}}\right)}{\eta_{\nu,N} J_{d/2}^2(\eta_{\nu,j}) k_m^\nu} \tilde{h}(r_j), \tag{3.38}$$

where $\nu = d/2 - 1$, $\tilde{h}(x) = x^\nu(\hat{g}(x) - 1)$, and the set of wavenumbers $\{k_m\}_m$ is fixed by (3.37). Finally, we can deduce from (3.37) that the minimal wavenumber of $\widehat{S}_{\mathrm{HBC}}$ (3.38) is $k_1 = k_1^{d/2-1} = \frac{\eta_{d/2-1,1}}{r_{\max}}$. Comparing k_1 with the minimal

wavenumber $k_{\min}$ (3.35) of $\widehat{S}_{\mathrm{HO}}$ (3.32), for the same number of points N and the same $r_{\max}$, we observe that $k_1 < k_{\min}$, as k_1 is proportional to the first zero of the Bessel function $J_{d/2-1}(x)$ while $k_{\min}$ is proportional to the N^{th} zero of $J_{d/2-1}(x)$. For the study of hyperuniformity, this suggests an advantage to using $\widehat{S}_{\mathrm{HBC}}$.

3.2.3 Demonstrating the estimators

We now demonstrate the strcuture factor estimators presented in Sections 3.2.1 and 3.2.2. All figures in this section can be reproduced using structure_factor, by following our demonstration notebook [4]; see also Section 3.4 for an introduction to structure_factor.

For this demonstration, we used $A = 50$ samples from the point processes KLY (Example 2.4), GPP (Example 2.2), PPP (Section 2.4), and TPP (Example 2.1). Recall that the KLY and the GPP are Class I HUPPs while PPP and TPP are not HUPPs. The KLY is obtained from matching $\mathbb{Z}^2$ with a PPP of intensity $\rho = 11$. Moreover, the KLY is of intensity $\rho_{\mathrm{KLY}} = 1$, while the GPP is of intensity $\rho_{\mathrm{Ginibre}} = 1/\pi$. The used PPP is of intensity $\rho_{\mathrm{Poisson}} = 1/\pi$, and TPP is of intensity $\rho_{\mathrm{Thomas}} = 1/\pi$, with $\rho_{\mathrm{parent}} = 1/(20\pi)$ and $\sigma = 2$. Our choice of observation window depends on the intensity and makes sure that we get samples of around 5800 points. The dimension is always $d = 2$.

Estimators assuming stationarity

The scattering intensity Figure 3.1 illustrates the scattering intensity estimator of Section 3.2.1. Columns respectively correspond to the KLY, GPP, PPP, and TPP. The first row contains a sample of each point process, observed in square windows. The second row shows the scattering intensity $\widehat{S}_{\mathrm{SI}}$ in (3.12) on arbitrary wavevectors $\mathbf{k}$, while in the third row, the estimators are only evaluated on a subset of the allowed wavevectors (3.11). The fourth and fifth rows illustrate the debiasing techniques, respectively the directly debiased scattering intensity $\widehat{S}_{\mathrm{DDT}}(t_0, \mathbf{k})$ from (3.21), and the undirectly debiased scattering intensity $\widehat{S}_{\mathrm{UDT}}(t_0, \mathbf{k})$ from (3.20).

[4] https://github.com/For-a-few-DPPs-more/structure-factor/tree/main/notebooks

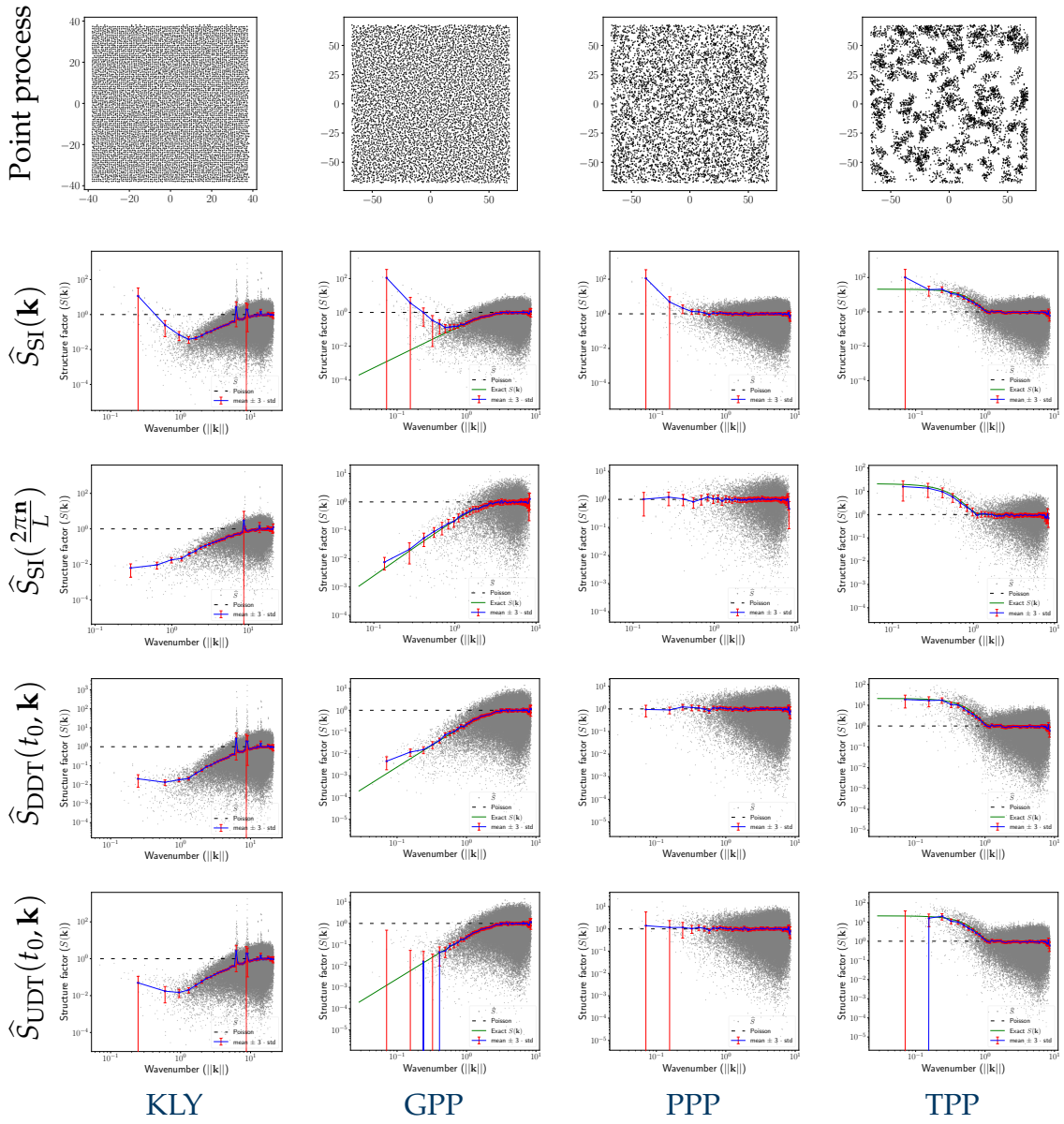


Figure 3.1: Variants of the scattering intensity estimator applied to four point processes. Plots obtained using the toolbox structure_factor.

The clouds of grey points are the approximated structure factors of the samples observed in the first row of the figure. For isotropic point processes, the structure factor S is a radial function, so we plot $k \mapsto \hat{S}(k)$, $k \in \mathbb{R}$ and not $\mathbf{k} \mapsto \hat{S}(\mathbf{k})$, $\mathbf{k} \in \mathbb{R}^2$. The KLY is the only non-isotropic example: in that case, we numerically average $\hat{S}(\mathbf{k})$ over vectors satisfying $\|\mathbf{k}\|_2 = k$. To regularize the obtained estimator, we bin the norm of the wavevectors regularly and provide the empirical

mean (in blue) and the empirical standard deviation of the mean (red bars indicate ± 3 such standard deviations). Note that the binning can be specified by the user in structure_factor. On each plot, the exact structure factor is represented by a green line when it is known. Finally, the dashed black lines are the structure factor of the PPP, for reference.

While we refer to Section 3.2.4 for a more detailed comparison, one can already observe from Figure 3.1 that the most accurate estimators are the scattering intensity $\widehat{S}_{\mathrm{SI}}$ (3.12) evaluated on the set of allowed wavevectors (3.11) and the debiased scattering intensity $\widehat{S}_{\mathrm{DDT}}(t_0, \mathbf{k})$ (3.21). The bias at small, non-allowed wavenumbers of the scattering intensity is visible in the second row. As for the undirectly debiased variant, it produces a few negative values, visible as large error bars on our log-log plot.

Using an alternate taper As mentioned in Section 3.2, the scattering intensity $\widehat{S}_{\mathrm{SI}}$ is a particular case of the tapered estimator $\widehat{S}_{\mathrm{T}}$, with the specific taper t_0. We are free to use other tapers verifying (3.18).

Figure 3.2 shows the estimated structure factors of the same four benchmark point processes (first row), using $\widehat{S}_{\mathrm{T}}$ (second row), the corresponding directly debiased version $\widehat{S}_{\mathrm{DDT}}$ (third row), and the undirectly debiased version $\widehat{S}_{\mathrm{UDT}}$ (last row). The taper used is the first sinusoidal taper $t_1(\mathbf{x}, W) = t(\mathbf{x}, \mathbf{p}^1, W)$ with $\mathbf{p}^1 = (1, 1)$ in (3.23). The same legend applies as for Figure 3.1.

First, the asymptotic bias of $\widehat{S}_{\mathrm{T}}$ at small wavenumber k is visible in the second row. Second, for the KLY (first column), the GPP (second column), and the PPP (third column) the estimator $\widehat{S}_{\mathrm{UDT}}$ (last row) returned a few negative values again, resulting in large inaccuracies in our log-log scale. The directly debiased estimator $\widehat{S}_{\mathrm{DDT}}$ yields the most accurate approximation of known structure factors, consistently across point processes.

Averaging over multiple tapers The multitapered estimator $\widehat{S}_{\mathrm{MT}}$ of (3.22) is now investigated in Figure 3.3, using the first four sinusoidal tapers, i.e., $(t_q)_{q=1}^4$ with $t_q(\mathbf{x}, W) = t(\mathbf{x}, \mathbf{p}^q, W)$ and $\mathbf{p}^q \in \{1, 2\}^2$ in (3.23). We also show the results of the corresponding directly and undirectly debiased versions, $\widehat{S}_{\mathrm{DDMT}}$ and $\widehat{S}_{\mathrm{UDMT}}$.

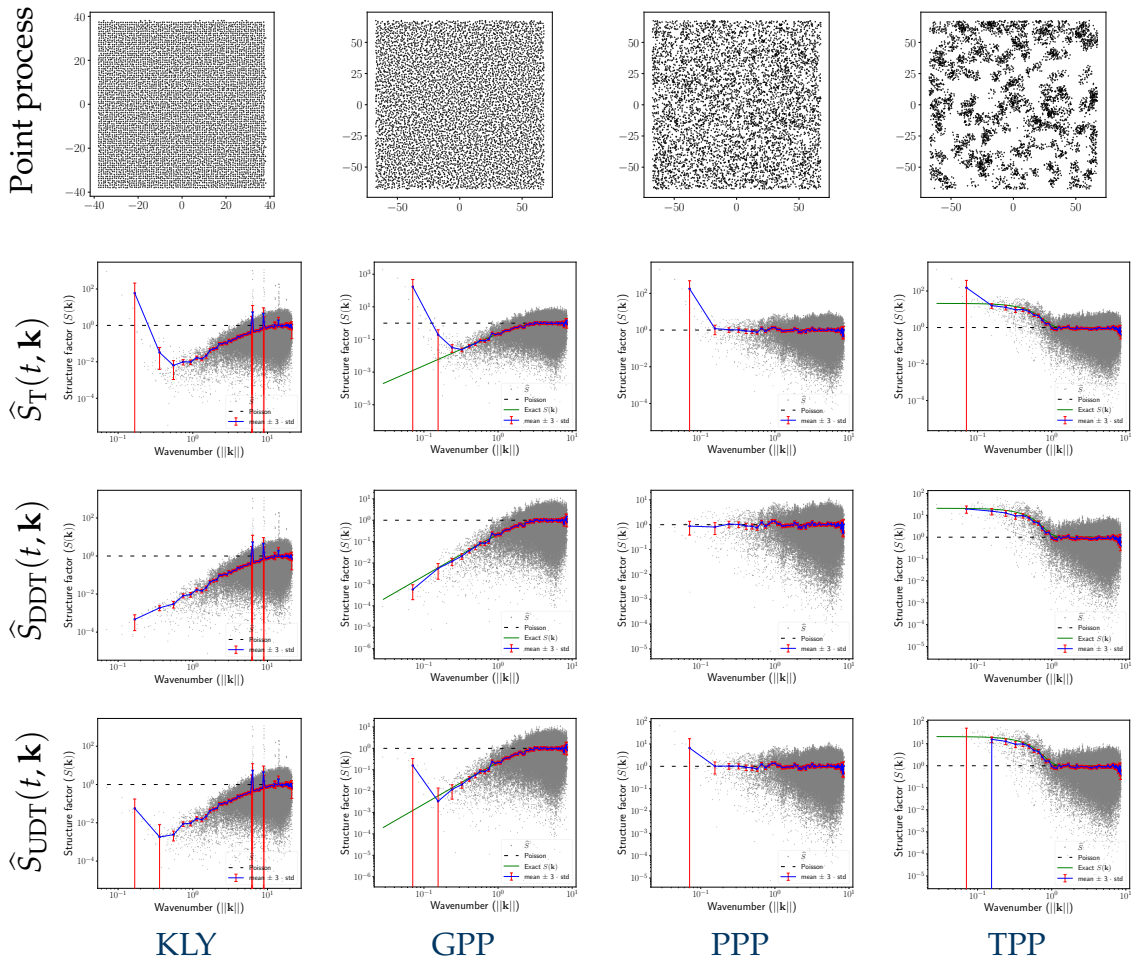


Figure 3.2: Tapered estimator and the corresponding debiased versions applied to four point processes. Plots obtained using the toolbox structure_factor.

We again observe the bias of $\widehat{S}_{\mathrm{MT}}$ at small wavenumbers (second row), and that the negative values output by $\widehat{S}_{\mathrm{UDMT}}$ at small wavenumbers (last row) make visual assessments of hyperuniformity less straightforward. Like with single tapers, the directly debiased estimator $\widehat{S}_{\mathrm{DDMT}}$ gives a consistently accurate approximation. Compared to Figure 3.2, however, it is not obvious whether multitapering yields a smaller mean square error than single tapers, and a more quantitative study will investigate this in Section 3.2.4.

Estimators assuming stationarity and isotropy

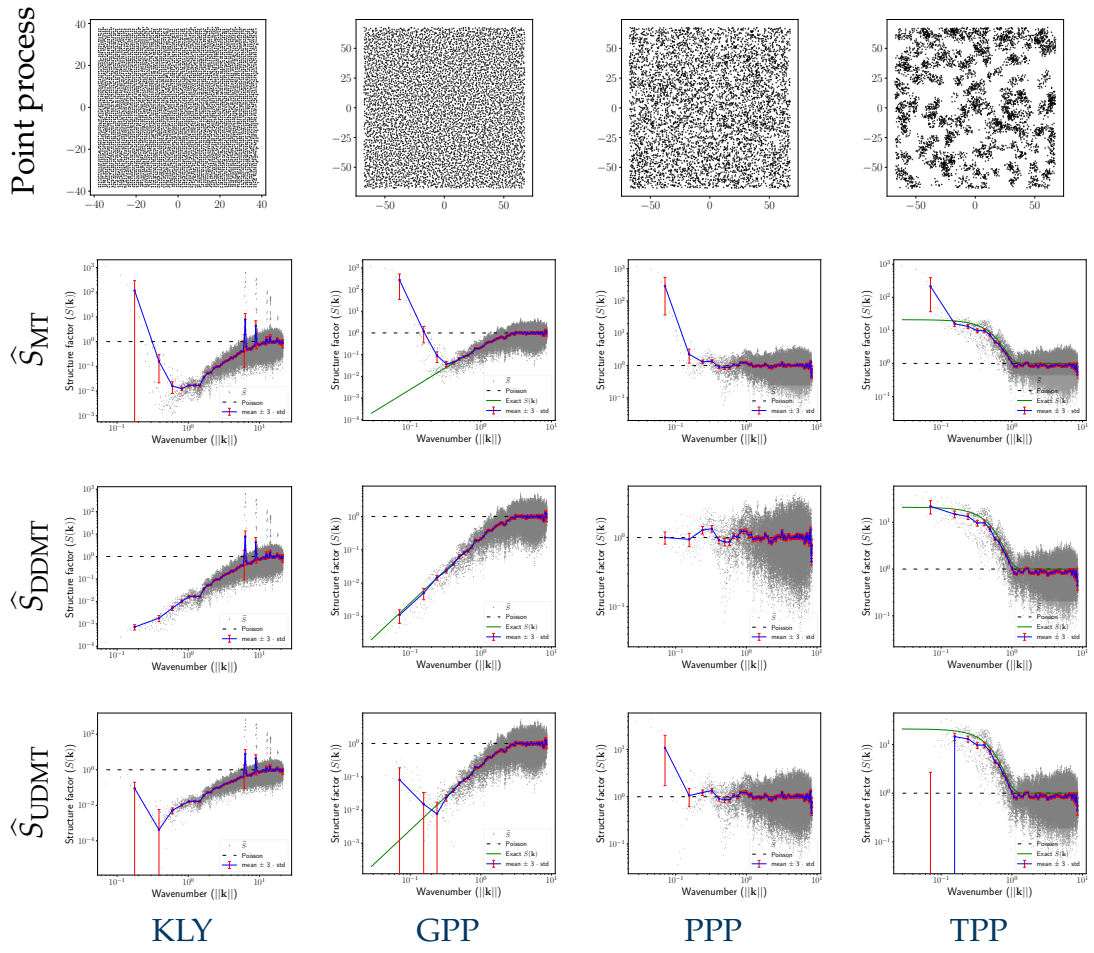


Figure 3.3: Multitapered estimator and the debiased versions applied to four point processes. Plots obtained using the toolbox structure_factor.

Bartlett's isotropic estimator Figure 3.4 illustrates Bartlett's isotropic estimator of Section 3.2.2. Columns respectively correspond to the KLY, GPP, PPP, and TPP. The first row contains a sample of each point process observed in a ball window. The second row shows $\widehat{S}_{\mathrm{BI}}$ on arbitrary wavenumbers k, while in the last row, the estimator is only evaluated on a subset of the Bessel-specific allowed wavenumbers (3.29).

First, we note that, unlike scattering intensity variants, plotting Bartlett's isotropic estimator $k \mapsto \widehat{S}_{\mathrm{BI}}(k)$ in (3.27) does not require binning. On the other hand, Bartlett's estimator is significantly costlier than its scattering intensity counterpart; see Section 3.2.4. Now, we comment on the accuracy of the estimator in Figure 3.4. Here again, small, non-allowed wavenumbers give rise to large biases

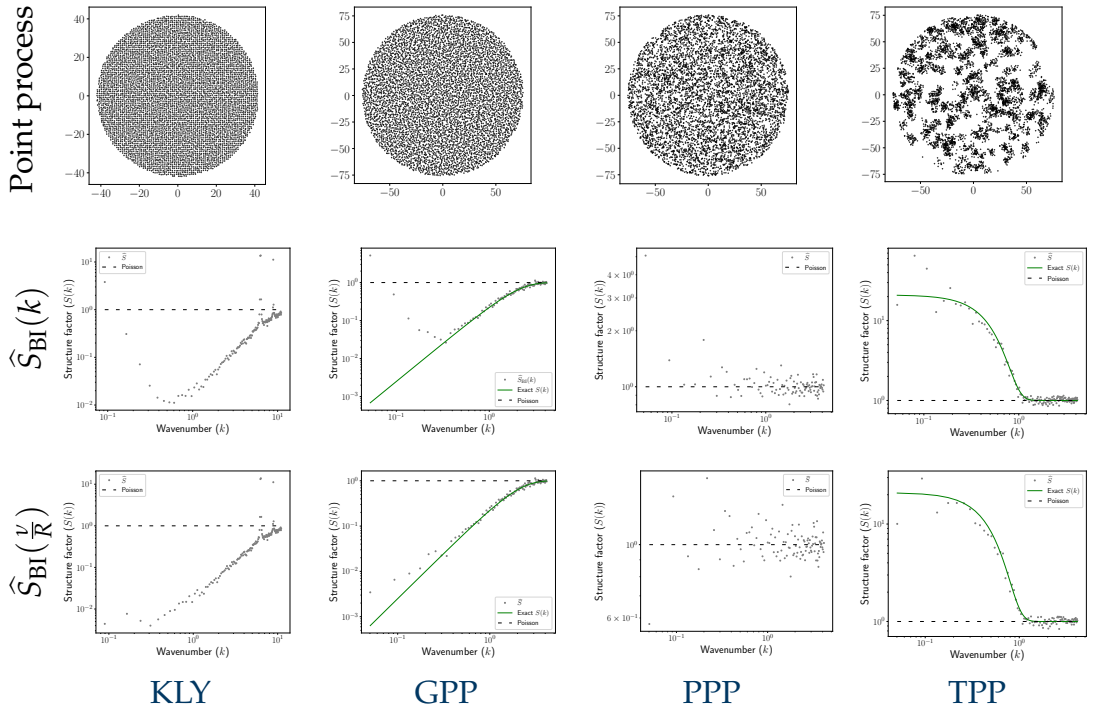


Figure 3.4: Bartlett's isotropic estimator applied to four point processes. Plots obtained using the toolbox structure_factor.

for $\widehat{S}_{\mathrm{BI}}(k)$, especially for the two HUPPs KLY and GPP. When applied to allowed wavenumbers, the estimator shows accuracy across all point processes, similarly to the directly debiased tapered estimators.

Estimating the pair correlation function The last two estimators of the structure factor are $\widehat{S}_{\mathrm{HO}}$ (3.32), and $\widehat{S}_{\mathrm{HBC}}$ (3.38). These estimators require an approximation of the pair correlation function g (2.5) of the point process. We thus quickly investigate standard estimators of the pair correlation function on our benchmark point processes.

There are two types of estimators of g for stationary isotropic point processes (Baddeley et al., 2015): kernel density estimators applied to pairwise distances and numerical derivatives of Ripley's K function. The R library spatstat implements both, respectively as `pcf.ppp`, which uses an Epanechnikov kernel and Stoyan's rule of thumb for bandwidth selection (Baddeley et al., 2015, Section 7.6.2), and `pcf.fv`, which computes the derivative of a polynomial estimator of

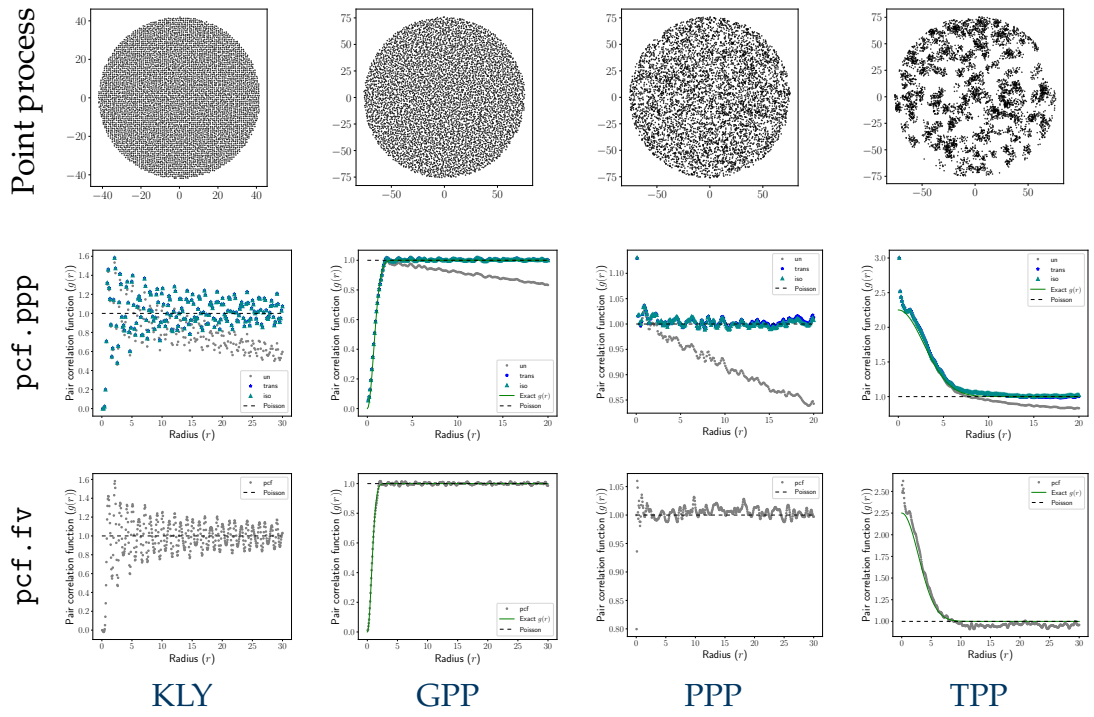


Figure 3.5: Approximated pair correlation function applied to four point processes. Plots obtained using the toolbox structure_factor.

Ripley's K function. The kernel density estimator behaves badly for small values of r: for many point processes, its variance becomes infinite when r goes to 0. The derivative estimator, `pcf.fv`, is recommended for large datasets, where direct estimation of the pair correlation function can be time-consuming (Baddeley et al., 2015, Section 7.6.2). Figure 3.5 shows the two estimators of the pair correlation function of the benchmark point processes. Note that we provide an independent, open-source Python interface[5] to the R library `spatstat`. The second row shows the estimation of the pair correlation function using `pcf.ppp`. This method provides a choice of boundary corrections, like `"trans"`, `"iso"`, or none (`"un"`). For more details see Baddeley et al. (2015, Sections 7.4.4 and 7.4.5). The last row of Figure 3.5 shows the estimation of the pair correlation function using `pcf.fv`.

We observe that, for the cardinalities considered here, the choice of edge correction method is irrelevant, as long as there is one. As expected, the uncorrected version `"un"` underestimates the pcf as r increases. This results from counting

[5] At `https://github.com/For-a-few-DPPs-more/spatstat-interface` and on PyPI.

only the pairs of points that fall inside the observation window, without correcting border effects. We also observe that the two methods for estimating the pair correlation function perform similarly, and we pick `pcf.fv` for the rest of this section. We manually remove undefined values (`NaN`, `-Inf`, or `Inf`), and we interpolate the obtained discrete approximation of g, in order to evaluate it at any point required by the quadratures of Section 3.2.2. Finally, note that the maximum radius $r_{\max}$ at which `spatstat` provides an approximation of g is limited by the size of the observation window. Typically, it should be less than half the window diameter for a ball window, and less than $1/4$ of the smaller side length of the window for a rectangular window; see the documentation of `spatstat`. For larger values than the $r_{\max}$ provided by `spatstat`, we manually set g to be identically 1, which has the effect of automatically truncating quadratures that evaluate $g - 1$, like Ogata's quadrature (3.32).

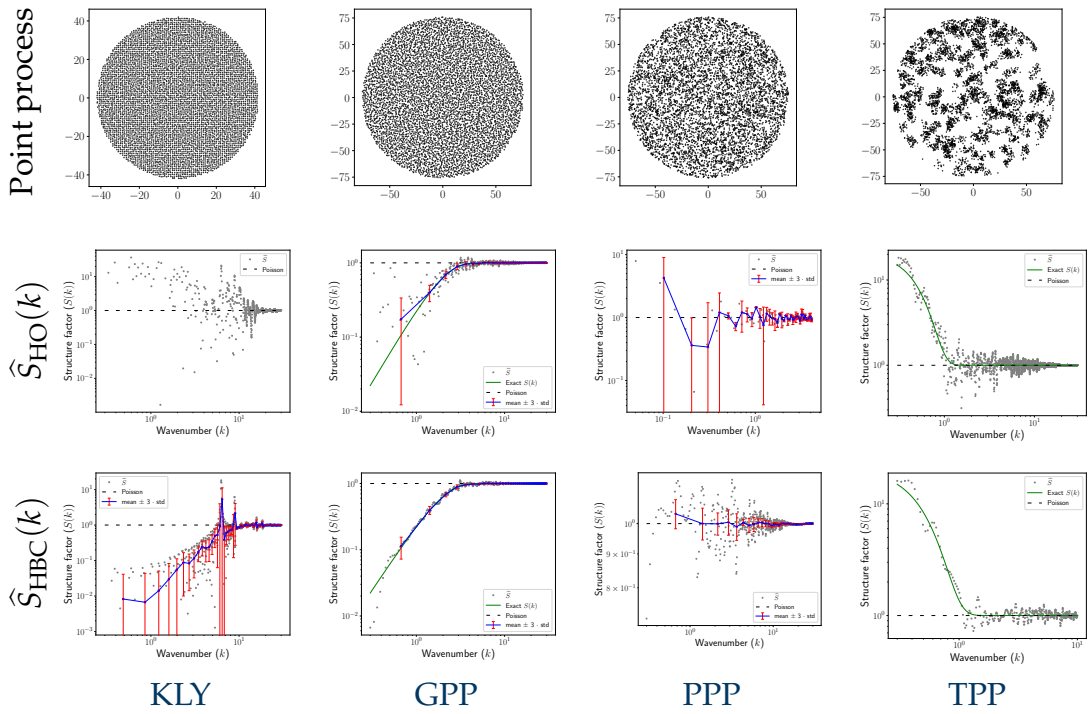


Figure 3.6: Structure factor estimation using Hankel transform quadratures applied to four point processes. Plots obtained using the toolbox structure_factor.

Hankel transform quadratures Figure 3.6 shows the results of Ogata's $\widehat{S}_{\mathrm{HO}}$ (second row) and Baddour-Chouinard's $\widehat{S}_{\mathrm{HBC}}$ (last row) estimators on the benchmark point processes KLY, GPP, PPP and TPP. The legend is the same as for Figure 3.1. For the accuracy of the estimators, we can see that $\widehat{S}_{\mathrm{HO}}$ failed to approximate the structure factor of the KLY. Even the results of $\widehat{S}_{\mathrm{HBC}}$ seem to be unreliable. The non-isotropy of the KLY may be the reason for the fluctuations of its approximated pair correlation function in Figure 3.5, leading to the inaccuracies of the quadratures. For the remaining point processes, $\widehat{S}_{\mathrm{HBC}}$ seems to give more accurate results than $\widehat{S}_{\mathrm{HO}}$.

3.2.4 Quantitative comparison of the estimators

We now compare the cost and accuracy of all estimators more quantitatively.

A note on computational costs

For a given wavevector, evaluating any single-tapered estimator requires a sum of N terms. Multitapering naturally multiplies the cost by the number of tapers, but it can be trivially parallelized, especially since the number of tapers remains low in practice (Rajala et al., 2020).

On the other hand, for a sample of N points, Bartlett's estimator $\widehat{S}_{\mathrm{BI}}$ is a sum of $O((N^2 - N)/2)$ evaluations of a Bessel function. This makes Bartlett's estimator significantly costlier than its scattering intensity counterpart. As an example, for a realization of $N = 10^4$ points from a point process in $\mathbb{R}^2$, using a modern laptop, the evaluation of $\widehat{S}_{\mathrm{SI}}$ at a single wavevector took about one millisecond, compared to 2 seconds for $\widehat{S}_{\mathrm{BI}}$. Note however that for a similar accuracy on an isotropic point process, $\widehat{S}_{\mathrm{BI}}$ needs to be evaluated at fewer wavevectors than, say, the scattering intensity. Indeed, a single value of the former should be compared to a binwise average of the latter.

Now, for Hankel quadrature estimators, the main bottleneck is the approximation of the pair correlation function. A kernel density estimator based on N points is again a quadratic computation. In our experience, combining the estimator

`pcf.fv`, based on numerically differentiating an estimated Ripley's K function, and a Hankel quadrature to build $\widehat{S}_{\mathrm{HO}}$ and $\widehat{S}_{\mathrm{HBC}}$ is the least expensive pipeline.

Measuring the accuracy of the estimators near zero

To confirm the intuitions gained from Sections 3.2.3, in particular, that the multitapered variant of the scattering intensity and Bartlett's estimator dominate the others when it comes to visual diagnostics of hyperuniformity, we now compare the mean square error of the different estimators of the structure factor, integrated near 0.

For a subdivision $[k_1, k_2) \cup \cdots \cup [k_{J-1}, k_J)$ of some interval $[k_1, k_J)$, the integrated MSE reads

$$\begin{aligned}\mathrm{iMSE}(\widehat{S}) &= \int_{k_1}^{k_J} \mathrm{MSE}(\widehat{S})(k)\mathrm{d}k \\ &= \int_{k_1}^{k_J} \mathbb{E}\left[\left(\widehat{S}(k) - S(k)\right)^2\right] \mathrm{d}k \\ &= \mathbb{E}\left[\int_{k_1}^{k_J} \left(\widehat{S}(k) - S(k)\right)^2 \mathrm{d}k\right].\end{aligned}$$

A crude numerical integration using the trapezoidal rule gives,

$$\mathrm{iMSE}(\widehat{S}) \approx \mathbb{E}\left[\sum_{j=1}^{J-1} \frac{1}{2}(k_{j+1} - k_j)\left[\left(\widehat{S}(k_{j+1}) - S(k_{j+1})\right)^2 + \left(\widehat{S}(k_j) - S(k_j)\right)^2\right]\right], \tag{3.39}$$

The expectation in (3.39) is under the investigated point process. Assuming it is easy to sample from the point process, we can estimate the iMSE in (3.39) by an empirical average $\widehat{\mathrm{iMSE}}$ over point process samples. Similarly, the difference of the iMSEs of two different estimators of S can be approximated by an empirical average of differences of $\widehat{\mathrm{iMSE}}$s. This yields a natural (paired) Student test to assess whether the difference of the iMSEs of two different estimators of S is 0.

Comparing variants of the scattering intensity

In Section 3.2.3, we derived the intuition that $\widehat{S}_{\mathrm{SI}}$ on its allowed values, $\widehat{S}_{\mathrm{DDT}}(t_0, \mathbf{k})$, $\widehat{S}_{\mathrm{DDT}}(t_1, \mathbf{k})$, and $\widehat{S}_{\mathrm{DDMT}}((t_q)_{q=1}^4, \mathbf{k})$ gave the most accurate approximations among the estimators of Section 3.2. We further expect $\widehat{S}_{\mathrm{DDMT}}((t_q)_{q=1}^4, \mathbf{k})$ to have the

smallest integrated MSE among them. To test this, we run three paired, one-sided Student tests, comparing $\widehat{S}_{\mathrm{DDMT}}((t_q)_{q=1}^4, \mathbf{k})$ to each of the three former estimators.

Specifically, we use 50 independent realizations of approximately 5800 points each, from the GPP, PPP, and TPP, for which we know the exact structure factor. The intensity of all processes is $\rho = 1/\pi$. The additional parameters of the TPP are, as fixed before, $\rho_{\mathrm{parent}} = 1/(20\pi)$, and $\sigma = 2$. We evaluated $\widehat{S}_{\mathrm{SI}}$ (3.12) on its allowed wavevectors (3.11), between $k_1 = 0.1$ and $k_J = 2.8$. For the other estimators, $\widehat{S}_{\mathrm{DDT}}(t_0, \mathbf{k})$, $\widehat{S}_{\mathrm{DDT}}(t_1, \mathbf{k})$ (3.21), and $\widehat{S}_{\mathrm{DDMT}}((t_q)_{q=1}^4, \mathbf{k})$ (3.22), we used arbitrary wavevectors keeping the same range of the wavenumbers used for $\widehat{S}_{\mathrm{SI}}$. The tapers used are t_0 (3.16), and the first 4 tapers $(t_q)_{q=1}^4$ of the family of sinusoidal tapers (3.23). As we got different approximated values of S for different wavevectors of the same wavenumber, for each wavenumber we set the estimated structure factor to the sample mean of these values.

Table 3.2: Paired t-tests

Estimators	T-score	p-value	T-score	p-value	T-score	p-value
$\widehat{S}_{\mathrm{DDMT}}$, $\widehat{S}_{\mathrm{SI}}$	-29.53	3×10^{-33}	-41.59	3×10^{-40}	-9.24	10^{-12}
$\widehat{S}_{\mathrm{DDMT}}$, $\widehat{S}_{\mathrm{DDT}}$	-22.40	10^{-27}	-30.42	8×10^{-34}	-6.38	2×10^{-8}
$\widehat{S}_{\mathrm{DDMT}}$, $\widehat{S}_{\mathrm{DDT}}$	-12.18	9×10^{-17}	-25.39	3×10^{-30}	-7.16	10^{-9}
	GPP		PPP		TPP	

Table 3.2 summarizes the results of the paired one-sided Student tests. For each point process, applying a Bonferroni correction [6], we can simultaneously reject at the level 0.01 the three hypotheses that there is no difference in mean between the estimated iMSEs. This confirms our intuition that $\widehat{S}_{\mathrm{DDMT}}$ yields the smallest iMSE among the considered variants of the scattering intensity. In particular, multitapering helps.

For information, the estimated iMSEs are given in Table 3.3, in the form of a confidence interval (CI) of the $\widehat{\mathrm{iMSE}}$s plus or minus 3 empirical standard deviations

[6] When conducting statistical hypothesis testing with multiple hypotheses, the probability of encountering an unusual event becomes higher increasing the risk of erroneously rejecting a null hypothesis. To mitigate this issue, Bonferroni (1936) proposed testing each hypothesis individually at a significance level of α divided by the number of hypotheses being tested, instead of the usual α.

of the mean. There is roughly a factor 4 between the $\widehat{\text{iMSE}}$ of the directly debiased multitapered estimator and each of the other three, which confirms its domination. For reference, we have also indicated the contribution of the sample integrated variance $\widehat{\text{i}\mathbb{V}\text{ar}}$ to each $\widehat{\text{iMSE}}$. Clearly, the variance is the biggest contributor, and the squared bias is at least one order of magnitude smaller, for all point processes.

Finally, Figure 3.7 shows the pointwise average of each estimator, across the 50 point process realizations. A reduction of bias at small k for $\widehat{S}_{\text{DDT}}(t_1, \mathbf{k})$ and $\widehat{S}_{\text{DDMT}}((t_q)_{q=1}^4, \mathbf{k})$ is visible for the GPP.

Table 3.3: Sample integrated variance and MSE

Estimators	$\widehat{\text{i}\mathbb{V}\text{ar}}$	CI[$\widehat{\text{iMSE}}$]	$\widehat{\text{i}\mathbb{V}\text{ar}}$	CI[$\widehat{\text{iMSE}}$]	$\widehat{\text{i}\mathbb{V}\text{ar}}$	CI[$\widehat{\text{iMSE}}$]
$\widehat{S}_{\text{SI}}(2\pi\mathbf{n}/L)$	0.32	0.32 ± 0.02	1.31	1.34 ± 0.06	69.51	70.71 ± 17.95
$\widehat{S}_{\text{DDT}}(t_0)$	0.32	0.33 ± 0.03	1.44	1.47 ± 0.1	72.15	73.63 ± 26.12
$\widehat{S}_{\text{DDT}}(t_1)$	0.34	0.35 ± 0.06	1.47	1.50 ± 0.14	79.29	80.51 ± 27.20
$\widehat{S}_{\text{DDMT}}((t_q)_1^4)$	0.08	$\mathbf{0.08} \pm 0.007$	0.37	$\mathbf{0.38} \pm 0.02$	17.90	$\mathbf{18.19} \pm 4.19$
	GPP		PPP		TPP	

Comparing estimators that assume isotropy

We now run the same comparison on $\widehat{S}_{\text{BI}}$ (3.27), $\widehat{S}_{\text{HO}}$ (3.32), and $\widehat{S}_{\text{HBC}}$ (3.38). For this study, we have sampled 50 independent realizations of approximately 5800 points each, from the GPP, PPP, and TPP, with the same parameters as in Section 3.2.4, except that the observation window for all point processes is now a 2-dimensional ball window centered at the origin. For $\widehat{S}_{\text{HO}}$, and $\widehat{S}_{\text{HBC}}$ we used the method `pcf.fv` to approximate the pair correlation function with maximal approximation radius $r_{\text{max}} = 30$.

We have noted in Section 3.2.2 that $\widehat{S}_{\text{HBC}}$ is more robust to non-isotropy than $\widehat{S}_{\text{HO}}$ and that $\widehat{S}_{\text{BI}}$ gave the tightest approximations in this family of estimators.

Table 3.4, summarizes the results of two one-sided paired t-tests per point process. For each point process, we can again simultaneously reject at level 0.01 the hypothesis that there is no difference between the estimates iMSEs. This confirms

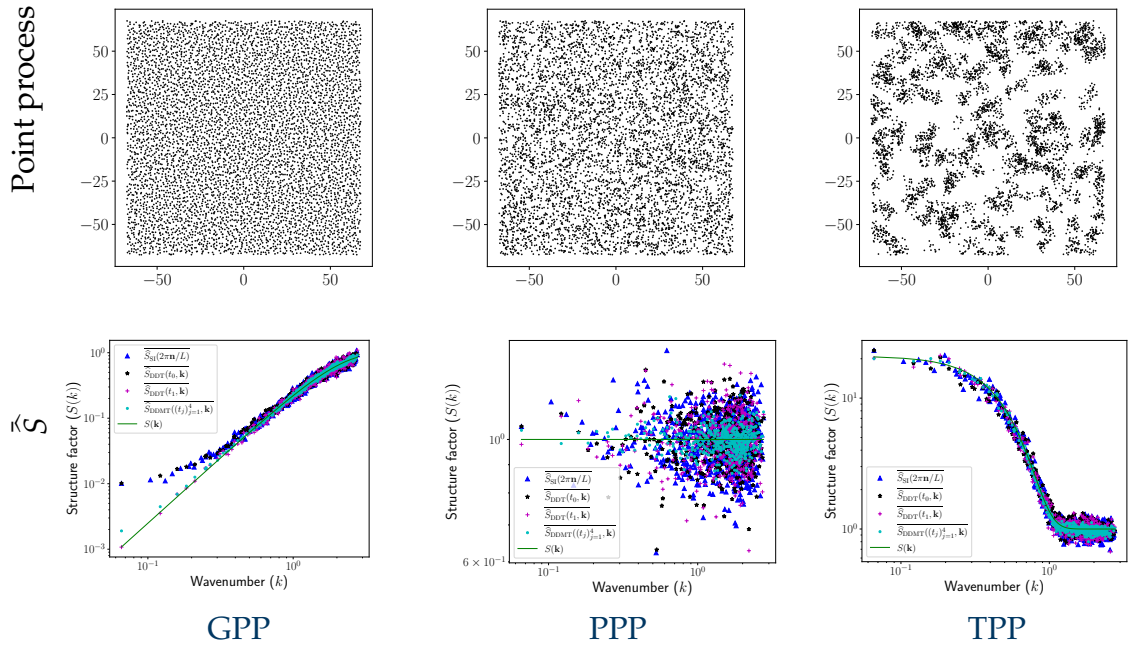


Figure 3.7: Pointwise averages of the variants of the scattering intensity, numerically rotation-averaged, across 50 independent realizations of the GPP, PPP, and TPP. Plots obtained using the toolbox structure_factor.

Table 3.4: Paired t-tests

Estimators	T-score	p-value	T-score	p-value	T-score	p-value
$\widehat{S}_{\mathrm{BI}}$ vs. $\widehat{S}_{\mathrm{HO}}$	-12.24	7×10^{-17}	-6.60	10^{-8}	-5.32	10^{-6}
$\widehat{S}_{\mathrm{BI}}$ vs. $\widehat{S}_{\mathrm{HBC}}$	-25.51	2×10^{-30}	-5.32	10^{-6}	-5.16	2×10^{-6}
	GPP		PPP		TPP	

the claim that Bartlett's estimator is the most accurate near 0, among estimators that assume isotropy. For reference, Table 3.5 gives the same summary statistics as Table 3.3 did for scattering intensity variants. Bartlett's estimator yields $\widehat{\mathrm{iMSE}}$s one order of magnitude smaller than both Hankel transform quadratures, for both GPP and PPP, and a factor three for TPP. Again, the iMSE is mostly variance.

Finally, Figure 3.8 shows the pointwise average of each estimator, across the 50 point processes realizations. The accuracy of Ogata's quadrature at small wavenumbers is poor. The bias of the two Hankel quadrature estimators is well

Table 3.5: Sample integrated variance and MSE

$\widehat{S}$	$\widehat{\text{iVar}}$	CI[$\widehat{\text{iMSE}}$]	$\widehat{\text{iVar}}$	CI[$\widehat{\text{iMSE}}$]	$\widehat{\text{iVar}}$	CI[$\widehat{\text{iMSE}}$]
$\widehat{S}_{\text{BI}}$	3.9×10^{-3}	$\mathbf{4.0 \times 10^{-3}} \pm 3 \times 10^{-4}$	0.057	$\mathbf{0.058} \pm 9 \times 10^{-3}$	11.25	$\mathbf{11.65} \pm 4.71$
$\widehat{S}_{\text{HO}}$	0.37	0.38 ± 0.09	2.12	2.14 ± 0.93	43.63	46.70 ± 18.40
$\widehat{S}_{\text{HBC}}$	0.03	0.03 ± 0.01	2.44	2.45 ± 1.33	57.62	63.02 ± 28.62
	GPP		PPP		TPP	

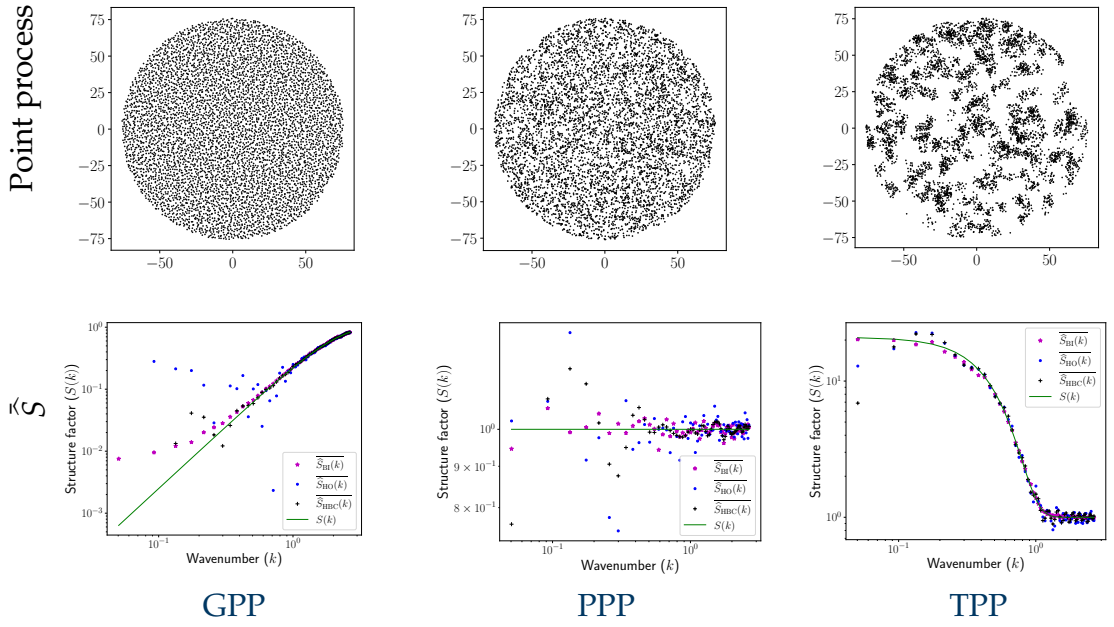


Figure 3.8: Pointwise averages of the three estimators that assume isotropy, across 50 independent samples of the GPP, PPP and TPP. Plots obtained using the toolbox structure_factor.

visible at small k for all point processes. This is likely due to poor estimation of the pair correlation function at large scales.

3.3 Hyperuniformity tests

The standard empirical diagnostic of hyperuniformity involves estimating and plotting the structure factor of the underlying point process (Torquato, 2018; Klatt et al., 2020). This graphical assessment is however not standardized and often

not described in full reproducible detail in the literature, with implementation choices and statistical properties often part of each field's folklore.

While presenting a preliminary draft of our paper (Hawat et al., 2023b) during the workshop "New trends in point process theory" at Karlsruhe in March 2022, a lively discussion with the participants about the existent hyperuniformity diagnostics revealed the need for a standardized statistical test of hyperuniformity. This further motivates us to contribute a statistical test of hyperuniformity [7].

In Section 3.3.1 we present the empirical diagnostics of hyperuniformity. Then, in Section 3.3.2 we present our statistical test of hyperuniformity (Hawat et al., 2023b, Section 4). To enhance standardization and accessibility, we have incorporated these tools into the toolbox structure_factor.

3.3.1 Empirical diagnostics

As mentioned in Section 3.1.1, the behavior in zero of the structure factor S quantifies hyperuniformity. Investigating whether one or several samples come from a HUPP is thus often carried out by estimating the structure factor, and then either visually inspecting the resulting plots around zero, or regressing the estimated values. Say for a motion-invariant point process, one option is to regress $\log(S)$ onto $\log(k)$ around zero, to assert a potential value for the decay rate α in Table 3.1.

Another criterion of *effective* hyperuniformity has been proposed (Torquato, 2018, Section 11.1.6); see also (Klatt et al., 2019, Supplementary material). For a motion-invariant HUPP, given a set of estimated values

$$\{(k_1, \widehat{S}(k_1)), \ldots, (k_n, \widehat{S}(k_n))\}$$

[7] We knew from personal communication during the workshop, that by then, Günter Last, Michael Andreas Klatt, and Norbert Henze were independently working on a hyperuniformity test, whose preprint (Klatt et al., 2022) came out as we were answering the referees on our manuscript (Hawat et al., 2023b). While our test was based on multiple realizations of the point process, interestingly, the test of Klatt et al. (2022), specifically tailored for Class I HUPPs, uses a single realization.

with $0 < k_1 < \cdots < k_n$, the H-index is defined by

$$H \triangleq \frac{\widehat{S}(0)}{\widehat{S}(k_{\text{peak}})}, \tag{3.40}$$

where $\widehat{S}(0)$ is a linear extrapolation of the structure factor in $k = 0$ based on the estimated values of S, and k_{peak} is the location of the first dominant peak value of the estimated structure factor, defined here as

$$k_{\text{peak}} = \inf_i \left\{k_i; \widehat{S}(k_i) > 1, \widehat{S}(k_{i-1}) < \widehat{S}(k_i), \text{ and } \widehat{S}(k_{i+1}) < \widehat{S}(k_i)\right\}.$$

If the set is empty, we set $\widehat{S}(k_{\text{peak}}) = 1$ in (3.40). When $H < 10^{-3}$, the process is called *effectively hyperuniform* by Torquato (2018, Section 11.1.6). Note that the linear extrapolation is chosen for simplicity and not based on model selection (Klatt et al., 2019, Supplementary materials). Like the threshold of 10^{-3}, the definition of a dominant peak is also somewhat arbitrary. In Section 3.3.2, we will introduce a statistical test that bypasses some of these arbitrary choices. Note that, Klatt et al. (2022) also introduced a statistical test of hyperuniformity specifically tailored for Class I HUPPs that bypasses some of these arbitrary choices.

3.3.2 Statistical test

In this section, we use the debiasing device of Rhee and Glynn (2015) to turn realizations of any of the nonnegative, asymptotically unbiased estimators of the structure factor S of Section 3.2, into an unbiased estimator of a truncated equivalent to $S(\mathbf{0})$. We then propose an asymptotically valid test of hyperuniformity. We group our assumptions and notations for this section in Assumption A3.

Assumption A3. *$\mathcal{X}$ is a stationary point process of $\mathbb{R}^d$ with intensity $\rho > 0$. Its pair correlation function g exists, and $\mathbf{r} \mapsto g(\mathbf{r}) - 1$ is integrable on $\mathbb{R}^d$.*

The coupled sum estimator

Consider a stationary point process $\mathcal{X}$ of $\mathbb{R}^d$, of which we observe the intersection of a single realization with multiple increasing windows. Formally, we consider an increasing sequence of sets $(\mathcal{X} \cap W_m)_{m \geq 1}$, with $W_s \subset W_r$ for all $0 < s < r$,

and $W_\infty = \mathbb{R}^d$. For simplicity, we assume that the windows are centered and either rectangular as in Assumption A1 or spherical as in Assumption A2 if $\mathcal{X}$ is isotropic.

We consider any of the positive, asymptotically unbiased estimators $\widehat{S}_m$ of S listed in Section 3.2, applied to $\mathcal{X} \cap W_m$. All such estimators are asymptotically unbiased on a set of values $\mathbb{B}_m$, possibly in the sense of Proposition 3.1, and there exists $\mathbf{k}_m^{\min} \in \mathbb{B}_m$ with $\mathbf{k}_m^{\min} \to \mathbf{0}$. The reader can keep in mind the scattering intensity of Section 3.2.1, with, say, the minimum restricted allowed value $\mathbf{k}_m^{\min} = (2\pi/L_{m_1}, \dots, 2\pi/L_{m_d}) \in \mathbb{Z}^d$, but the proofs hold more generally.

We define the sequence of random variables

$$Y_m = 1 \wedge \widehat{S}_m(\mathbf{k}_m^{\min}), \quad m \geq 1. \tag{3.41}$$

We cap the estimators at 1 arbitrarily to make them uniformly bounded. The idea is to use the sequence (Y_m) to test whether $S(\mathbf{0}) = 0$.

Following (Rhee and Glynn, 2015, Section 2), we first define a new sequence

$$Z_m = \sum_{j=1}^{m \wedge M} \frac{Y_j - Y_{j-1}}{\mathbb{P}(M \geq j)}, \quad m \geq 1, \tag{3.42}$$

where M is an $\mathbb{N}$-valued random variable such that $\mathbb{P}(M \geq j) > 0$ for all j, and $Y_0 = 0$ by convention. Rhee and Glynn (2015) observed that

$$\mathbb{E}[Z_m] = \mathbb{E}[Y_m] \quad \text{and} \quad Z_m \xrightarrow[m\to\infty]{\text{a.s.}} Z,$$

where Z is the *coupled sum estimator*

$$Z = \sum_{j=1}^{M} \frac{Y_j - Y_{j-1}}{\mathbb{P}(M \geq j)}. \tag{3.43}$$

Under some conditions, Rhee and Glynn (2015) proved that if $Y_m \to Y$ in L^2, then $Z \in L^2$ is an unbiased estimator of $\mathbb{E}[Y]$ with a square root convergence rate. For our choice of $(Y_m)_m$, the assumptions of (Rhee and Glynn, 2015, Theorem 1), especially the L^2 convergence of Y_m, may not hold for non-HUPPs. We thus use weaker assumptions that are still enough to build a hyperuniformity test.

Proposition 3.3. *Under Assumption A3, with Z defined in (3.43), assume that $M \in L^p$ for some $p \geq 1$. Then $Z \in L^p$ and $Z_m \to Z$ in L^p. Moreover,*

1. *If $\mathcal{X}$ is hyperuniform, then* $\mathbb{E}[Z] = 0$.

2. *If $\mathcal{X}$ is not hyperuniform and*

$$\sup_m \mathbb{E}\left[\widehat{S}_m^2(\mathbf{k}_m^{\min})\right] < \infty, \tag{3.44}$$

then $\mathbb{E}[Z] \neq 0$.

Proof. Let $M \in L^p$ with $p > 0$. We first prove that $Z_m \to Z$ in L^p. As we have $Z_m \to Z$ a.s., it is enough to show that Z_m is uniformly bounded in L^p. For a realization M' of M we have,

$$|Z_m| \leq \sum_{j=1}^{m \wedge M'} \frac{|Y_j - Y_{j-1}|}{\mathbb{P}(M \geq j)} \leq \frac{M'}{\mathbb{P}(M \geq M')}.$$

By assumption $M \in L^p$ so Z_m is uniformly bounded in L^p. This proves the first part of the proposition.

Before proving the additional two points, note that, since S is continuous,

$$\mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\right] \xrightarrow[m\to\infty]{} S(\mathbf{0}). \tag{3.45}$$

Now, let us prove the first point of the proposition. Assume that $M \in L^1$ and $\mathcal{X}$ is hyperuniform, so that $S(\mathbf{0}) = 0$. Since $\hat{S}_m$ is nonnegative, Equation (3.45) yields

$$\widehat{S}_m(\mathbf{k}_m^{\min}) \xrightarrow[m\to\infty]{L^1} 0.$$

Moreover, letting $f : x \mapsto 1 \wedge x$, $|f(x)| \leq x$ on $\mathbb{R}^+$, so that

$$\mathbb{E}\left[|f(\widehat{S}_m(\mathbf{k}_m^{\min}))|\right] \leq \mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\right] \to 0,$$

and

$$Y_m = f\left(\widehat{S}_m(\mathbf{k}_m^{\min})\right) \xrightarrow[m\to\infty]{L^1} 0. \tag{3.46}$$

Since $\mathbb{E}[Y_m] = \mathbb{E}[Z_m]$ and Z_m converges in L^1 to Z, by unicity of the limit, we have $\mathbb{E}[Z] = 0$.

It remains to show the last point of the proposition. Assume again that $M \in L^1$, but that $\mathcal{X}$ is not hyperuniform, so that $S(\mathbf{0}) > 0$. Reasoning by contradiction, assume that $\mathbb{E}[Z] = 0$. As $\mathbb{E}[Y_m] = \mathbb{E}[Z_m]$ and Z_m converges in L^1 to Z, we get

$$\mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\mathbb{1}_{\{\widehat{S}_m(\mathbf{k}_m^{\min})<1\}}\right] \xrightarrow[m\to\infty]{} 0, \tag{3.47}$$

and

$$\mathbb{E}\left[\mathbb{1}_{\{\widehat{S}_m(\mathbf{k}_m^{\min})\geq 1\}}\right] \xrightarrow[m\to\infty]{} 0. \tag{3.48}$$

Meanwhile,

$$\mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\right] \xrightarrow[m\to\infty]{} S(\mathbf{0}) > 0.$$

Using Equation (3.47), we get

$$\mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\mathbb{1}_{\{\widehat{S}_m(\mathbf{k}_m^{\min})\geq 1\}}\right] \xrightarrow[m\to\infty]{} S(\mathbf{0}) > 0. \tag{3.49}$$

Finally, Cauchy-Schwarz, together with Condition (3.44) and Equation (3.48) yield

$$\begin{aligned}\mathbb{E}\left[\widehat{S}_m(\mathbf{k}_m^{\min})\mathbb{1}_{\{\widehat{S}_m(\mathbf{k}_m^{\min})\geq 1\}}\right] &\leq \mathbb{E}^{1/2}\left[\widehat{S}_m^2(\mathbf{k}_m^{\min})\right] \times \mathbb{E}^{1/2}\left[\mathbb{1}_{\{\widehat{S}_m(\mathbf{k}_m^{\min})\geq 1\}}\right] \\ &\longrightarrow 0,\end{aligned}$$

which contradicts Equation (3.49) and ends the proof. □

Assumption (3.44) bears on the estimator that we use for the structure factor and the point process. We believe it not to be too strong and we prove it for the scattering intensity of a PPP.

Proof (validity of Assumption (3.44)). We show that Assumption (3.44) is satisfied for $\mathcal{X}$ a PPP of intensity ρ, $\widehat{S} = \widehat{S}_{\mathrm{SI}}$, and W_m are increasing rectangular windows.

Let $N_m = |\mathcal{X} \cap W_m|$, we have

$$\rho^2|W_m|^2\mathbb{E}\left[\widehat{S}_m^2(\mathbf{k})\right] =$$

$$\mathbb{E}\left(\left|\sum_{\mathbf{x}\in\mathcal{X}\cap W_m} e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}\rangle}\right|^2\right)^2 =$$

$$\mathbb{E}\left[N_m^2 + \left(\sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}} e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle}\right)^2 + 2N_m \sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}} e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle}\right] =$$

$$\begin{aligned}\mathbb{E}\Bigg[2N_m^2 - N_m + \sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}} 2N_m e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}2\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + \sum_{\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}\neq\mathbf{z}\neq\mathbf{t}} e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}+\mathbf{z}-\mathbf{t}\rangle} \\ + \sum_{\mathbf{x},\mathbf{y},\mathbf{z}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}\neq\mathbf{z}} 2e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}\langle\mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle} + e^{\mathrm{i}\langle\mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle}\Bigg] =\end{aligned}$$

$$\begin{aligned}\mathbb{E}\Bigg[2N_m^2 - N_m + \sum_{\mathbf{x},\mathbf{y}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}} 4e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}2\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + \sum_{\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}\neq\mathbf{z}\neq\mathbf{t}} e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}+\mathbf{z}-\mathbf{t}\rangle} \\ + \sum_{\mathbf{x},\mathbf{y},\mathbf{z}\in\mathcal{X}\cap W_m}^{\mathbf{x}\neq\mathbf{y}\neq\mathbf{z}} 4e^{-\mathrm{i}\langle\mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}\langle\mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle} + e^{\mathrm{i}\langle\mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle}\Bigg].\end{aligned}$$

Using the definition of the n-th order intensity measure $\mu^{(n)}$ (2.3) and that for any $n \geq 1$, the n-th order intensity function $\rho^{(n)}$ of $\mathcal{X}$ simplifies to ρ^n (Section 2.4) we get

$$\begin{aligned}
&\rho^2|W_m|^2\mathbb{E}\left[\widehat{S}_m^2(\mathbf{k})\right] = \\
&\mathbb{E}\left[2N_m^2 - N_m\right] + \int_{W_m\times W_m}\left(4e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}2\langle \mathbf{k},\mathbf{x}-\mathbf{y}\rangle}\right)\rho^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \\
&+ \int_{W_m\times W_m\times W_m}\left(4e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}-\mathbf{y}\rangle} + e^{-\mathrm{i}\langle \mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle} + e^{\mathrm{i}\langle \mathbf{k},2\mathbf{x}-\mathbf{y}-\mathbf{z}\rangle}\right)\rho^3\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{z} \\
&+ \int_{W_m\times W_m\times W_m\times W_m} e^{-\mathrm{i}\langle \mathbf{k},\mathbf{x}-\mathbf{y}+\mathbf{z}-\mathbf{t}\rangle}\rho^4\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{z}\mathrm{d}\mathbf{t}.
\end{aligned}$$

Now, using the parity of $\mathbb{1}_{W_m}$ and that N_m is a Poisson r.v., we get

$$\begin{aligned}
\mathbb{E}\left[\widehat{S}_m^2(\mathbf{k}_m^{\min})\right] =& \frac{1}{(\rho|W_m|)^2}\Bigg[\rho|W_m| + 2(\rho|W_m|)^2 + \rho^4\mathcal{F}^4(\mathbb{1}_{W_m})(\mathbf{k}_m^{\min}) \\
&+ \rho^2\left(4\mathcal{F}^2(\mathbb{1}_{W_m})(\mathbf{k}_m^{\min}) + \mathcal{F}^2(\mathbb{1}_{W_m})(2\mathbf{k}_m^{\min})\right) \\
&+ \rho^3\left(4|W_m|\mathcal{F}^2(\mathbb{1}_{W_m})(\mathbf{k}_m^{\min}) + 2\mathcal{F}(\mathbb{1}_{W_m})(2\mathbf{k}_m^{\min})\mathcal{F}^2(\mathbb{1}_{W_m})(\mathbf{k}_m^{\min})\right)\Bigg].
\end{aligned} \tag{3.50}$$

Upon noting that $\mathbf{k}_m^{\min} = (\frac{2\pi}{L_1}, \cdots, \frac{2\pi}{L_d})$ and

$$\mathcal{F}(\mathbb{1}_{W_m})(\mathbf{k}) = \prod_{j=1}^{d}\frac{\sin(k_jL_j/2)}{k_j/2}.$$

Equation (3.50) simplifies to

$$\mathbb{E}\left[\widehat{S}_m^2(\mathbf{k}_m^{\min})\right] = \frac{1}{\rho|W_m|} + 2.$$

Thus Assumption (3.44) holds. □

A multiscale test

We apply Proposition 3.3 with $p = 2$, say M is a Poisson random variable with mean $\lambda > 0$. In particular, $\mathbb{V}\mathrm{ar}[Z] < \infty$, and we can apply the central limit theorem to build a standard test comparing $\mathbb{E}[Z]$ to zero.

Consider A i.i.d. pairs $(\mathcal{X}_a, M_a)_{a=1}^A$ of realizations of $(\mathcal{X}, M)$, and let $Z_1, \ldots, Z_A$ be the A corresponding i.i.d. copies of Z. Now, denote the sample mean and

sample standard deviation of Z by $\bar{Z}_A$ and $\bar{\sigma}_A$. Since the variance of Z is finite, Slutsky's lemma yields the usual asymptotic confidence interval $CI[\mathbb{E}[Z]]$ of level ζ for $\mathbb{E}[Z]$,

$$\left[\bar{Z}_A - z\bar{\sigma}_A A^{-1/2}, \bar{Z}_A + z\bar{\sigma}_A A^{-1/2}\right], \tag{3.51}$$

where z is chosen such that $\mathbb{P}(-z < \mathrm{N}(0,1) < z) = \zeta$, and $\mathrm{N}(0,1)$ denotes the standard normal distribution. By Proposition 3.3, for an estimator of the structure factor satisfying (3.44), a test of hyperuniformity of asymptotic level ζ consists in assessing whether 0 lies in the interval (3.51). Since the estimator Z is based on windows of different sizes, we call the test *multiscale*.

3.3.3 Demonstrating the tests

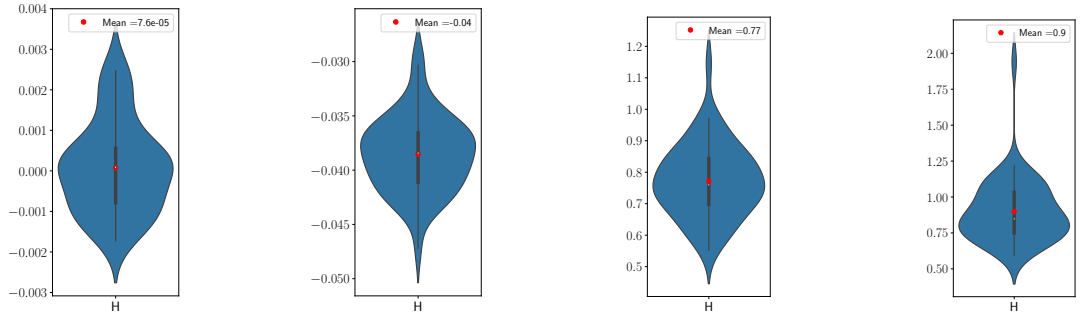
We now demonstrate the estimation of the hyperuniformity index H from Section 3.3.1, the multiscale hyperuniformity test from Section 3.3.2, and the estimation of the decay rate α from Section 3.1.2. All figures in this section can be reproduced using the toolbox structure_factor, by following our demonstration notebook [8]. For this demonstration, we used $A = 50$ samples from the point processes KLY, GPP, PPP and TPP with the same setup used for the experiments in Section 3.2.3. We additionally used thinned versions of the GPP with various retaining probabilities to test the accuracy of the multiscale hyperuniformity test; see Section 2.6.1 and Example 3.2. Recall that the KLY and the GPP are Class I hyperuniform while PPP, TPP and the thinned GPP are not hyperuniform.

Effective hyperuniformity

Figure 3.9 illustrates the violin plots[9] of the index H (3.40) across $A = 50$ samples from the benchmark point processes. We used the results of $\widehat{S}_{\mathrm{BI}}$, across A samples of roughly 10^4 points each (see Section 3.2.2). To fit the line required to compute H, we considered wavenumbers up to 0.6 for the TPP and 1 for the remaining

[8] https://github.com/For-a-few-DPPs-more/structure-factor/tree/main/notebooks

[9] A violin plot gathers a box plot and a kernel density estimator of the assumed underlying density. The former shows the median (white point), the interquartile range (thick black bar in the center), and the rest of the distribution except for points determined as outliers (thin black line in the center). We also add the mean (red dot).



(a) KLY (b) GPP (c) PPP (d) TPP

Figure 3.9: Violin plots of H across 50 samples from the KLY, GPP, PPP, and TPP. Note the different y-scales. Plots obtained using the toolbox structure_factor.

point processes. These values were chosen manually: the trade-off is to remain close to zero while including enough data points to fit a line. The violin plots of Figure 3.9 indicate that consistently across the realizations of the PPP and TPP, H is larger than, say, 0.5. This is a strong hint that these point processes are not hyperuniform. On the contrary, for the GPP, H is even slightly negative, hinting at hyperuniformity. For the KLY, although H is close to zero, we note that a threshold of 10^{-3} would not lead to the same answer across all 50 realizations.

Multiscale hyperuniformity test

Table 3.6: Multiscale hyperuniformity test

	$\bar{Z}_A$	$CI[\mathbb{E}[Z]]$	$\bar{Z}_A$	$CI[\mathbb{E}[Z]]$
KLY	0.003	$[-0.003, 0.009]$	0.003	$[-0.0003, 0.007]$
GPP	0.015	$[-0.021, 0.051]$	0.007	$[-0.003, 0.011]$
PPP	0.832	$[0.444, 1.220]$	0.781	$[0.560, 1.001]$
TPP	0.928	$[0.788, 1.068]$	1	$[0.999, 1]$
$\widehat{S}$	$\widehat{S}_{\mathrm{SI}}$		$\widehat{S}_{\mathrm{BI}}$	

Table 3.6 summarizes the results of the multiscale hyperuniformity test of Section 3.3.2, for the scattering intensity and Bartlett's isotropic estimators, on rectangular windows for the former and ball windows for the latter. To compute $\bar{Z}_A$, we used A draws of $(\mathcal{X}, M)$. In practice, the GPP cannot be sampled on arbitrarily large windows on a personal computer. We thus proceed as follows. Let

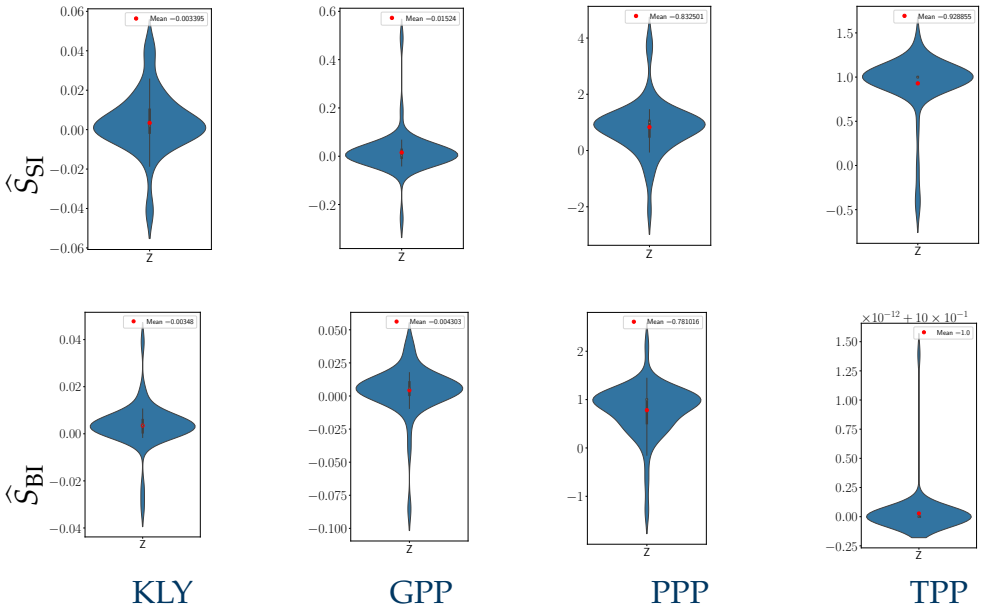


Figure 3.10: Violin plots of Z across 50 samples from the KLY, GPP, PPP, and TPP using $\widehat{S}_{\mathrm{SI}}$ (first line) and $\widehat{S}_{\mathrm{BI}}$ (last line). Plots obtained using the toolbox structure_factor.

$L_{\max}$ be the maximum sidelength, respectively maximum radius $R_{\max}$, of the box- (respectively ball-) shaped window on which the point process can be sampled in practice. Since M is a Poisson random variable with parameter λ, we choose λ such that the probability that W_M is larger than $W_{L_{\max}}$ (respectively $W_{R_{\max}}$) is small, say smaller than 10^{-4}. Precisely, for the scattering intensity, $\lambda = 85$, and the sidelength of the box window ranges from $L_{\min} = 20$ to $L_{\max} = 140$, with a unit stepsize for the GPP, the PPP and the TPP. The wavevectors used $\{\mathbf{k}_m^{\min}\}_{m\geq 1}$ are the minimum wavevectors of (3.13) corresponding to $\{W_m\}_{m\geq 1}$. Finally, the asymptotic confidence interval (3.51), denoted by $CI[\mathbb{E}[Z]]$ in Table 3.6, has a 99.7% asymptotic level since we use three standard deviations. For Bartlett's estimator, we use ball windows with a minimum radius $R_{\min} = 30$, a maximum radius $R_{\max} = 100$, a unit step size, $\lambda = 50$, and the minimum wavenumbers of (3.29) corresponding to the subwindows for the GPP, the PPP and the TPP. For the KLY having bigger intensity than the other benchmark point processes, we use smaller parameters $L_{\max} = 80$, $R_{\max} = 56$ and $L_0 = R_0 = 20$.

We can see from Table 3.6 that the test successfully rejects the hyperuniformity for the PPP and the TPP, and does not reject for the GPP and the KLY, as expected. Moreover, we note that $\widehat{S}_{\mathrm{BI}}$ provides tighter confidence intervals. Yet, one should bear in mind that, because $\widehat{S}_{\mathrm{SI}}$ uses a rectangular window and the GPP is naturally sampled on disk-like windows, $\widehat{S}_{\mathrm{BI}}$ has access here to a sample of larger cardinality than $\widehat{S}_{\mathrm{SI}}$. Moreover, $\widehat{S}_{\mathrm{BI}}$ is also computationally more expensive to evaluate.

Note that the coverage of the confidence interval on which our test is based is only controlled when A goes to infinity. For a fixed λ, we thus recommend choosing A as large as possible. In particular, the overly wide confidence interval for the PPP in Table 3.6 calls for increasing A. Indeed, the violin plot of the 50 realizations of Z (Figure 3.10) has large support. Increasing A naturally reduces the size of that support; see Figure 3.11. Note in passing how $\bar{Z}_A$ does not converge to $S(\mathbf{0}) = 1$, which is an effect of our capping the estimated structure factor in (3.41). Overall,

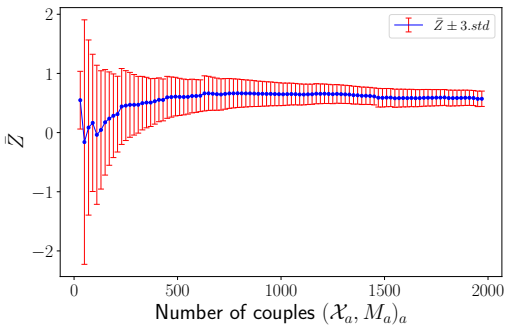


Figure 3.11: Estimated confidence intervals of $\bar{Z}$ as a function of the number of realizations used to get $\bar{Z}$ for a PPP. Z is based on the scattering intensity.

there is no free lunch: our test might fail in diagnosing hyperuniformity if A or λ is too small.

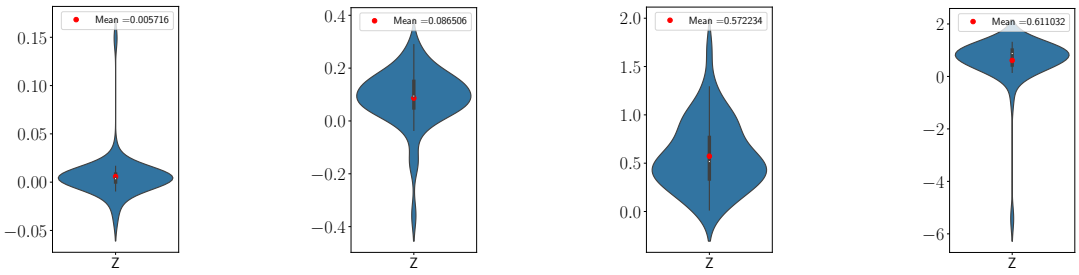
Finally, we demonstrate the test on three thinned versions of the GPP samples with retaining probabilities $p = \{0.9, 0.5, 0.1\}$. The corresponding point processes have respectively $S(0) = \{0.1, 0.5, 0.9\}$; see Example 3.2. We have noticed that the estimator $\widehat{S}_{\mathrm{BI}}$ provides tighter confidence intervals, and thus focused here on $\widehat{S}_{\mathrm{BI}}$ to compute $\bar{Z}_{50}$. We keep the same parameters as before. Note that we use the same 50 realizations of the GPP across the four different values of p (including

$p = 1$), so that, strictly speaking, the asymptotic confidence intervals are to be understood with a Bonferroni correction.

Table 3.7: Multiscale hyperuniformity test obtained using $\widehat{S}_{\mathrm{BI}}$ on thinned GPP

	$\bar{Z}_A$	$CI[\mathbb{E}[Z]]$
GPP	0.0057	$[-0.0042, 0.0156]$
Thinning $p = 0.9$	0.0865	$[0.0411, 0.1318]$
Thinning $p = 0.5$	0.5722	$[0.4227, 0.7217]$
Thinning $p = 0.1$	0.611	$[0.2082, 1.0137]$

Table 3.7 summarizes the obtained results, and Figure 3.12 shows the corresponding violin plots. The test successfully rejects the hyperuniformity for the three thinned versions. Note that the test is more sensitive when using $\widehat{S}_{\mathrm{SI}}$ within different trials. Nevertheless, the failure of the test can be noticed from the wide confidence intervals obtained in general. As mentioned before, getting a wide confidence interval call for increasing A. We recommend using $\widehat{S}_{\mathrm{BI}}$, if possible. Finally, as expected for the window sizes/intensity that we consider, the case $0.9 < p < 1$ remains difficult: in preliminary experiments, we did not reject hyperuniformity without hand-tuning the test's parameters to reach the desired conclusion. We leave this critical case to future work.

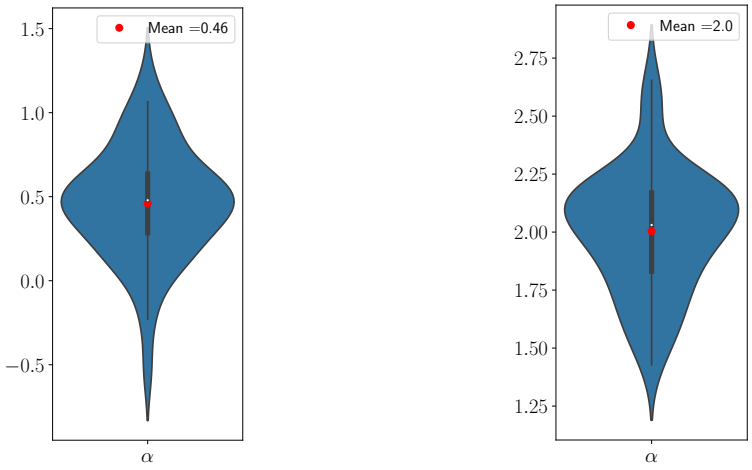


(a) GPP (b) $p = 0.9$ (c) $p = 0.5$ (d) $p = 0.1$

Figure 3.12: Violin plots of Z obtained using $\widehat{S}_{\mathrm{BI}}$ across $A = 50$ samples from the GPP and the corresponding independent thinned point process with retaining probability p. Plots obtained using the toolbox structure_factor.

Hyperunifomity class

Figure 3.13 shows the violin plots of the estimated power decay α (Section 3.1.2) across $A = 50$ samples, of roughly 10^4 points each, of the KLY and GPP. We used the results of $\widehat{S}_{\mathrm{BI}}$ (Section 3.2.3). To approximate the decay rate α of the structure factor, the maximum wavenumber used to fit the polynomial is equal to 0.45 for the GPP, and to 0.6 for the KLY. Again, these thresholds are fit manually and represent a trade-off between having enough points to fit our monomial and being close to zero. Figure 3.13 shows the obtained results. There is limited evidence that the KLY is indeed hyperuniform: while the support of the distribution is large and includes 0, most estimated values of α are positive and concentrate around 0.5. This could be further taken as a hint that KLY belongs to Class III in Table 3.1. Nevertheless, KLY is known to be Class I, so we'd expect α to be bigger than 1. The misleading concentration around 0.5 may result from the non-isotropy of the KLY and calls for caution when making claims on the hyperuniformity class from regression diagnostics. For the GPP, the concentration of values of α around 2 successfully reflects the known power decay (Section 3.1.2).



(a) KLY (b) GPP

Figure 3.13: Violin plots of α across A samples of KLY and GPP. Plots obtained using the toolbox structure_factor.

3.4 Python toolbox structure_factor

All the illustrations and experiments of this chapter were obtained using the Python toolbox structure_factor[10]. The toolbox is available on GitHub and released within the software repository for the Python programming language PyPI. For the installation procedure, we refer to the dedicated guidelines provided on the toolbox's GitHub README page.

The main objective of structure_factor is to estimate the structure factor of a stationary point process $\mathcal{X}$ using any of the examined estimators in Section 3.2 as well as to explore the hyperuniformity of $\mathcal{X}$ through the tests outlined in Section 3.3. This section offers a concise overview of the three core components within structure_factor essential for achieving these objectives. In Section 3.4.1, we present the tools designed for generating a point pattern of $\mathcal{X}$ in a suitable format. This prepared sample is subsequently used in Section 3.4.2 to estimate the structure factor of $\mathcal{X}$. Finally, Section 3.4.3 is concerned with testing the hyperuniformity of $\mathcal{X}$. For more details, readers are encouraged to explore the online documentation[11] as well as the tutorial Jupyter Notebook[12].

3.4.1 Generating a `PointPattern`

Consider a stationary point process $\mathcal{X}$ of intensity ρ. In a nutshell, an object of type `PointPattern` of $\mathcal{X}$ encapsulates a sample of $\mathcal{X} \cap W = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, the observation window W, and the intensity ρ (optional). Some comments are in order. First, one can obtain a `PointPattern` of the PPP, TPP, GPP, and KLY using the corresponding objects accessible within the module `point_processes`. Second, the currently supported window shapes are spherical via the object `BallWindow` and rectangular via `BoxWindow`. Third, if the intensity of $\mathcal{X}$ is not provided by the user, it is automatically approximated by the asymptotically unbiased estimator $\widehat{\rho} = \frac{N}{|W|}$. Finally, to comply with the requirements of the structure factor estima-

[10] https://github.com/For-a-few-DPPs-more/structure-factor

[11] https://for-a-few-dpps-more.github.io/structure-factor/

[12] https://github.com/For-a-few-DPPs-more/structure-factor/blob/main/notebooks/tutorial_structure_factor.ipynb

tors on specific windows, as outlined in Assumptions A1 and A2, we provide a `restrict_to_window` method for the class `PointPattern`. This feature enables the restriction of the realization of $\mathcal{X}$ within `PointPattern` to another observation window specified by the user.

3.4.2 Estimating the structure factor

All the structure factor estimators from Section 3.2 are implemented as methods of the class `StructureFactor`. The class constructor takes as input an object of type `PointPattern`; see Section 3.4.1.

Given that certain isotropic estimators necessitate the calculation of the pair correlation function, the module `pair_correlation_function` hosts the functions `estimate`, and `interpolate`. `estimate` is tailored to gauge the pair correlation function through the estimator `pcf.ppp`, or `pcf.fv` originally implemented in the R package spatstat. Meanwhile, `interpolate` is dedicated to interpolating the obtained discrete approximation. Further details are mentioned in Section 3.2.3.

Finally, note that each estimator is complemented by a `plot` function that enhances the visualization of the resultant outcomes with an option to get the same output as in the figures of Section 3.2.3.

3.4.3 Testing the hyperuniformity

The hyperuniformity tests outlined in Section 3.3 are functions of the module `hyperuniformity`. Within this module, the function `multiscale_test` implements the statistical test of hyperuniformity (Section 3.3.2). `multiscale_test` is complemented by the function `subwindows` designed to derive a list of subwindows, from a window specified by the user. It also provides the allowed wavevectors corresponding to the output subwindows. These subwindows and allowed wavevectors are required inputs of the function `multiscale_test`.

The module `hyperuniformity` also contains the function `hyperuniformity_class` devised for estimating the power decay α of the structure factor near the origin. As highlighted in Section 3.1.2, the value of α indicates the hyperuniformity class.

The estimation procedure to get α involves a simple polynomial fitting around the origin of an estimation of the structure factor that can be obtained using the class `StructureFactor` as outlined in Section 3.4.2.

3.5 Conclusion

In this chapter, we surveyed estimators of the structure factor of a stationary point process, along with numerical diagnostics of hyperuniformity. The two estimators that fared best in our benchmark are the multitapered variant of the scattering intensity and, in the case of isotropic point processes, Bartlett's isotropic estimator. Further comparing these two estimators is a natural avenue for further work, but needs careful thinking. For starters, the two estimators are not defined on similar windows, and they do not require the same number of evaluations for a similar accuracy. Selecting the optimal tapers for the multitaper estimator is currently unclear. On the other hand, the computational cost of Bartlett's estimator should be lowered, e.g. by subsampling pairs of points. Estimators based on Hankel transform quadratures, which rely on first estimating the pair correlation function, comparatively showed poor performance. But it is still possible that estimating the pair correlation first can help estimating the structure factor if, say, edge correction plays an important role. If the user only has data collected on a non-rectangular, non-ball window, we would then recommend trying the Hankel-Baddour-Chouinard estimator. For our benchmarks, though, the large cardinalities and the regular windows involved do not build upon this strong point of quadrature-based estimators. On another note, since the structure factor is a Fourier transform, one might be tempted to periodize one's sample before computing estimators. However, we do not see a clear argument for (or against) periodization yet. First, if one is using the scattering intensity applied on any of the restricted allowed wavectors in $\mathbb{A}_L^{res}$, then periodizing the point process has no effect on the estimator. The case of wavevectors in $\mathbb{A}_L$, and of other estimators, is less straightforward. Second, adding periodic boundaries forces an arbitrary regularity at long distances, which intuitively would impact any empirical diagnostic of hyperuniformity.

We also proposed an (asymptotic) test of hyperuniformity, based on the debiasing techniques of (Rhee and Glynn, 2015). The test relies on the coupled sum estimator Z obtained from a sequence of estimation of the structure factor derived from the point pattern observed in a list of increasing-size windows. We proved that $\mathbb{E}[Z]$ is zero if and only if the corresponding point process is hyperuniform (under mild assumptions) and proposed employing Z to test the hyperuniformity in practice. The test we proposed involves finding a confidence interval for $\mathbb{E}[Z]$ and checking whether zero falls within this interval as a sign of hyperuniformity. Our experiments demonstrated the test's power of rejecting the hyperuniformity for point processes with $S(\mathbf{0})$ exceeding 0.1 along moderate sample sizes. However, for $S(\mathbf{0})$ below this threshold, further careful investigation is required. We postpone this investigation for future work. Moreover, the actual proposed test requires many realizations from the point process which may be a limitation in some practical applications. The next natural step is to explore the possibility of employing a single realization of moderate size.

Finally, we provided an open-source Python toolbox structure_factor containing all discussed estimators of the structure factor, diagnostics of hyperuniformity, and our statistical test of hyperuniformity. We used structure_factor to benchmark the estimators, diagnostics, and tests. Moreover, we present an introduction to the main objects of the toolbox. We hope that our survey and software can contribute to the standardization and reproducibility of empirical investigations involving structure factors, including the study of hyperuniform point processes.

4 — Repelled point processes

Contents

Linear statistics of point processes yield Monte Carlo estimators of integrals. While the simplest approach relies on a homogeneous Poisson point process (PPP), more regularly spread point processes, such as scrambled low-discrepancy sequences or determinantal point processes, yield Monte Carlo estimators with fast-decaying root mean square error (RMSE); see Section 2.7.

Motivated by the efficiency of regular configurations in numerical integration, in this chapter, we introduce the repulsion operator, which aims to reduce clustering in a configuration by slightly pushing its points away from each other. Namely, given a configuration of points X and a parameter $\varepsilon > 0$, we explain how to construct another configuration $\Pi_\varepsilon(\mathsf{X})$, called the *repelled configuration*, using the repulsion operator Π_ε. Keeping in mind our motivation for numerical integration, we want Π_ε to be (i) computationally cheap to apply. Moreover, when applied to a random configuration, Π_ε should (ii) preserve stationarity and intensity, and (iii) reduce the variance of linear statistics compared to a PPP. While the first condition ensures that sampling from $\Pi_\varepsilon(\mathsf{X})$ remains relatively tractable, the latter two guarantee that $\Pi_\varepsilon(\mathsf{X})$ offers a Monte Carlo method with reduced RMSE compared to the crude Monte Carlo. To achieve these objectives, the repulsion operator employs a force that leads the interaction between the points of the configuration drawing inspiration from differential equation schemes used in gravitational allocations; see Example 2.5. In a few words, Π_ε can be seen as performing a *single* step of a numerical scheme to integrate a differential equation corresponding to a gravitational allocation.

The investigation described in this chapter was carried out in our paper (Hawat et al., 2023a). In Section 4.1, we introduce the repulsion operator, while Section 4.2 is dedicated to examining the properties of the force function which is the driving gear of the repulsion operator. The crux of this chapter lies in Section 4.3 where we analyze the properties of the repulsion operator when applied to a PPP. The obtained point process is denoted repelled Poisson point process (RPPP). We present our main result, showing that the RPPP yields an unbiased Monte Carlo method with *lower* variance than under the PPP. Additionally, we describe a sampling procedure that we implemented within the Python toolbox MCRPPy[1]. We present an experimental illustration of the variance reduction and estimate the pair correlation function, and the structure factor of the RPPP using the toolbox structure_factor. Moreover, we explore iterating the repulsion operator several times and draw an analogy to gravitational allocations. Lastly, we put the Monte Carlo with the RPPP in context, by conducting a comparison with standard Monte Carlo methods on synthetic integration tasks. In Section 4.4, we

[1] https://github.com/dhawat/MCRPPy

explore applying the repulsion operator to already repulsive point processes. The experiments suggest that the variance reduction also holds when the operator is applied to motion-invariant point processes other than the PPP. All the proofs are gathered in Section 4.5. Finally, Section 4.6 concludes this chapter with a few research directions.

4.1 The repulsion operator

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

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

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 $F_{\mathsf{X}}(\mathbf{x})$ 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 4.2 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. When $d = 3$, $-F_{\mathcal{P}}$ corresponds to the ordinary Newtonian gravitational force, upon which, Chatterjee et al. (2010) build a gravitational allocation from $\mathcal{P}$ to the Lebesgue measure; see also Example 2.5. Further information regarding the characteristics of $F_{\mathcal{P}}$ can be found in Section 4.2.

Call $\mathsf{X} \in \mathfrak{N}$ a *valid* configuration if, for all $\mathbf{x}$, the limit defining $F_{\mathsf{X}}(\mathbf{x})$ in (F_1) exists. For $\varepsilon \in \mathbb{R}$, we define the (Coulomb) repulsion[2] operator Π_ε, acting on valid configurations, through

$$\Pi_\varepsilon : \mathsf{X} \longmapsto \{\mathbf{x} + \varepsilon F_{\mathsf{X}}(\mathbf{x}) : \mathbf{x} \in \mathsf{X}\}. \tag{4.1}$$

[2]While we generally speak of "repulsion", note that when $\varepsilon < 0$ the dynamics become attractive instead of repulsive.

There are two formal caveats to our definition (4.1). 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 Π_ε. Anticipating a bit, Corollary 4.1 shows that these two caveats are irrelevant when Π_ε is applied to a PPP $\mathcal{P}$. 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 $\Pi_\varepsilon \mathcal{X}$ of a more general point process $\mathcal{X}$, although its existence needs to be discussed. Once the existence is verified, we have the following motion-invariance property.

Proposition 4.1 (Motion-invariance)**.** *Let $\mathcal{X}$ be a point process that is almost surely valid, and $\varepsilon \in \mathbb{R}$. If $\mathcal{X}$ is motion-invariant, then $\Pi_\varepsilon \mathcal{X}$ is also motion-invariant.*

The proof of this proposition is deferred to Section 4.5.2.

4.2 Properties of the force

In this section, we discuss key characteristics of the random function $F_{\mathcal{P}}$, when $\mathcal{P}$ is a PPP.

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 4.2. *Let $\mathcal{P}$ be a homogeneous Poisson point process of intensity ρ 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 [3] of index $\alpha = \frac{d}{d-1}$.*

[3] A random vector $\mathbf{x} \in \mathbb{R}^d$ has an $\alpha-$ stable distribution iff there exists $\alpha \in (0, 2]$ such that for any $\mathbf{x}_1, ..., \mathbf{x}_n$ i.i.d. copies of $\mathbf{x}$, there exits $\mathbf{b}_n \in \mathbb{R}^d$ such that $\sum_{i=1}^n \mathbf{x}_i \overset{d}{=} n^{1/\alpha}\mathbf{x} + \mathbf{b}_n$. If in additional $\mathbf{x} \overset{d}{=} -\mathbf{x}$, $\mathbf{x}$ is said $\alpha-$ symmetric stable distribution (Abdul-Hamid and Nolan, 1998). While the

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 ρ 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 $-\mathcal{P} \overset{d}{=} \mathcal{P}$. Proposition 4.2 implies that $\mathbb{E}\left[F_{\mathcal{P}}(\mathbf{x})\right] = 0$ and $\mathbb{E}\left[\|F_{\mathcal{P}}(\mathbf{x})\|_2^{\nu}\right] < \infty$ iff $\nu < \alpha$. 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 4.3. *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 $\mathbf{c} \in \mathbb{R}^d$*

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

The proof of this proposition is deferred to Section 4.5.1. We note that an additional result regarding the joint density of $(F_{\mathcal{P}}(\mathbf{x}), F_{\mathcal{P}}(\mathbf{y}))$ 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

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

where κ_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 the summation. Chatterjee et al. (2010, Proposition 5) proved that the expressions (F_1) and (F_2) are equivalent when $\mathcal{P}$

multivariate stable distribution is a significant theoretical concept as it extends the multivariate normal distribution, it is not widely employed in practical applications. This is primarily because of its heavy-tailed behavior and the absence of a closed-form expression for its density function.

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

$$F_{\mathcal{X}}^{(q,p)}(\mathbf{x}) \triangleq \sum_{\substack{\mathbf{z} \in \mathcal{X} \setminus \{\mathbf{x}\} \cap A^{(q,p)}(\mathbf{x}) \\ \|\mathbf{x}-\mathbf{z}\|_2 \uparrow}} \frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|_2^d}, \tag{4.2}$$

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 the 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

$$\Pi_{\varepsilon}^{(q,p)} \mathcal{X} \triangleq \{\mathbf{x} + \varepsilon F_{\mathcal{X}}^{(q,p)}(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}. \tag{4.3}$$

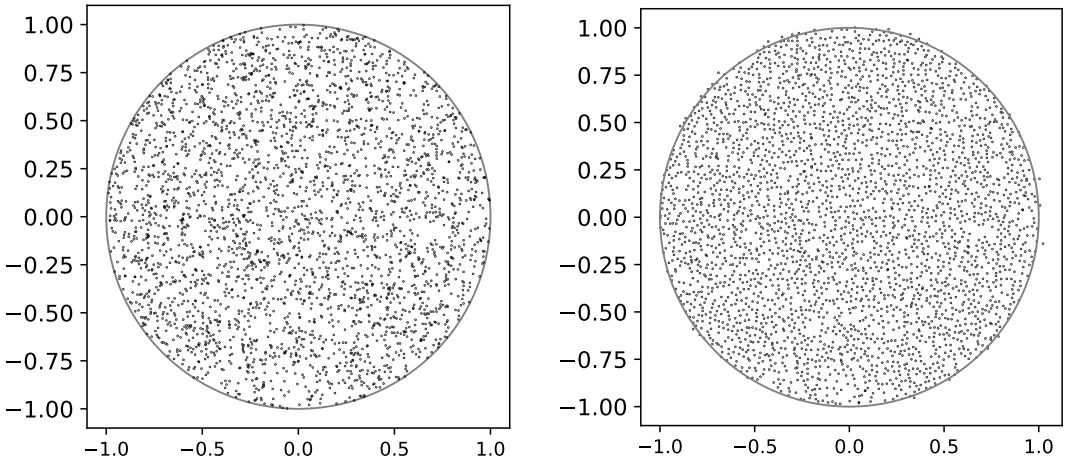
4.3 The repelled Poisson point process

We have seen in Section 4.1 that a PPP is a valid configuration for the operator Π_ε (Equation (4.1)). In this section, we study $\Pi_\varepsilon \mathcal{P}$, the repelled Poisson point process (RPPP) obtained by applying Π_ε to a PPP $\mathcal{P}$ of intensity $\rho > 0$. The first panel of Figure 4.1 displays a sample from a PPP of intensity $\rho = 1000$ in two dimensions[4], intersected with a disk-shaped observation window. We illustrate the corresponding RPPP in the second panel.

4.3.1 Properties

In this section, we state some properties of the RPPP. Of particular importance to us is the following corollary.

[4]Note that for graphical convenience we present a two-dimensional construction, but we are not making any convergence claim for the force (F_1) in $d = 2$.



(a) PPP (b) RPPP

Figure 4.1: A sample from a PPP of intensity 1000 and the corresponding RPPP. Plot obtained using the toolbox MCRPPy.

Corollary 4.1. *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 ρ.

The proof of this corollary is deferred to Section 4.5.2. According to Corollary 4.1, $\Pi_\varepsilon \mathcal{P}$ is of intensity ρ, the same intensity as $\mathcal{P}$. Consequently, for any integrable function f of compact support K, Equation (2.11) yields

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

where $I(f)$ is the integral of f defined in (2.22). In particular

$$\widehat{I}_{\Pi_\varepsilon\mathcal{P}}(f) \triangleq \frac{1}{\rho}\sum_{\mathbf{x}\in\Pi_\varepsilon\mathcal{P}} f(\mathbf{x}) \tag{4.4}$$

is an unbiased Monte Carlo estimator of $I(f)$; see also Section 2.7. We shall also consider the so-called "self-normalized" estimator

$$\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}(f) \triangleq \frac{|K|}{\Pi_\varepsilon\mathcal{P}(K)}\mathbb{1}_{\{\Pi_\varepsilon\mathcal{P}(K)>0\}}\sum_{\mathbf{x}\in\Pi_\varepsilon\mathcal{P}\cap K} f(\mathbf{x}), \tag{4.5}$$

where $\Pi_\varepsilon\mathcal{P}(K)$ is the number of points of $\Pi_\varepsilon\mathcal{P}$ in K. Compared to (4.4), (4.5) replaces ρ by an unbiased estimator. Self-normalized estimators are frequent in spatial statistics, and one can expect a (small) variance reduction in (4.5) at the price of a small bias.

Remark 4.1. The self-normalized estimator (4.5) of $I(f)$ is biased. Indeed

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

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

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

where the void probability $\mathbb{T}_{\Pi_\varepsilon\mathcal{P}}$ is defined by Equation (2.1). As K grows, the bias thus decreases. It is actually reasonable to expect that it vanishes exponentially fast with the size of K.

The following result ensures that the RPPP has moments of any order.

Proposition 4.4 (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 Section 4.5.3, and we note that a quantitative upper bound of the expectation can be deduced from the proof. A direct consequence of Proposition 4.4 is that for any continuous function f of compact support K, we have

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

Once the existence of the variance of linear statistics under $\Pi_\varepsilon\mathcal{P}$ is established, we delve further into our investigation. At this point, we present the following variance reduction result. Its proof is deferred to Section 4.5.4.

Theorem 4.5 (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*

$$\mathbb{V}ar\left[\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f)\right] = \mathbb{V}ar\left[\widehat{I}_{\mathcal{P}}(f)\right](1 - 2d\kappa_d \rho \varepsilon) + O(\varepsilon^2), \tag{4.6}$$

where $\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f)$ is defined in (4.4),

$$\widehat{I}_{\mathcal{P}}(f) \triangleq \frac{1}{\rho} \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}), \tag{4.7}$$

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

Several remarks are in order.

Remark 4.2. Upon noting that $\Pi_0 \mathcal{P} = \mathcal{P}$, Equation (4.6) implies a negative derivative of the variance of $\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f)$ at $\varepsilon = 0$. Actually,

$$\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\Pi_\varepsilon \mathcal{P}}(f)\right] < \mathbb{V}\mathrm{ar}\left[\widehat{I}_{\mathcal{P}}(f)\right] = \rho^{-1} I(f^2), \tag{4.8}$$

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 Sections 4.2 and 4.5.4.

Remark 4.3. Taking ε equal to

$$\varepsilon_0 \triangleq \frac{1}{2d\kappa_d \rho} \tag{4.9}$$

makes the term of order ε in (4.6) vanish. Note also that ε_0 does not depend on the integrand f.

Remark 4.4. When $\varepsilon < 0$, we obtain a positive first-order derivative of the variance at $\varepsilon = 0$, so for $|\varepsilon|$ small enough we have

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

This result is expected as the behavior of Π_ε shifts from repulsive to attractive.

Remark 4.5. (Truncating the force) The proof of Theorem 4.5 holds even when we replace $F_{\mathcal{P}}$ with its truncated version $F_{\mathcal{P}}^{(0,p)}$ (4.2), as long as p is larger than the diameter of the support K of the integrand.

Remark 4.6. A key element of the proof of Theorem 4.5 is the super-harmonicity[5] 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

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

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 also Remark 4.8.

Remark 4.7. 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 potential

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

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

4.3.2 Sampling

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 $\operatorname{diam}(K)$. In this section, we propose two approaches to approximately sampling from $\Pi_\varepsilon\mathcal{P} \cap K$. By Corollary 4.1, $\Pi_\varepsilon\mathcal{P}$ is stationary, and we henceforth assume that $K \subseteq B(\mathbf{0}, p)$, where $p = \operatorname{diam}(K)/2$.

Our first approach is simply to sample $\Pi_\varepsilon^{(0,p)}\mathcal{P} \cap K$. This seems reasonable in the context of application to numerical integration since the variance reduction result also holds for $\Pi_\varepsilon^{(0,p)}\mathcal{P}$; see Remark 4.5. The corresponding pseudo-code is provided by Algorithm 1. In words, we use the points of $\mathcal{P}$ that fall in the larger ball $B(\mathbf{0}, 2p)$ to displace the points of $\mathcal{P} \cap K$. Informally, for large 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

[5] A C^2 function ϕ on an open set A is super-harmonic if and only if the Laplacian $\Delta\phi$ of ϕ is negative on A.

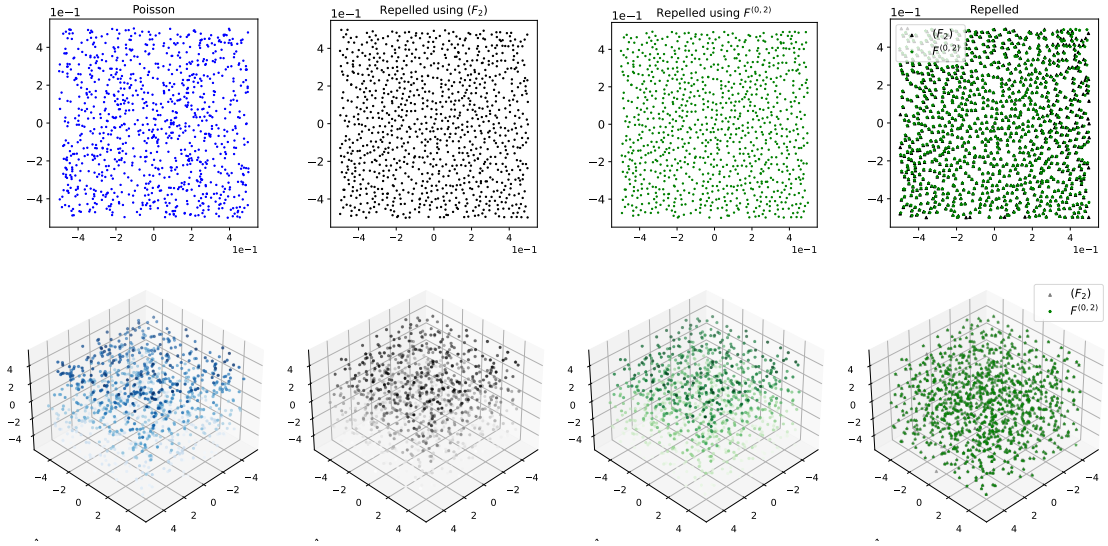


Figure 4.2: PPP samples (first column) and the corresponding RPPP samples, obtained using Algorithm 2 (second column) and using Algorithm 1 (third column). The first row corresponds to $d = 2$ and the second row to $d = 3$, with ε set in each row to the value $\varepsilon_0 = \varepsilon_0(d)$ in (4.9). The last column shows the two repelled samples superimposed. Plots obtained using the toolbox MCRPPy.

than p 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 find the points of $\mathcal{P}$ located in $B(\mathbf{x}, p)$. While storing the initial sample of $\mathcal{P} \cap B(\mathbf{0}, 2p)$ 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.

Indeed, our second approach stems from the fact that the partial sums of (F_2) use the same points of $\mathcal{P}$, independently of $\mathbf{x}$. The correction term in (F_2) seems partially taking into account the effects of using a fixed order in the sum, so we propose to sample $\mathcal{P} \cap B(\mathbf{0}, p)$ and use the points in the latter sample to displace the points of $\mathcal{P} \cap K$ using (F_2) as described by Algorithm 2. We still need to compute for each $\mathbf{x} \in \mathcal{P} \cap K$ the distances between $\mathbf{x}$ and the points of $\mathcal{P} \cap B(\mathbf{0}, p)$, which in total cost $O(NM)$ operations, where $M = \mathcal{P}(K)$ and $N = \mathcal{P}(B(\mathbf{0}, p))$.

Some comments are in order. First, the expected number of points of $\Pi_\varepsilon \mathcal{P} \cap K$ is equal to $\rho|K|$. Second, in both sampling approaches, the points of $\Pi_\varepsilon \mathcal{P} \cap K$ can be sampled concurrently, resulting in a reduction in computational time roughly

Algorithm 1 : Sampling a repelled point pattern of a point process $\mathcal{P}$ in a centered observation window K using (F_1)

fix ε
set $p = \mathrm{diam}(K)/2$, and $k = 1$
sample a point pattern $\{\mathbf{x}_i\}_{i=1}^N$ from the point process $\mathcal{P}$ in the centered ball window $B(\mathbf{0}, 2p)$
sort $\{\mathbf{x}_i\}_{i=1}^N$ in a KD-tree
parfor $i : 1 \to N$ **do**
    if $\mathbf{x}_i \in K$ **then**
        search in the KD-Tree the points $\{\mathbf{x}_j\}_j$ located in $B(\mathbf{x_i}, p) \setminus \{\mathbf{x}_i\}$
        use $\{\mathbf{x}_j\}_j$ to compute the truncated force $F_{\mathcal{P}}^{(0,p)}(\mathbf{x}_i)$ at $\mathbf{x}_i$ via Equation (4.2)
        set $\mathbf{y}_k = \mathbf{x}_i + \varepsilon F_{\mathcal{P}}^{(0,p)}(\mathbf{x}_i)$, and $k = k + 1$
    end if
end parfor
return $\{\mathbf{y}_k\}_k \cap K$

proportional to the number of available processors. This parallelization appears in Algorithm 1, and 2 through a *parfor* loop. Third, currently, we do not have a strategy in place to mitigate border effects without increasing the computational cost. We recommend, if possible, sampling $\Pi_\varepsilon \mathcal{P} \cap W$, where W is a window slightly larger than K, and then restricting the obtained sample to the target window K. In our upcoming experiments, we use $W = B(\mathbf{0}, \mathrm{diam}(K)/2)$. Finally, we provide a Python package, called MCRPPy, available on GitHub[6], which implements the two sampling methods.

Figure 4.2 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 (Algorithm 2), while $F_{\mathcal{P}}^{(0,2)}$ was used in the third column (Algorithm 1). The last column is a superposition of the samples obtained in columns 2 and 3, displaying very close agreement. Finally, note that Figures 4.1 and 4.3 were obtained using Algorithm 2, and we

[6] https://github.com/dhawat/MCRPPy

will keep using this simulation method in the next sections for sampling from the repelled point process of a stationary point process which may not necessarily be the PPP.

Algorithm 2 : Sampling a repelled point pattern of a point process $\mathcal{P}$ in a centered observation window K using (F_2)

fix ε
set $p = \mathrm{diam}(K)/2$, and $k = 1$
sample a point pattern $\{\mathbf{x}_i\}_{i=1}^N$ from the point process $\mathcal{P}$ in the centered ball window $B(\mathbf{0}, p)$
parfor $i : 1 \to N$ **do**
    if $\mathbf{x}_i \in K$ **then**
        use $\{\mathbf{x}_i\}_{i=1}^N$ to compute the force $F_{\mathcal{P}}(\mathbf{x}_i)$ at $\mathbf{x}_i$ via Equation (F_2)
        set $\mathbf{y}_k = \mathbf{x}_i + \varepsilon F_{\mathcal{P}}(\mathbf{x}_i)$, and $k = k + 1$
    end if
end parfor
return $\{\mathbf{y}_k\}_k \cap K$

4.3.3 An experimental illustration of the variance reduction

In this section, we present a numerical experiment to confirm the variance reduction found in Theorem 4.5. Additional experiments can be found in Section 4.3.6.

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

$$f_1(\mathbf{x}) \triangleq \left(1 - 4\|\mathbf{x}\|_2^2\right)^2 \exp\left(\frac{-2}{1 - 4\|\mathbf{x}\|_2^2}\right) \mathbb{1}_{B(\mathbf{0},1/2)}(\mathbf{x}),$$

$$f_2(\mathbf{x}) \triangleq \mathbb{1}_{B(\mathbf{0},1/2)}(\mathbf{x}), \quad \text{and} \quad f_3(\mathbf{x}) \triangleq \prod_{i=1}^{d} \cos^3(\pi x_i) \sin(\pi x_i) \mathbb{1}_K(\mathbf{x}). \tag{4.10}$$

Both f_1 and f_3 satisfy the requirements of Theorem 4.5, while the indicator function f_2 is discontinuous on $\partial B(\mathbf{0}, 1/2)$. For each of these functions, Figure 4.3 shows the estimated standard deviations $\widehat{\sigma}(\widehat{I}_{s,\Pi_\varepsilon \mathcal{P} \cap K}(.))$ of the self-normalized

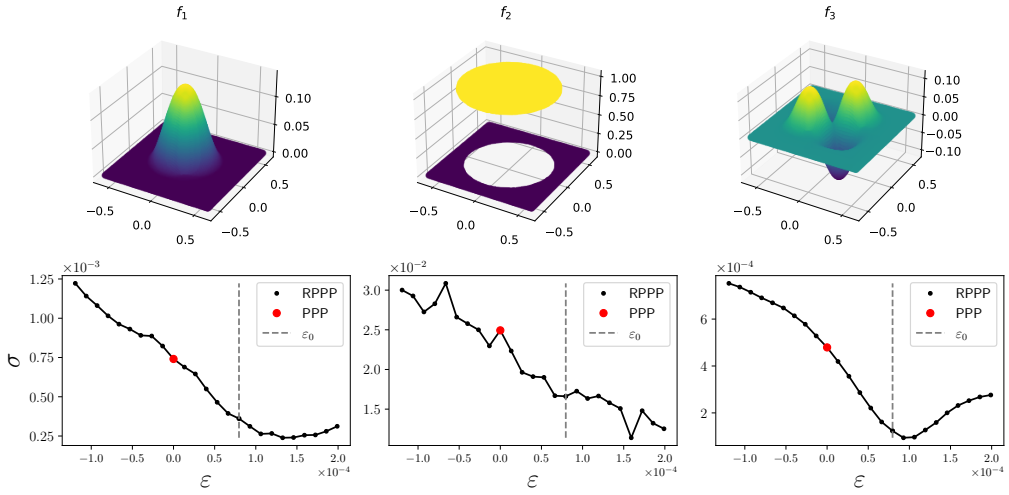


Figure 4.3: Estimated standard deviations of $\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}$ with respect to ε, for f_1, f_2, and f_3, in $d = 3$. Plots obtained using the toolbox MCRPPy.

estimator $\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}$ (4.5) for varying values of ε in $d = 3$. We conducted the analysis using 50 independent samples of $\mathcal{P}$, a PPP 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 $\widehat{\sigma}(\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}(.))$. The dashed lines indicate ε_0 in (4.9). Note that within the range of ε 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 ε, $\widehat{\sigma}(\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}(.))$ are bigger than $\widehat{\sigma}(\widehat{I}_{s,\mathcal{P}\cap K}(.))$, for the three functions. This behavior is expected because the operator Π_ε is attractive for negative values of ε. Second, for positive values of ε, up to ε_0, we observe that $\widehat{\sigma}(\widehat{I}_{s,\Pi_\varepsilon\mathcal{P}\cap K}(.))$ are lower than $\widehat{\sigma}(\widehat{I}_{s,\mathcal{P}\cap K}(.))$. This result aligns with our theoretical expectations and provides evidence for the variance reduction in Theorem 4.5. Third, for f_1 and f_3 we observe an interesting trend when ε exceeds ε_0. The standard deviations decrease until reaching a minimum value. This minimum value, particularly for f_3, is relatively close to ε_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 ε_0 is a reasonable choice for ε, regardless of the integrand, although not necessarily the optimal threshold for a specific integrand.

4.3.4 Second-order properties

We have seen in Chapter 2 that the second-order characteristics of a point process, such as the pair correlation function (2.5) and the structure factor (2.8), offer valuable insights into the regularity of the point process. Specifically, a pair correlation function less than one indicates repulsion within the points of the point process (Section 2.5). On the other hand, a structure factor equal to zero at the origin indicates hyperuniformity, as seen in Chapter 3. In this section, we estimate the pair correlation function and the structure factor of the RPPP to examine whether they reveal any indications of repulsion or hyperuniformity.

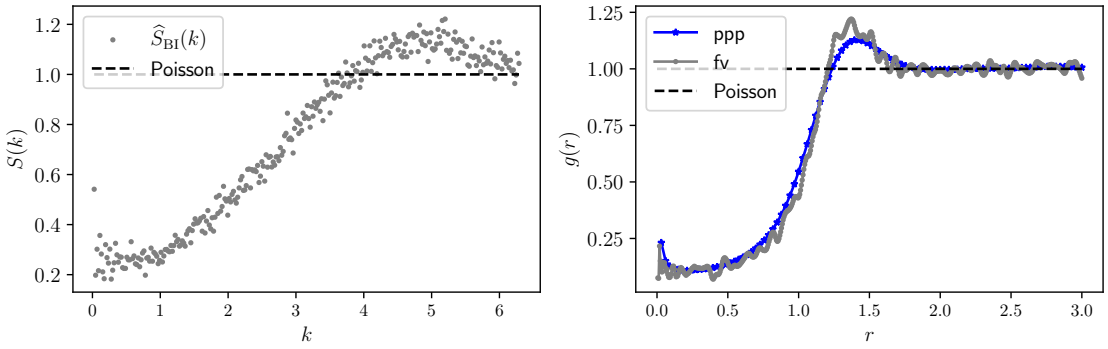


Figure 4.4: Estimated structure factor S (left) and pair correlation function g (right) of a RPPP of $\mathbb{R}^2$ of intensity $1/\pi$. Plots obtained using the toolbox structure_factor.

Figure 4.4 shows the estimated pair correlation function (2.6) and structure factor (2.8) of a RPPP of intensity $1/\pi$ in dimension $d = 2$ obtained using the toolbox structure_factor. We used a sample from the RPPP observed in a centered ball of radius $r = 150$; see Section 4.3.2. Bartlett's isotropic estimator $\widehat{S}_{\mathrm{BI}}$ (Equation (3.27)) was used to estimate the structure factor on the corresponding set of allowed wavenumbers k (Equation (3.29)). The estimators pcf.ppp and pcf.fv (Baddeley et al., 2015, Sections 7.4.4 and 7.4.5) were used to estimate the pair correlation function, as in Section 3.2.3. For reference, dashed black lines represent the structure factor and pair correlation function of the PPP. As expected, the estimated values of the pair correlation function are smaller than 1 up to a certain value of $r > 1$, indicating repulsion between the points at small scales. However,

the values of $\widehat{S}_{\mathrm{BI}}(k)$ are bigger than 0.2 at small k, suggesting that the RPPP may not be a hyperuniform point process.

4.3.5 Iterating the repulsion

Our repulsion operator Π_ε in (4.1) can be seen as performing one step of a numerical discretization scheme for a system of differential equations describing gravitational allocation schemes. In this section, we investigate iterating the application of Π_ε, by drawing an analogy to gravitational allocations (Example 2.5).

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

$$\frac{\partial}{\partial t} Y_{\mathbf{x}}(t) = -F_{\mathcal{P}}(Y_{\mathbf{x}}(t)) \quad \text{with} \quad Y_{\mathbf{x}}(0) = \mathbf{x}. \tag{4.11}$$

The solution $t \mapsto Y_{\mathbf{x}}(t)$ of the differential equation (4.11), 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 also Example 2.5. 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 (4.11), 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 (4.11), 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 (4.11). 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 gravitational 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 thesis.

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

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \varepsilon F_{\mathcal{P}}(\mathbf{x}_{t-1}), \quad t = 1, \dots, M. \tag{4.12}$$

Call $\Pi_{\varepsilon,t} \mathcal{P}$ the set of $t-$th iterates (4.12) of the points of $\mathcal{P}$. We expect that for $\varepsilon < 0$ and large t, 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 4.5; see also Figure C.1. It is important to note

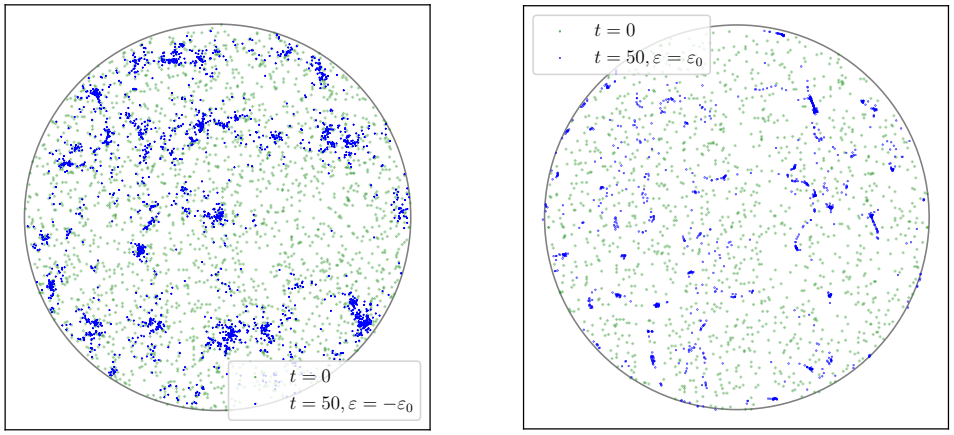


Figure 4.5: The green points represent a sample from a PPP $\mathcal{P}$ of unit intensity, the blue points correspond to $\Pi_{-\varepsilon_0,50}\mathcal{P}$ (left), and $\Pi_{+\varepsilon_0,50}\mathcal{P}$ (right). Plots obtained using the toolbox MCRPPy.

that some points of $\Pi_{-\varepsilon_0,50}$ are situated outside the observation window and are thus not visible in Figure 4.5. Indeed, following the arguments of Chatterjee et al. (2010), one can prove that the differential equation (4.11) 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 $t \to \tau_{\mathbf{x}_0}$. 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 (4.12), where we employed a fixed stepsize ε, 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 and give a new way to generate a clustered point process.

The scenario where $\varepsilon > 0$ and M is large is less straightforward. This case can be associated with the reverse dynamics of Equation (4.11). 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 a basin of the gravitational allocation from Lebesgue to $\mathcal{P}$; see the right panel of Figure B.1. By the same analogy as before, we can expect that for $\mathbf{x}_0 \in \mathcal{P}$ and ε small enough,

as $t \to \infty$, each point $\mathbf{x}_t$ will eventually approach the boundary of a basin 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 4.5, 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 C.1.

4.3.6 Monte Carlo with the repelled Poisson point process

This section focuses on examining and comparing the performance of $\widehat{I}_{\Pi_{\varepsilon_0}\mathcal{P}}(f)$ in (4.4), for $\varepsilon = \varepsilon_0$ in (4.9), with the Monte Carlo methods outlined in Section 2.7.2.

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 BPP using Algorithm 2, 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 $\widehat{I}_{\text{MCCV}}$ by *not* accounting for the evaluations of the integrand necessary to estimate the coefficient (2.28).

We use $M = 100$ samples from each of the point processes with N ranging roughly from 50 to 1000. We examine the functions f_1, f_2, and f_3 defined in Section 4.3 by Equation (4.10). Their integrals are

$$I(f_2) = \kappa_d/2^d, \qquad I(f_3) = 0,$$

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

Figure 4.6 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 squares linear regressions (OLSs). The slopes and standard deviations of the slopes are indicated in the legend.

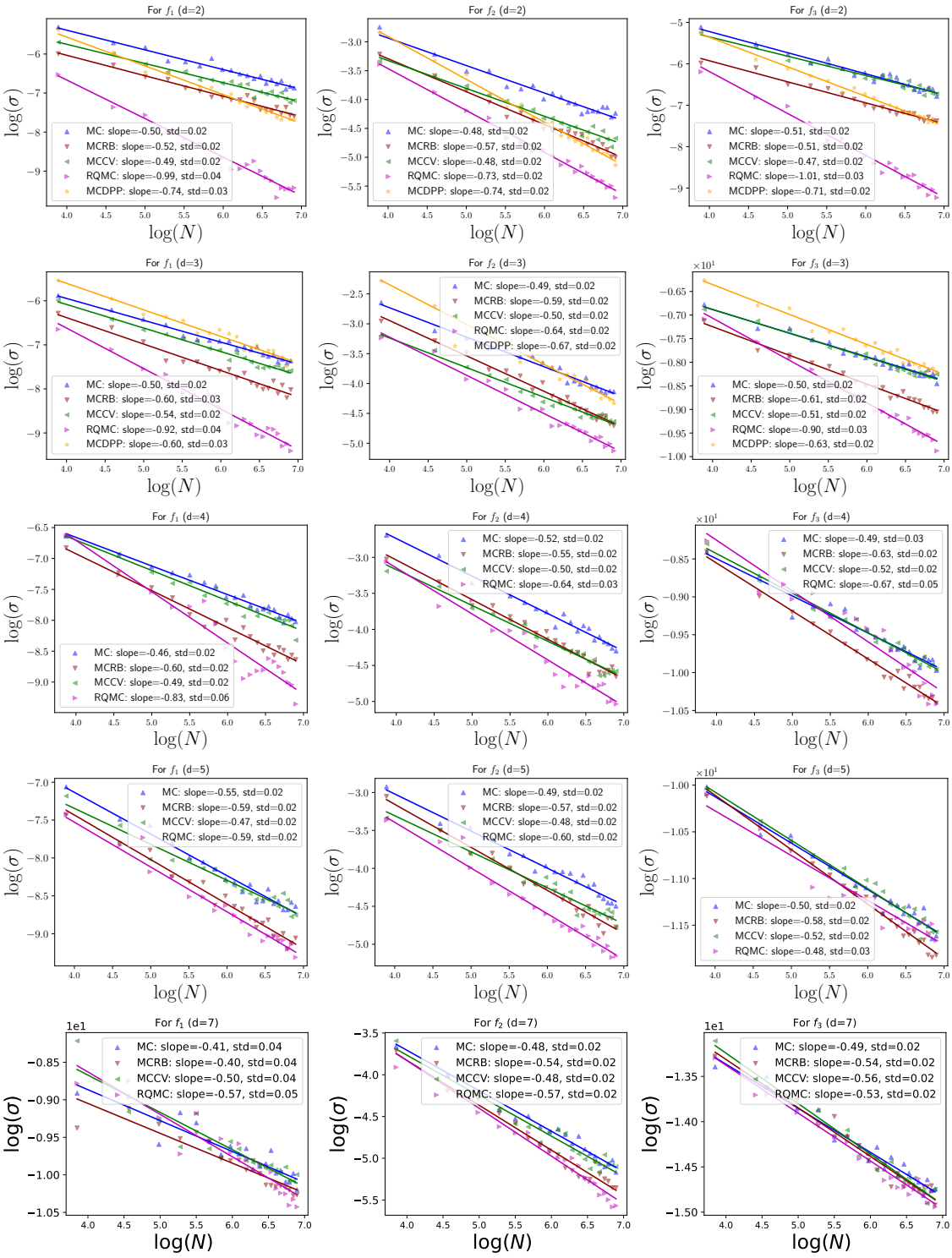


Figure 4.6: Estimated standard deviation σ of various Monte Carlo methods for f_1, f_2, and f_3 across different dimensions $d \in \{2, 3, 4, 5, 7\}$. Plots obtained using the toolbox MCRPPy.

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_3 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 variance of $\widehat{I}_{\text{MCDPP}}$ is $O(N^{-1-1/d})$, it seems that the hidden constant increases significantly with the dimension compared to the other methods. 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 $\widehat{I}_{\text{MCRP}}$. 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. 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_3 in $d = 5$. Finally, 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.

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

We use the Shapiro-Wilk test (Shapiro and Wilk, 1965) to assess whether the residuals of each OLS of $\log(\sigma)$ of $\widehat{I}_{\text{MCRB}}$ over $\log(N)$ shown in Figure 4.6 are normally distributed[7]. We also computed confidence intervals of the slopes of the corresponding OLS. Table 4.1 summarizes the results. The null hypothesis of Shapiro-Wilk's test is that the residuals are normally distributed. As the statistics

[7]To determine the significance of the variance of the slope in a homoscedastic OLS, it is necessary to ensure that the residuals are normally distributed. The normality of the residual allows getting a confidence interval for the slope using the estimated standard deviations via the Central Limit Theorem (Fox, 2015, Chapters 6 and 12).

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 use the estimated standard deviations of the slopes to obtain confidence intervals for these slope values. Table 4.1 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 $\widehat{I}_{\text{MCRB}}$ may decrease slightly faster than the N^{-1} convergence rate of the variance of crude Monte Carlo.

Finally, to account for the slight possible bias of $\widehat{I}_{\text{MCRB}}$, we present the errors – difference between $I(f)$ and its estimated value – obtained in the experiment illustrated by Figure 4.6. Figure 4.7 displays the box plots of the error of $\widehat{I}_{\text{MC}}$, $\widehat{I}_{\text{MCRB}}$, $\widehat{I}_{\text{MCCV}}$, and $\widehat{I}_{\text{RQMC}}$, for the functions f_2 and f_3 for d in $\{2,3,4,5,7\}$ and for $\widehat{I}_{\text{MCDPP}}$ when d in $\{2,3\}$. There is no clear evidence that $\widehat{I}_{\text{MCRB}}$ exhibits any notable bias.

4.4 Other repelled point processes

In this section, we perform numerical investigations to determine whether the variance reduction identified in Theorem 4.5 remains valid when the initial point process is more regular than a PPP like the Ginibre ensemble (GPP) (Example 2.2) and the scrambled Sobol sequence (SSS) (Section 2.7.2).

Recall that the GPP exhibits strong regularity being a Class I HUPP; see Section 3.1.2. Figure 4.8 displays a sample of a GPP observed in $B(\mathbf{0},50)$ and the corresponding repelled sample with $\varepsilon = \varepsilon_0$. The sampling methodology follows the description provided in Section 4.3.2.

Another model of interest is the SSS, which is a typical example of a randomized low-discrepancy sequence; see Section 5.1. Randomized low-discrepancy sequences play a crucial role in numerical integration being the backbone of the powerful Randomized Quasi-Monte Carlo method; see Section 2.7.2. Figure 4.9 displays a sample of the SSS observed in $B(\mathbf{0},50)$ and the corresponding repelled

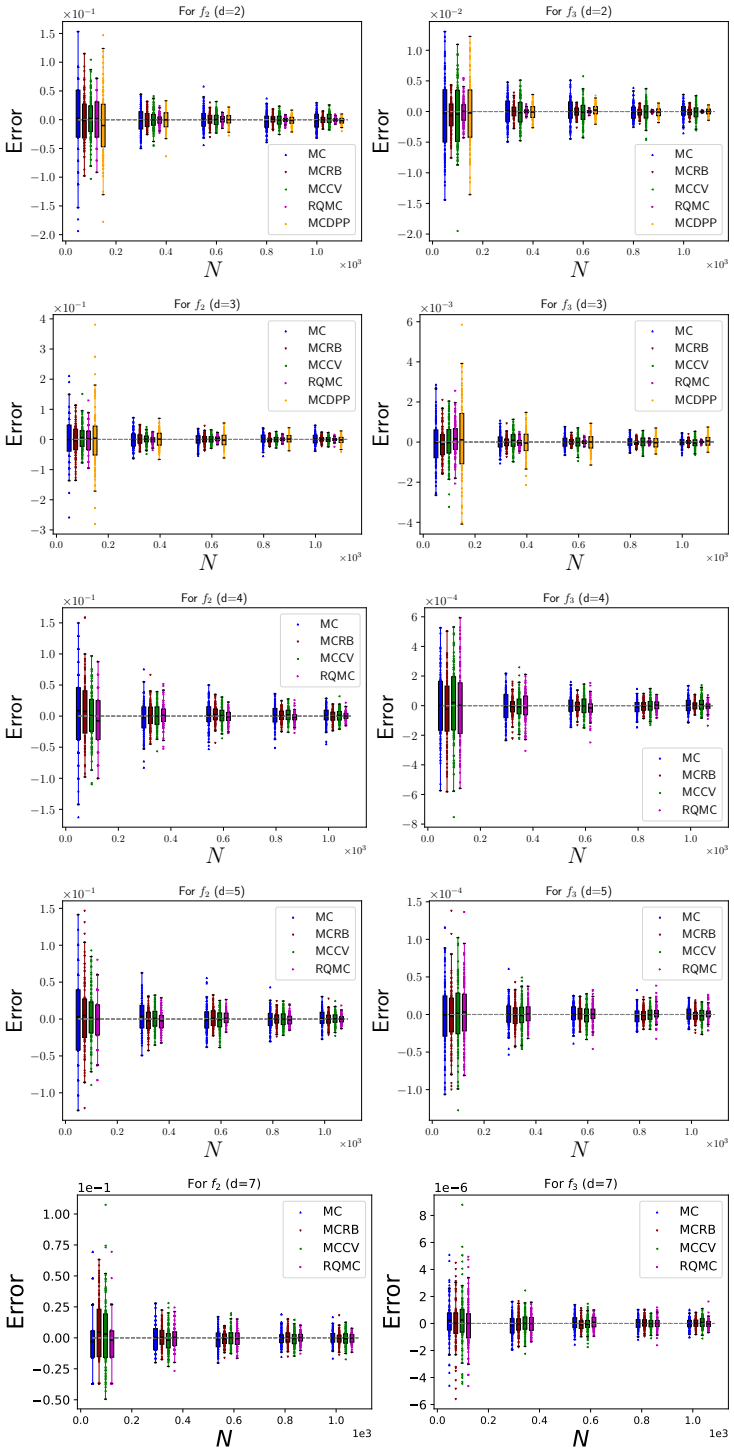
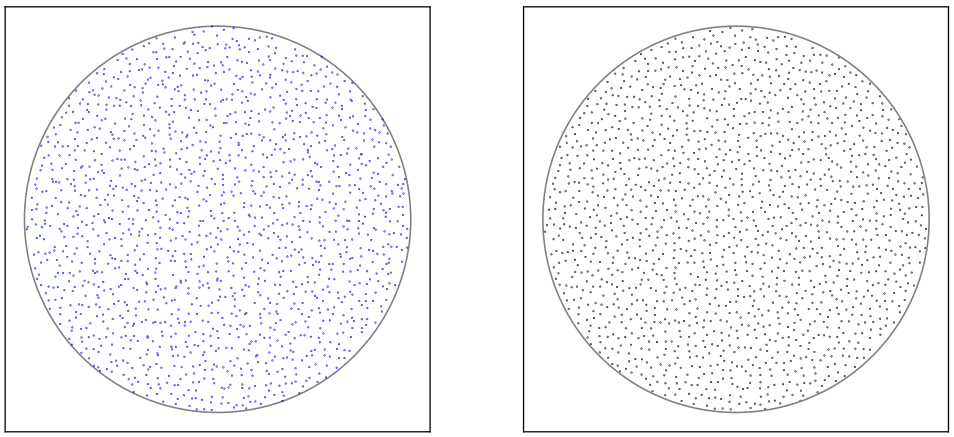


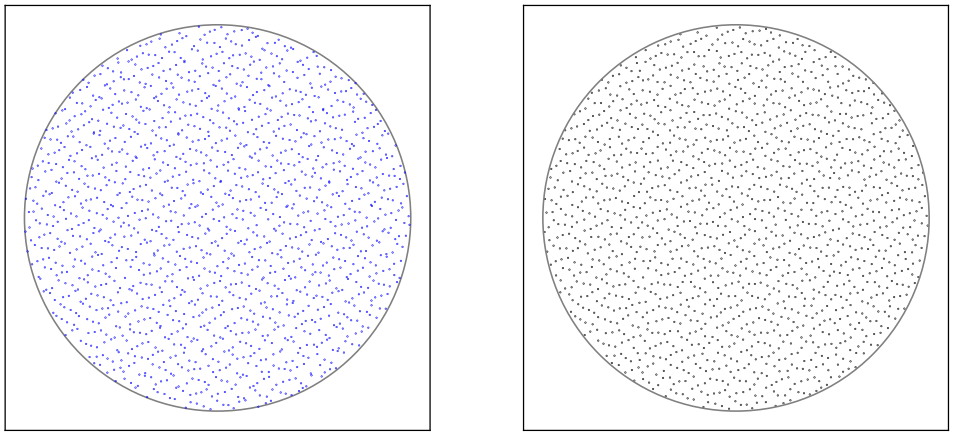
Figure 4.7: Estimated error of various Monte Carlo methods for f_2, and f_3 and $d \in \{2, 3, 4, 5, 7\}$.

sample, with $\varepsilon = \varepsilon_0$. Like for GPP, the repelled SSS displays a high level of regularity.



(a) GPP (b) Repelled GPP

Figure 4.8: A sample from the GPP (left) and the corresponding repelled sample (right) with $\varepsilon = \varepsilon_0$. Plots obtained using the toolbox MCRPPy.



(a) SSS (b) Repelled SSS

Figure 4.9: A sample from the SSS (left) and the obtained repelled sample (right) with $\varepsilon = \varepsilon_0$. Plots obtained using the toolbox MCRPPy.

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. ε, where $\mathcal{G}$ represents a GPP and $\mathcal{S}$ denotes a SSS. We use the functions f_1, f_2 and f_3 defined in Equation (4.10). Figure 4.10 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 ε. We conducted the experiments using 50 independent samples of $\mathcal{G}$ and $\mathcal{S}$ of intensity $\rho = 500$ obtained using Algorithm 2. GPP's samples were rescaled to achieve $\rho = 500$ as the original intensity of the GPP is equal to $1/\pi$; see Example 2.2. 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, 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 ε_0 defined in Equation (4.9). For the C^2 functions f_1 and f_3, we observe a behavior similar to the RPPP experiments depicted in Figure 4.3, indicating a variance reduction within a certain range of positive values of ε. However, for f_2, the variance decreases as ε increases for $\Pi_\varepsilon\mathcal{G}$ in a manner similar to Figure 4.3, 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.

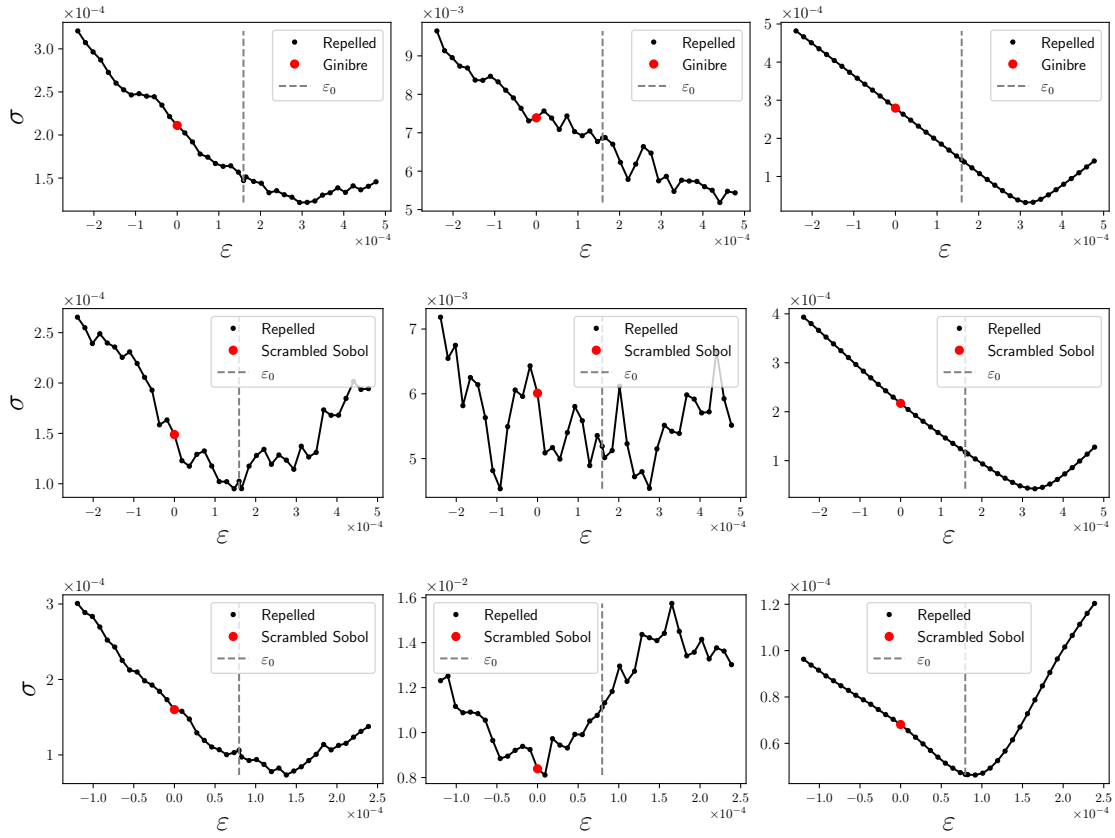


Figure 4.10: Estimated standard deviations of $\widehat{I}_{s,\Pi_\varepsilon\mathcal{G}\cap K}$ (Ginibre) and $\widehat{I}_{s,\Pi_\varepsilon\mathcal{S}\cap K}$ (scrambled Sobol) with respect to ε for f_1 (first column), f_2 (second column), and f_3 (last column). The first row shows the obtained results for the GPP in $d = 2$, the second row for the SSS in $d = 2$, and the last one for the SSS in $d = 3$. Plots obtained using the toolbox MCRPPy.

4.5 Proofs

In this section, we present the proofs of the results mentioned in Sections 4.1, 4.2, and 4.3.1. The proof of Proposition 4.3 can be found in Section 4.5.1. The motion invariance results stated by Proposition 4.1 and Corollary 4.1 are proven in Section 4.5.2. The most intricate proof, which is for Proposition 4.4, is presented in Section 4.5.3. Finally, Section 4.5.4 contains the proof of the variance reduction result formulated in Theorem 4.5.

4.5.1 Proposition 4.3

In this section, our objective is to prove Proposition 4.3. This proposition will play a crucial role in establishing the first part of Corollary 4.1.

Proof of Proposition 4.3. 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 $\Omega_1, \Omega_2, \ldots$ 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 Ω_n occur, and for $\mathbf{c} \in \mathbb{R}^d$,

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

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,

$$\mathbb{P}\left(F_{\mathcal{P}}(\mathbf{x}) - F_{\mathcal{P}}(\mathbf{y}) = \mathbf{c}\right) \leq \sum_{n\geq 1} \mathbb{P}\left(\mathbf{X}_n = \mathbf{c} - \mathbf{Y}_n, \Omega_n\right). \tag{4.13}$$

Now, for all n, conditionally on Ω_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 $\mathbb{P}\left(\mathbf{X}_n = \mathbf{c} - \mathbf{Y}_n, \Omega_n\right) = 0$. By (4.13), we conclude that

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

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

$$g_i \colon B_n \setminus \{\mathbf{x}, \mathbf{y}\} \to \mathbb{R}$$
$$\mathbf{z} \mapsto \frac{x_i - z_i}{\|\mathbf{x} - \mathbf{z}\|_2^d} - \frac{y_i - z_i}{\|\mathbf{y} - \mathbf{z}\|_2^d}.$$

Actually, for $\mathbf{z} \triangleq \mathcal{P} \cap B_n$, $g_i(\mathbf{z})$ 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 Ω_n, $\mathbf{z}$ is uniformly distributed on B_n, so

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

□

4.5.2 Proposition 4.1 and Corollary 4.1 (motion-invariance)

Throughout this Section, we fix $\varepsilon \in \mathbb{R}$ and we prove Proposition 4.1 and Corollary 4.1.

First, to prove Proposition 4.1, we show that for a point process $\mathcal{X}$ satisfying the conditions of Proposition 4.1, the void probability (Equation 2.1) of a translation, respectively rotation of $\Pi_\varepsilon \mathcal{X}$ is equal two the void probability of $\Pi_\varepsilon \mathcal{X}$ for any Borel set. The main point of contention is that $F_{\mathcal{X}}(\mathbf{x})$ (Equation F_1) is invariant under both translations and rotations, for any $\mathbf{x} \in \mathbb{R}^d$.

Proof of Proposition 4.1. Suppose that $\mathcal{X}$ is motion-invariant. As the law of $\Pi_\varepsilon \mathcal{X}$ is defined by the system of void probabilities (2.1), 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 $\mathfrak{r}$,

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

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

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

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

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

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

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

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

which is equal to $\mathbb{T}_{\Pi_\varepsilon(\mathcal{X})}(B)$ 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 Π_ε with its truncated version $\Pi_\varepsilon^{(q,p)}$ (4.3), with $0 \le q < p$.

Now, to prove Corollary 4.1, we first note that a PPP $\mathcal{P}$ is almost surely a valid configuration, as mentioned in Section 4.2. In view of Proposition 4.1, it is enough to prove that, almost surely, the images of two distinct points from $\mathcal{P}$ under Π_ε remain distinct and that $\Pi_\varepsilon\mathcal{P}$ possesses the same intensity as $\mathcal{P}$.

Proof of Corollary 4.1. To show that almost surely the images of two distinct points from $\mathcal{P}$ under Π_ε remain distinct,[8] we need to show that

$$\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x}+\varepsilon F_{\mathcal{P}}(\mathbf{x})=\mathbf{y}+\varepsilon F_{\mathcal{P}}(\mathbf{y})\}}(\mathbf{x},\mathbf{y})\right] = 0. \tag{4.14}$$

Using the extended Slivnyak-Mecke theorem (2.12) we get

$$\begin{aligned}
&\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x}+\varepsilon F_{\mathcal{P}}(\mathbf{x})=\mathbf{y}+\varepsilon F_{\mathcal{P}}(\mathbf{y})\}}(\mathbf{x},\mathbf{y})\right] \\
&= \mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}}^{\neq} \mathbb{1}_{\{\mathbf{x}+\varepsilon F_{\mathcal{P}\setminus\{\mathbf{x},\mathbf{y}\}}(\mathbf{x})+\varepsilon\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2^d}=\mathbf{y}+\varepsilon F_{\mathcal{P}\setminus\{\mathbf{x},\mathbf{y}\}}(\mathbf{y})+\varepsilon\frac{\mathbf{y}-\mathbf{x}}{\|\mathbf{y}-\mathbf{x}\|_2^d}\}}(\mathbf{x},\mathbf{y})\right] \\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathbb{E}\left[\mathbb{1}_{\{\mathbf{x}+\varepsilon F_{\mathcal{P}}(\mathbf{x})+\varepsilon\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2^d}=\mathbf{y}+\varepsilon F_{\mathcal{P}}(\mathbf{y})+\varepsilon\frac{\mathbf{y}-\mathbf{x}}{\|\mathbf{y}-\mathbf{x}\|_2^d}\}}(\mathbf{x},\mathbf{y})\right]\rho^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathbb{P}\left(F_{\mathcal{P}}(\mathbf{x})-F_{\mathcal{P}}(\mathbf{y}) = (\mathbf{y}-\mathbf{x})\left(\varepsilon^{-1}+2\|\mathbf{y}-\mathbf{x}\|_2^{-d}\right)\right)\rho^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}.
\end{aligned} \tag{4.15}$$

By Proposition 4.3, 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 (4.15) yields (4.14).

It remains to show that the intensity of $\Pi_\varepsilon\mathcal{P}$ is equal to ρ. Consider a compact K of $\mathbb{R}^d$. Based on the previous reasoning, almost surely, when the repulsion operator Π_ε 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 (2.12) we get

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

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

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

[8] 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 Proposition 6.7 from (Last and Penrose, 2017).

Exchanging the integral and the expectation using Tonelli's theorem gives

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

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

4.5.3 Proposition 4.4 (existence of the moments)

In this section, let $\mathcal{P} \subset \mathbb{R}^d$ be a PPP of intensity ρ, with $d \geq 3$ and $\varepsilon \in (-1, 1)$. Our objective is to demonstrate the existence of the moments of $\Pi_\varepsilon \mathcal{P}$. 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_2) 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

$$F_{\mathcal{P}}(\mathbf{x}) = F_{\mathcal{P}}^{(0,1)}(\mathbf{x}) + F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x}), \tag{4.16}$$

where the truncated forces are defined in (4.2). We refer to the first term in the right-hand side of (4.16) 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

$$\{\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) \in B(\mathbf{0}, R)\} \subset \left\{\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}) \in B(\mathbf{0}, r(\mathbf{x}))\right\} \cup \left\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 \geq \frac{r(\mathbf{x}) - R}{|\varepsilon|}\right\}. \tag{4.17}$$

To see the validity of this inclusion, note that if $\mathbf{x}$ is not in the right-hand side of (4.17), 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 (4.17), we get

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

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

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\Pi_\varepsilon\mathcal{P}} \mathbb{1}_{B(\mathbf{0},R)}(\mathbf{x})\right)^m\right] \leq 3^{m-1}\Bigg(&\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}} \mathbb{1}_{B(\mathbf{0},R')}(\mathbf{x})\right)^m\right] \\ &+ \mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')^c} \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}))}(\mathbf{x}+\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}))\right)^m\right] \\ &+ \mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x})-R)/|\varepsilon|\}}(\mathbf{x})\right)^m\right]\Bigg).\end{aligned} \tag{4.18}$$

The first term on the right-hand side of (4.18) 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 4.2 and 4.3.

We first focus on the term in (4.18) involving the external force.

Lemma 4.1. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, 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*

$$\mathbb{P}(\|F_{\mathcal{P}}^{(q,p)}(\mathbf{0})\|_2 > t) \leq c_1 \exp\left(-c_2 q^{d-1} t \log\left(\frac{c_3 t}{q}\right)\right). \tag{4.19}$$

Note that by translation-invariance, the choice of $\mathbf{0}$ in (4.19) is arbitrary. In addition, when $\rho = 1$, Equation (4.19) 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, following Chatterjee et al. (2010),

we provide here a proof of Lemma 4.1, 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 4.1. 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

$$\mathbb{P}\left(G_i^{(q,p)}(\mathbf{0}) > t\right) \leq C_1 \exp\left(-C_2 q^{d-1} t \log\left(\frac{C_3 t}{q}\right)\right), \tag{4.20}$$

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

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

and combining this with (4.20) we obtain (4.19), with $c_1 = 2dC_1$, $c_2 = C_2/d$, and $c_3 = C_3/d$. Thus, we only need to verify Equation (4.20).

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

$$\begin{aligned}\mathbb{P}\left(G_i^{(q,p)}(\mathbf{0}) > t\right) &= \mathbb{P}\left(\exp\left(\theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d}\right) > \exp(\theta t)\right) \\ &\leq \mathbb{E}\left[\exp\left(\theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2^d}\right)\right] \exp(-\theta t). \end{aligned} \tag{4.21}$$

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

$$\mathbb{E}\left[\exp\left(\theta \sum_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \frac{z_i}{\|\mathbf{z}\|_2}\right) \;\middle|\; \mathcal{P}(A^{(q,p)}) = N\right] =$$

$$\mathbb{E}\left[\prod_{\substack{\mathbf{z} \in A^{(q,p)} \cap \mathcal{P} \\ \|\mathbf{z}\|_2 \uparrow}} \exp\left(\theta \frac{z_i}{\|\mathbf{z}\|_2^d}\right) \;\middle|\; \mathcal{P}(A^{(q,p)}) = N\right] =$$

$$\mathbb{E}\left[\exp\left(\theta \frac{u_i}{\|\mathbf{u}\|_2^d}\right)\right]^N = \left(1 + \mathbb{E}\left[\sum_{k \geq 2} \frac{1}{k!} \theta^k \frac{u_i^k}{\|\mathbf{u}\|_2^{kd}}\right]\right)^N.$$

In particular,

$$\mathbb{E}\left[\exp\left(\theta \sum_{\substack{\mathbf{z}\in A^{(q,p)}\cap\mathcal{P}\\ \|\mathbf{z}\|_2\uparrow}} \frac{z_i}{\|\mathbf{z}\|_2}\right) \;\middle|\; \mathcal{P}(A^{(q,p)}) = N\right] \leq \left(1 + \sum_{k\geq 2}\frac{1}{k!}\theta^k \mathbb{E}\left[\|\mathbf{u}\|_2^{-k(d-1)}\right]\right)^N. \tag{4.22}$$

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

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

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 (4.21), we obtain

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

Now, remembering that $\mathcal{P}(A^{(q,p)})$ is a Poisson random variable of parameter $\rho|A^{(q,p)}|$, we obtain[9]

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

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

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

[9] If X is a Poisson random variable of parameter λ, then for any γ the mean of the random variable $(1+\gamma)^X$ is equal to $\exp(\lambda\gamma)$.

which ends the proof. □

Lemma 4.1 has the following corollary.

Corollary 4.2. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, 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*

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')^c} \mathbb{1}_{\left\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x})-R}{|\varepsilon|}\right\}}(\mathbf{x})\right)^m\right] \leq \sum_{k=1}^m a_k \left(\int_{B(\mathbf{0},R')^c} \exp\left(-b_k g_k(\mathbf{x}) \log(c_k g_k(\mathbf{x}))\right) \rho \mathrm{d}\mathbf{x}\right)^k, \tag{4.23}$$

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 (4.16). Indeed, as $\|\mathbf{x}\|_2 \to \infty$ we have

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

Thus, for any positive integer m

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x})-R}{|\varepsilon|}\}}(\mathbf{x})\right)^m\right] < \infty. \tag{4.24}$$

Proof of Corollary 4.2. Fix a positive integer m. There exists m constants $(d_i)_{i=1}^m$ such that

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')^c} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 > \frac{r(\mathbf{x})-R}{|\varepsilon|}\}}(\mathbf{x})\right)^m\right] = \sum_{k=1}^m d_k \underbrace{\mathbb{E}\left[\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_k\in\mathcal{P}\cap B(\mathbf{0},R')^c}^{\neq} \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x}_1)\|_2 > \frac{r(\mathbf{x}_1)-R}{|\varepsilon|}\}}(\mathbf{x}_1) \cdots \mathbb{1}_{\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x}_k)\|_2 > \frac{r(\mathbf{x}_k)-R}{|\varepsilon|}\}}(\mathbf{x}_k)\right]}_{\triangleq E_k}. \tag{4.25}$$

Using Lemma 4.1 we will show that for any $k \geq 1$ there exists positive constants a_k, b_k and c_k such that

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

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

$$\mathbb{1}_{\left\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x})-R)/|\varepsilon|\right\}}(\mathbf{x}) \leq \mathbb{1}_{\left\{\|F_{\mathcal{P}\setminus\{\mathbf{x},\mathbf{y}\}}^{(1,\infty)}(\mathbf{x})\|_2 \geq 2g_2(\mathbf{x})\right\}}(\mathbf{x}). \tag{4.26}$$

Generalizing Equation (4.26) to k points gives

$$\mathbb{1}_{\left\{\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{x})\|_2 \geq (r(\mathbf{x})-R)/|\varepsilon|\right\}}(\mathbf{x}) \leq \mathbb{1}_{\left\{\|F_{\hat{\mathcal{P}}(k)}^{(1,\infty)}(\mathbf{x})\|_2 \geq kg_k(\mathbf{x})\right\}}(\mathbf{x}). \tag{4.27}$$

Using Equation (4.27) and the extended Slivnyak-Mecke theorem (2.12) we get

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

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

$$\begin{aligned} E_k &\leq \int_{(B(\mathbf{0},R')^c)^k} \min_{j\in\{1,\dots,k\}} \mathbb{P}\left(\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq kg_k(\mathbf{x}_j)\right)\rho^k \mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_k \\ &= \int_{(B(\mathbf{0},R')^c)^k} \mathbb{P}\left(\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq \max_{j\in\{1,\dots,k\}} kg_k(\mathbf{x}_j)\right)\rho^k \mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_k \\ &\leq \int_{(B(\mathbf{0},R')^c)^k} \mathbb{P}\left(\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 \geq \sum_{j=1}^{k} g_k(\mathbf{x}_j)\right)\rho^k \mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_k. \end{aligned} \tag{4.28}$$

For $\mathbf{x_1}, \dots, \mathbf{x}_k \in B(\mathbf{0}, R')^c$, Lemma 4.1 with $q = 1$ and $p = \infty$ guarantees the existence of C_1, C_2, $C_3 > 0$ such that

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

Plugging back into (4.28), and then in (4.25), 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

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

which concludes the proof. □

Now we switch focus to bounding the contribution from the internal force to (4.16). Again, we work with a lemma and a corollary.

Lemma 4.2. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, with $d \geq 3$. Let $R > 0$ and $\varepsilon \in (-1, 1)$. Consider a function $r : \mathbb{R}^d \to \mathbb{R}^+ \setminus \{0\}$ 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*

$$\mathbb{P}\left(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{0}, r(\mathbf{x})) \right) \leq \frac{c_1}{h(\mathbf{x}) g(\mathbf{x})^{1+\gamma}} + c_3 \exp\left(-c_4 g(\mathbf{x}) \log(c_5 g(\mathbf{x})) \right), \tag{4.29}$$

where $h(\mathbf{x}) = \max\left(1, c_2 \left(\|\mathbf{x}\|_2 / r(\mathbf{x}) - 1 \right)^{d-1} - 1\right)$, 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 (4.29) is arbitrary. Second, by choosing $r(\mathbf{x}) = \|\mathbf{x}\|_2^{\beta}$, with $0 < \beta < 1/(2(d-1))$,

the upper bound in Equation (4.29) is $o(\|\mathbf{x}\|_2^{-d})$ as $\|\mathbf{x}\|_2$ goes to infinity. Thus, $\mathbb{P}\left(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B\left(\mathbf{0}, r(\mathbf{x})\right)\right)$ is integrable over $B(\mathbf{0}, R)^c$.

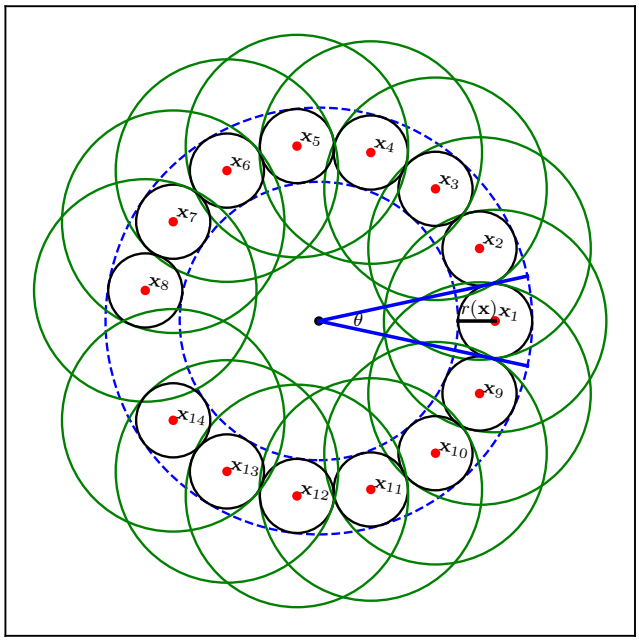


Figure 4.11: Illustration of the proof idea

Proof of Lemma 4.2. 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 $\mathbf{y}$ on $S(\mathbf{0}, \|\mathbf{x}\|_2)$. For any $i \in \{1, \ldots, m\}$

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

Thus, we have $\mathbf{z} \in S(\mathbf{0}, \|\mathbf{x}\|_2)$ and $B(\mathbf{z}, r(\mathbf{x})) \in A^{(q,p)} \setminus \bigcup_{i=1}^m B(\mathbf{x}_i, r(\mathbf{x}))$ 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 4.11 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 $i \in \{1, ..., m\}$, $\mathbf{x}_i$ can be chosen inductively as the rotation of $\mathbf{x}_{i-1}$ of angle $\theta = \mathrm{Arcsin}(r(\mathbf{x})/\|\mathbf{x}\|_2)$ 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

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

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

$$\mathbb{P}\left(\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in A^{(q,p)}\right) \geq m\mathbb{P}\left(\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_1, r(\mathbf{x}))\right). \tag{4.30}$$

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

$$\begin{aligned}h\colon [\|\mathbf{x}\|_2 - r(\mathbf{x}), \|\mathbf{x}\|_2 + r(\mathbf{x})] &\to \mathbb{R} \\ x &\mapsto x^d\end{aligned}$$

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

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

where $C_2 = (2d)/3^d$. Thus Equation (4.30) 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

$$\begin{aligned}\mathbb{P}\left(\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{0}) \in B(\mathbf{x}_1, r(\mathbf{x}))\right) &\leq \frac{\mathbb{P}\left(|\varepsilon|\|F_{\mathcal{P}}(\mathbf{0}) - F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 > \|\mathbf{x}\|_2 - r(\mathbf{x})\right)}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right)} \\ &\leq \frac{\mathbb{P}\left(\|F_{\mathcal{P}}(\mathbf{0})\|_2 > \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|}\right) + \mathbb{P}\left(\|F_{\mathcal{P}}^{(1,\infty)}(\mathbf{0})\|_2 > \frac{\|\mathbf{x}\|_2 - r(\mathbf{x})}{2|\varepsilon|}\right)}{\max\left(1, C_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right)}\end{aligned} \tag{4.31}$$

As a result of Proposition 4.2, for any $0 < \gamma < 1/(d-1)$ we have $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^{1+\gamma}]$ is finite. Applying Markov's inequality to the first part of Equation (4.31), and using Lemma 4.1 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

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

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 $C_6 = c_3/2$. □

Note that under the assumptions and definitions of Lemmas 4.1, 4.2, for $\mathbf{x} \in B(\mathbf{0}, R^{1/\beta})^c$, there exists positive constants $c_1, \ldots, c_5$ such that

$$\begin{aligned}\mathbb{P}(\mathbf{x} + \varepsilon F_{\mathcal{P}}(\mathbf{x}) \in B(\mathbf{0}, R)) \leq &\frac{c_1}{\max\left(1, c_2\left(\frac{\|\mathbf{x}\|_2}{r(\mathbf{x})} - 1\right)^{d-1} - 1\right) g(\mathbf{x})^{1+\gamma}} \\ &+ c_3 \exp\left(-c_4 g(\mathbf{x}) \log(c_5 g(\mathbf{x}))\right),\end{aligned} \tag{4.32}$$

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

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

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 (4.32) 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 4.2.

Corollary 4.3. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, 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*

$$E_k \triangleq \mathbb{E}\left[\sum_{\mathbf{x}_1,\dots,\mathbf{x}_k\in\mathcal{P}\cap B(\mathbf{0},R)^c}^{\neq} \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}_1))}(\mathbf{x}_1 + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1)) \dots \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}_k))}(\mathbf{x}_k + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_k))\right]. \tag{4.33}$$

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

$$E_1 \leq \int_{B(\mathbf{0},R)^c} \frac{c_1}{\max\left(1, c_2\left(\|\mathbf{x}\|_2^{1-\beta} - 1\right)^{d-1} - 1\right) g(\mathbf{x})^{1+\gamma}} + c_3 \exp\left(-c_4 g(\mathbf{x}) \log\left(c_5 g(\mathbf{x})\right)\right) \rho \mathrm{d}\mathbf{x}, \tag{4.34}$$

and

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)^c} \mathbb{1}_{B(\mathbf{0},f(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}))\right)^m\right] \leq \sum_{k=1}^{m-1}\left(a_k E_k + b^{k-1} E_1 E_{m-k}\right) + b^{m-1} E_1. \tag{4.35}$$

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

Taking $0 < \beta < \frac{\gamma}{d-1}$, the integrand in Equation (4.34) is $o(\|\mathbf{x}\|_2^{-d})$ as $\|\mathbf{x}\|_2 \to \infty$ implying that

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)^c} \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}))}(\mathbf{x} + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}))\right)^m\right] < \infty, \tag{4.36}$$

for any positive integer m.

Proof of Corollary 4.3. We will show the validity of Equation (4.35) by induction on $m \geq 1$.

To begin with, for $m = 1$ using the extended Slivnyak-Mecke theorem (2.12) and that the law of $F_{\mathcal{P}}^{(0,1)}(\mathbf{x})$ is translation-invariant (w.r.t. $\mathbf{x}$) we have

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

By Lemma 4.2, for any $0 < \gamma < 1/(d-1)$ there exists positive constants $c_1, \ldots, c_5$ such that

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

Next, suppose that Equation (4.35) 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

$$\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)^c} \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}))}(\mathbf{x}+\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}))\right)^{m+1}\right] = \sum_{k=1}^{m} d_k E_k + E_{m+1}. \tag{4.37}$$

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

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

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

$$+$$

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

First, for A_{m+1} the extended Slivnyak-Mecke theorem (2.12) followed by a change of variables gives

$$A_{m+1} \leq \mathbb{E}\left[\sum_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_{m+1} \\ \max_{1<i} \|\mathbf{x}_i + \mathbf{x}_1\| \leq 2}}^{\neq} \mathbb{1}_{B(\mathbf{0}, r(\mathbf{x}_1))}\left(\mathbf{x}_1 + \varepsilon F_{\hat{\mathcal{P}}^{(m+1)}}^{(0,1)}(\mathbf{x}_1) + \varepsilon \sum_{j=2}^{m+1} \frac{\mathbf{x}_1 - \mathbf{x}_j}{\|\mathbf{x}_1 - \mathbf{x}_j\|_2^d}\right)\right] =$$

$$\int_{B(\mathbf{0},R)^c} \int_{B(\mathbf{x}_1,2)^m} \mathbb{P}\left(\mathbf{x}_1 + \varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}_1) \in B\left(\varepsilon \sum_{j=2}^{m+1} \frac{\mathbf{x}_j - \mathbf{x}_1}{\|\mathbf{x}_j - \mathbf{x}_1\|_2^d}, r(\mathbf{x}_1)\right)\right) \rho^{m+1} \mathrm{d}\mathbf{x}_{m+1} \ldots \mathrm{d}\mathbf{x}_1 =$$

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

By the definition of the first intensity measure (2.2) the last equation is equal to

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

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

$$A_{m+1} \leq \int_{B(\mathbf{0},2)^m} \mathbb{E}\left[\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)^c} \mathbb{1}_{B(\mathbf{0},r(\mathbf{x}))}(\mathbf{x}+\varepsilon F_{\mathcal{P}}^{(0,1)}(\mathbf{x}))\right]\rho^m \mathrm{d}\mathbf{t}_1\ldots\mathrm{d}\mathbf{t}_m$$
$$= (2^d\kappa_d\rho)^m E_1. \tag{4.38}$$

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

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

Applying the extended Slivnyak-Mecke theorem (2.12) then, using the independence of $F_{\mathcal{P}}^{(0,1)}(\mathbf{x})$ and $F_{\mathcal{P}}^{(0,1)}(\mathbf{y})$ whenever $\|\mathbf{x}-\mathbf{y}\|_2>2$ yield

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

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

$$A_k \leq$$

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

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

$$(2^d \kappa_d \rho)^{k-1} E_1 E_{m+1-k}. \tag{4.39}$$

Inserting (4.38), and (4.39) in (4.37) we get

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

$$\leq c_1^m E_1 + \sum_{k=1}^{m} \left(d_k E_k + c_1^{k-1} E_1 E_{m+1-k} \right),$$

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

4.5.4 Theorem 4.5 (variance reduction)

In this section, we prove Theorem 4.5. 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 $\varepsilon \to 0$ 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 4.5.

Lemma 4.3. *Let $d > 2$ and $g \in C^2(\mathbb{R}^d)$ have compact support. For all $\mathbf{x} \in \mathbb{R}^d$, we have*

$$g(\mathbf{x}) = \frac{1}{(2-d)d\kappa_d} \int_{\mathbb{R}^d} \Delta g(\mathbf{y}) \frac{1}{\|\mathbf{x}-\mathbf{y}\|_2^{d-2}} \mathrm{d}\mathbf{y} = \frac{1}{d\kappa_d} \int_{\mathbb{R}^d} \nabla g(\mathbf{y}) \cdot \frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2^d} \mathrm{d}\mathbf{y}. \tag{4.40}$$

Proof of Lemma 4.3. First, recall Green's "integration by parts" formula (Axler et al., 2001, Chapter 1)

$$\int_A f(\mathbf{x}) \Delta g(\mathbf{x}) \mathrm{d}\mathbf{x} = \int_{\partial A} (f \nabla g.\mathbf{n} + g \nabla f.\mathbf{n}) \, \mathrm{d}S - \int_A \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{4.41}$$

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, and S is the surface-area measure on ∂A. Let the support of g be a compact $K \subset A$, (4.41) simplifies and we have

$$\int_A f(\mathbf{x})\Delta g(\mathbf{x})\mathrm{d}\mathbf{x} = \int_{\partial A} f\nabla g.\mathbf{n}\,\mathrm{d}S - \int_A \nabla f(\mathbf{x})\cdot\nabla g(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{4.42}$$

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

$$\frac{1}{(2-d)d\kappa_d}\int_{\mathbb{R}^d}\frac{\Delta g(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|_2^{d-2}}\mathrm{d}\mathbf{y} = \lim_{\varepsilon\to 0}\frac{1}{(2-d)d\kappa_d}\int_{B(\mathbf{0},R)\setminus B(\mathbf{x},\varepsilon)}\frac{\Delta g(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|_2^{d-2}}\mathrm{d}\mathbf{y}. \tag{4.43}$$

Now, using (4.42) 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

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

Plugging into (4.43) evaluating the limit, we obtain the desired limit. □

We use Lemma 4.3 to prove the following result, which takes us closer to controlling the variance of linear statistics under the RPPP.

Lemma 4.4. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, 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*

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathbb{E}\left[\left(\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R')} f(\mathbf{x}+\varepsilon F_{\mathcal{P}}(\mathbf{x}))\right)^2 - \left(\sum_{\mathbf{x}\in\mathcal{P}} f(\mathbf{x})\right)^2\right] = -2d\kappa_d\rho^2 I(f^2), \tag{4.44}$$

where $I(.)$ is defined in (2.22).

Proof of Lemma 4.4. 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

$$\lim_{\varepsilon \to 0} \mathbb{E}[X_\varepsilon] = -2d\kappa_d \rho^2 I(f^2).$$

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 $Y \in L^1$. To see why, note that the extended Slivnyak-Mecke theorem (2.12) yields

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

where the last inequality results from Hölder's inequality with $1 < a < d/(d-1)$ and $b = a/(a-1)$. As indicated in Section 4.2, $\mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^\gamma] < \infty$ 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

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

Expanding the sum, it comes

$$\lim_{\varepsilon\to 0}\mathbb{E}[X_\varepsilon] =$$
$$2\mathbb{E}\left[\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)} f(\mathbf{x})\nabla f(\mathbf{x})\cdot F_{\mathcal{P}}(\mathbf{x})+\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}\cap B(\mathbf{0},R)}^{\neq} f(\mathbf{x})\nabla f(\mathbf{y})\cdot F_{\mathcal{P}}(\mathbf{y})\right] =$$
$$2\mathbb{E}\left[\sum_{\mathbf{x}\in\mathcal{P}\cap B(\mathbf{0},R)} f(\mathbf{x})\nabla f(\mathbf{x})\cdot F_{\mathcal{P}\setminus\{\mathbf{x}\}}(\mathbf{x})\right]-2\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}\cap B(\mathbf{0},R)}^{\neq} f(\mathbf{x})\nabla f(\mathbf{y})\cdot\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2^d}\right]$$
$$+2\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in\mathcal{P}\cap B(\mathbf{0},R)}^{\neq} f(\mathbf{x})\nabla f(\mathbf{y})\cdot F_{\mathcal{P}\setminus\{\mathbf{y},\mathbf{x}\}}(\mathbf{y})\right]. \tag{4.45}$$

Using Slivnyak-Mecke (2.12) and Equation (2.11), we obtain

$$\begin{aligned}\lim_{\varepsilon\to 0}\mathbb{E}\left[X_\varepsilon\right] &= 2\int_{B(\mathbf{0},R)}\mathbb{E}\left[f(\mathbf{x})\nabla f(\mathbf{x})\cdot F_{\mathcal{P}}(\mathbf{x})\right]\rho\mathrm{d}\mathbf{x}\\ &\quad-2\int_{B(\mathbf{0},R)\times B(\mathbf{0},R)} f(\mathbf{x})\nabla f(\mathbf{y})\cdot\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|_2^d}\rho^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\ &\quad+2\int_{B(\mathbf{0},R)\times B(\mathbf{0},R)}\mathbb{E}\left[f(\mathbf{x})\nabla f(\mathbf{y})\cdot F_{\mathcal{P}}(\mathbf{y})\right]\rho^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}.\end{aligned} \tag{4.46}$$

But the first and last terms in (4.46) are zero due to $F_{\mathcal{P}}$ being a centered process, and the second term is $-2d\kappa_d\rho^2 I(f^2)$ by Lemma 4.3. □

Remarkably, the proof of Lemma 4.4 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 (4.45) remains valid.

Lemma 4.5. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, 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 $R' > (2R+2)^{1/\beta}$*

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

The proof is based on Corollary 4.2 and 4.3.

Proof of Lemma 4.5. 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 (4.17), and write

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

By Corollary 4.3 with $m = 2$, there exists $a_1 > 0$ such that for any $0 < \gamma < 1/(d-1)$

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

where E_1 is defined by Equation (4.33), for $k = 1$ and

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

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

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

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 ε approaches zero, so that $E_{\text{int}}(\varepsilon) = o(\varepsilon)$.

It remains to show that $E_{\text{ext}}(\varepsilon) = o(\varepsilon)$. By Corollary 4.2 with $m = 2$ there exists $a_5, \ldots, a_8 > 0$ such that

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

with

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

In particular

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

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 $Y_\varepsilon = o_{\varepsilon\to 0}(1)$, which ends the proof. □

Again, the proof's validity is unaffected by replacing $F_\mathcal{P}$ with its truncated counterpart $F_\mathcal{P}^{(0,p)}$, where $p > 0$. The upcoming lemma is the final tool required to demonstrate Theorem 4.5.

Lemma 4.6. *Consider a homogeneous Poisson point process $\mathcal{P} \subset \mathbb{R}^d$ of intensity ρ, with $d \geq 3$. Let f be a $C^2(\mathbb{R}^d)$ function of compact support $K \subset B(\mathbf{0}, R)$ with $R > 0$, and $\varepsilon \in (-1, 1)$. For any $R' \geq R$ we have*

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

Proof of Lemma 4.6. 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 $Z_\varepsilon = o_{\varepsilon\to 0}(1)$. Since f is C^2 with compact support, it is Lipschitz and we denote its Lipschitz constant by $L \geq 0$. For $\mathbf{x} \in B^c(0, R')$, $f(\mathbf{x}) = 0$ 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

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

The last equality was obtained using the extended Slivnyak-Mecke theorem (2.12). Next, we employ Hölder's inequality with $1 < p < \frac{d}{d-1}$ and $q = p/(p-1)$ to obtain

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

When attempting to compute the limit of the last equation as ε 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

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

Thus

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

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

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

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

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

and $\int_{\mathbb{R}^d} Y(\mathbf{u})\mathrm{d}\mathbf{u} = \kappa_d R'^d \mathbb{E}[\|F_{\mathcal{P}}(\mathbf{0})\|_2^p]$, which is finite since $p < d/(d-1)$; see Section 4.2. The dominated convergence theorem thus concludes the proof. □

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 4.5. We will use Lemmas 4.4, 4.5, and 4.6.

Proof of Theorem 4.5. Consider a function $f \in C^2(\mathbb{R}^d)$ of compact support $K \subseteq B(0,R)$ for some $R > 0$. We start by proving that

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

First, Proposition 4.4 implies the existence of $\mathbb{V}\mathrm{ar}\left[\widehat{I}_{\Pi_\varepsilon\mathcal{P}\cap K}(f)\right]$ 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

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

Using Lemmas 4.4, 4.5, and 4.6 we get

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

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

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

which ends the proof. □

Remark 4.8. Substituting $F_{\mathcal{P}}$ with its truncated version $F_{\mathcal{P}}^{(0,p)}$, where $p \geq \mathrm{diam}(K)$, preserves the validity of the proof of Theorem 4.5. 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 4.4. This highlights the possibility that the variance reduction may be attributed to the singularity of $F_{\mathcal{P}}$ at the points of $\mathcal{P}$.

4.6 Conclusion

Motivated by the variance reduction of Monte Carlo methods with regular point processes and inspired by gravitational allocations, we introduced the repulsion operator Π_ε. 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 and controlled by ε.

We provided a detailed theoretical study of $\Pi_\varepsilon \mathcal{P}$, where the repulsion operator is applied to a homogeneous Poisson point process $\mathcal{P}$, and suggested a practical value for the parameter ε. In particular, we proved that the repelled Poisson point process is stationary and isotropic having the same intensity as the father Poisson point process. We also proved the existence of all the moments under the repelled Poisson point process. The main theoretical result is a variance reduction for smooth linear statistics. Numerical experiments support the variance reduction and make the repelled Poisson point process a promising alternative to crude Monte Carlo if one can afford the quadratic cost of computing pairwise distances. Exploratory experiments suggest that variance reduction is also achieved when applying the repulsion operator to other point processes, such as the (hyperuniform) Ginibre ensemble and the (randomized low-discrepancy sequence) scrambled Sobol sequence. 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.

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.

Moreover, we proposed two algorithms for approximately sampling from repelled point processes. Many open questions remain in this area like developing solutions to prevent points from exiting the observation window or implementing suitable border corrections, and exploring the feasibility of further reducing the computational complexity, e.g. by employing fast multipole methods.

In another direction, it would be interesting to explore the attractive case of the operator Π_ε, achieved by selecting a negative parameter ε. 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.

5 — Perspectives

Contents

In this thesis, we studied hyperuniform point processes and we introduced the repelled point processes, both of which are regular point processes. Our motivation comes from numerical integration where using regular point processes as nodes shows improvement over the classical homogeneous Poisson point process. In this section, we present some directions for future research. Additional avenues are also delineated in the conclusions of Chapter 3 and Chapter 4, each of which is closely aligned with the specific investigations conducted in the corresponding chapter.

5.1 Hyperuniformity and low discrepancy sequences

We observed in section 3.1.3 that Monte Carlo methods with hyperuniform point processes (HUPPs) admit (under mild assumptions) fast decaying RMSE compared to crude Monte Carlo when estimating integrals of indicator functions. In particular, for $K = B(\mathbf{0}, R)$ and $\mathcal{Y}_{\overline{N}} = \mathcal{Y} \cap W$, where W is any compact of $\mathbb{R}^d$ containing K and $\mathcal{Y}$ is the rescaled version (3.3) of a Class I HUPP, we have

$$\mathbb{E}\left[\left(\frac{|W|}{\overline{N}} \sum_{\mathbf{x} \in \mathcal{Y}_{\overline{N}}} \mathbb{1}_K(\mathbf{x}) - |K|\right)^2\right]^{1/2} = O\left(\overline{N}^{-1/2-1/(2d)}\right), \tag{5.1}$$

where $\overline{N}$ is the expected number of points of $\mathcal{Y}$ in W. Moreover, mild technical assumptions allow extending this result to K other than spherical windows, (Coste, 2021). A promising avenue for future research is exploring the extension of this result to other classes of functions. One potential starting point is deriving an analogous to the Koksma-Hlawka inequality –which is tailored to low-discrepancy sequence– for HUPPs, to generalize the results in (5.1) to smooth functions.

To elaborate, recall that a random sequence $\mathcal{X}_N$ of N points in $W = [0,1]^d$ is called a *randomized low-discrepancy sequence* if each point of $\mathcal{X}_N$ is uniformly distributed over W and $\mathcal{X}_N$ exhibits low-discrepancy, i.e. there exists two positive constants M and B such that

$$\mathbb{P}\left(D(\mathcal{X}_N) < BN^{-1}\log(N)^d\right) = 1,$$

for all $N \geq M$, where $D(\mathcal{X}_N)$ is the (star) discrepancy of $\mathcal{X}_N$

$$D(\mathcal{X}_N) \triangleq \sup_{\mathbf{a}\in[0,1)^d} \left| \frac{1}{N} \sum_{\mathbf{x}\in\mathcal{X}_N} \mathbb{1}_{[\mathbf{0},\mathbf{a})}(\mathbf{x}) - |[\mathbf{0},\mathbf{a})| \right|, \tag{5.2}$$

and $[\mathbf{0},\mathbf{a})$ is the rectangular window $[0,a_1) \times \cdots \times [0,a_d)$ of volume $|[\mathbf{0},\mathbf{a})|$; see (Owen, 2013, Chapters 15 and 17). The points of a randomized low-discrepancy sequence are commonly used as nodes in randomized Quasi-Monte Carlo methods, which can achieve a convergence rate of $\log(N)^{(d-1)/2}N^{-3/2}$ under strong regularity assumptions on the integrand f; see also Section 2.7.2. In fact, a bound on the discrepancy (5.2) can be turned into a bound on integration errors using the Koksma-Hlawka inequality (Koksma, 1942/1943; Hlawka, 1962). In particular, for any function f of bounded variations in the *sense of Hardy and Krause*, we have

$$\mathbb{E}\left[\left(\frac{1}{N}\sum_{\mathbf{x}\in\mathcal{X}_N} f(\mathbf{x}) - \int_{[0,1]^d} f(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2\right]^{1/2} \leq CV_{\mathrm{HK}}(f)\mathbb{E}\left[D(\mathcal{X}_N)^2\right]^{1/2}, \tag{5.3}$$

for large N, where $V_{\mathrm{HK}}(f)$ is the total Hardy-Krause variation of f; see (Owen, 2013, Section 17.1).

Remark that the notion of discrepancy D in (5.2) bears a resemblance to hyperuniformity (5.1). Yet, D appears to be a more stringent criterion, as it quantifies the maximum error across all window sizes. On the other hand, the convergence rate in (5.1) achieved by a Class I HUPP is the optimal convergence rate for the

class of indicator functions, as discussed in Section 3.1.3, while the optimal rate for discrepancies remains unknown (Owen, 2013, Section 15.3). Hence, it's possible that the discrepancy D represents a more stringent criterion than required in this context, and hyperuniformity alone may be sufficient to achieve the optimal convergence rate for a Monte Carlo method. Investigating further the relationship between a HUPP, a randomized low-discrepancy sequence, and the estimation error in a Monte Carlo method could help understanding better the error of a Monte Carlo method. This exploration may also pave the way for deriving a counterpart of equation (5.3) for HUPPs, potentially allowing to generalize the results in (5.1) to smooth functions. If such a result is obtained, the incorporation of HUPPs into Monte Carlo methods could lead to a qualitative enhancement in both performance and theoretical foundations.

5.2 Extension of repelled point processes

In Chapter 4 we introduced the repulsion operator Π_ε in (4.1) acting on a (valid) *stationary* point process $\mathcal{X}$ via the force function $F_\mathcal{X}$ defined in (F_1). We proved in Theorem 4.5 that the repelled Poisson point process (RPPP) obtained by applying the repulsion operator to a homogeneous Poisson point process (PPP) yields a variance reduction of the corresponding linear statistics compared to the PPP. This results in particular a Monte Carlo method with a smaller variance than crude Monte Carlo. We have also verified the variance reduction in experiments. Now, we present two interesting directions for future work.

First, it's valuable for practical applications to investigate the behavior of the repulsion operator (4.1) when applied to additional points. For the sake of clarity, let's consider an example. Let $\mathcal{X}_N$ be a realization of N i.i.d. points in a compact K and $\Pi_\varepsilon \mathcal{X}_N$ the corresponding repelled sample. Suppose that we have estimated the integral $I(f)$ of a function f supported on K using the Monte Carlo estimator with nodes the points of $\Pi_\varepsilon \mathcal{X}_N$ and we found that to achieve higher accuracy, we need to increase the total number of nodes to $N + M$. The first scenario is to resample $N + M$ i.i.d. points $\mathcal{X}_{N+M}$, find $\Pi_\varepsilon \mathcal{X}_{N+M}$, and use it as the set of nodes. However, when the evaluation of the function is expensive this scenario is not desirable as we have to start all over again computing f at $N + M$ nodes. It is

preferable to sample only M additional i.i.d. points $\mathcal{X}_M$ in K, repel them, and evaluate f only on those M repelled points. Hence, an interesting avenue is to investigate whether the configuration resulting from the union of the points of $\Pi_\varepsilon \mathcal{X}_N$ and the additional repelled points of $\mathcal{X}_M$

$$\Pi_\varepsilon \mathcal{X}_N \cup \{\mathbf{x} + \varepsilon F_{\mathcal{X}_{N+M}}(\mathbf{x})\}_{\mathbf{x}\in\mathcal{X}_M},$$

is capable of producing variance reduction, akin to what we would get with $\Pi_\varepsilon \mathcal{X}_{N+M}$.

Second, the exploratory experiments of Section 4.4 suggest that variance reduction is also achieved when applying the repulsion operator to point processes other than the PPP. Proving this is a natural next step for future work, as well as proving variance reduction when applying the repulsion operator to non-stationary point processes. However, in the case of a non-stationary point process $\mathcal{X}$ of intensity function ρ we suggest adapting the force function $F_\mathcal{X}$ in F_1 to account for the inhomogeneity. One way to achieve this is by modifying the force function as follows

$$F^{\text{inh}}_\mathcal{X}(\mathbf{x}) = \sum_{\mathbf{z}\in\mathcal{X}\setminus\{\mathbf{x}\},\, \|\mathbf{x}-\mathbf{z}\|\uparrow} \frac{\beta}{\rho(\mathbf{x})\rho(\mathbf{z})} \frac{\mathbf{x}-\mathbf{z}}{\|\mathbf{x}-\mathbf{z}\|^d}, \tag{5.4}$$

where $\beta > 0$ is a tunning parameter. For $\mathcal{P}$ an inhogeneous Poisson point process of intensity function ρ, assuming the series defining $F^{\text{inh}}_\mathcal{P}$ converges, it's easy to verify that the random vector $F^{\text{inh}}_\mathcal{P}$ is centered, i.e. $\mathbb{E}[F^{\text{inh}}_\mathcal{X}(\mathbf{x})] = \mathbf{0}$ like the homogeneous counterpart (F_1). However, for $\mathbf{x} \in \mathbb{R}^d$, $F^{\text{inh}}_\mathcal{P}(\mathbf{x})$ does not mainly follow a symmetric stable distribution. Consequently, demonstrating a counterpart to Theorem 4.5 in the non-stationary case may necessitate a distinct and tailored technical approach than what we used to prove Theorem 4.5. However, once such a result is obtained under sufficiently weak assumptions, applying the repulsion operator could become a default postprocessing in many Monte Carlo integration tasks.

Bibliography

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.

G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, Cambridge, 2009. ISBN 9780511801334. doi: 10.1017/CBO9780511801334.

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. doi: 10.1201/B19708.

N. Baddour and U. Chouinard. Theory and operational rules for the discrete Hankel transform. *Journal of the Optical Society of America A*, 32(4):611, apr 2015. ISSN 1084-7529. doi: 10.1364/JOSAA.32.000611.

N. S. Bakhvalov. On the approximate calculation of multiple integrals. *Vestnik MGU, Ser. Math. Mech. Astron. Phys. Chem. in Russian, (English translation: Journal of Complexity, 2015)*, 31(4):3–18, 1959. ISSN 0885-064X. doi: https://doi.org/10.1016/j.jco.2014.12.003.

N. S. Bakhvalov. On the optimality of linear methods for operator approximation in convex classes of functions. *USSR Computational Mathematics and Mathemati-*

cal Physics, 11(4):244–249, 1971. ISSN 0041-5553. doi: https://doi.org/10.1016/0041-5553(71)90017-6.

R. Bardenet and A. Hardy. Monte Carlo with determinantal point processes. *The Annals of Applied Probability*, 30(1), feb 2020. ISSN 1050-5164. doi: 10.1214/19-AAP1504.

M. S. Bartlett. The spectral analysis of two-dimensional point processes. *Biometrika*, 51(3-4):299–311, mar 1964. ISSN 0006-3444. doi: 10.1093/biomet/51.3-4.299.

J. Beck. Irregularities of distribution. I. *Acta Mathematica*, 159(0):1–49, 1987. ISSN 0001-5962. doi: 10.1007/BF02392553.

A. Belhadji, R. Bardenet, and P. Chainais. Kernel quadrature with DPPs. volume 32, pages 12927–12937, 2019.

J. L. Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, sep 1975. ISSN 0001-0782.

C. A. N. Biscio and R. Waagepetersen. A general central limit theorem and a subsampling variance estimator for α-mixing point processes. *Scandinavian Journal of Statistics*, 46(4):1168–1190, dec 2019. ISSN 0303-6898. doi: 10.1111/sjos.12389.

C. Bonferroni. *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze. Seeber, 1936.

J. Boursier. Optimal local laws and CLT for 1D long-range Riesz gases. *ArXiv e-prints*, Dec 2021. doi: 10.48550/arXiv.2112.05881.

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.

S. N. Chiu, D. Stoyan, W. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley Series in Probability and Statistics. Wiley, aug 2013. ISBN 9780470664810. doi: 10.1002/9781118658222.

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 $[0,1]^i$, $i = 1,\dots,d$, 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.

D. Daley and R. Vesilo. Long range dependence of point processes, with queueing examples. *Stochastic Processes and their Applications*, 70(2):265–282, 1997. ISSN 0304-4149. doi: https://doi.org/10.1016/S0304-4149(97)00045-8.

D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer, 2 edition, 2003. ISBN 978-0-387-21564-8.

B. Delyon and F. Portier. Integral approximation by kernel smoothing. *Bernoulli*, 22(4):2177 – 2208, 2016.

D. Dereudre and D. Flimmel. Hyperuniformity of gibbs point processes with short range interaction. 2023.

P. J. Diggle. On parameter estimation for spatial point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(2):178–181, 1978. ISSN 00359246.

P. J. Diggle, D. J. Gates, and A. Stibbard. A nonparametric estimator for pairwise-interaction point processes. *Biometrika*, 74(4):763–770, dec 1987. ISSN 0006-3444. doi: 10.1093/BIOMET/74.4.763.

I. Dimov. *Monte Carlo Methods for Applied Scientists*. World Scientific, 2008. ISBN 9789812779892.

S. M. Ermakov and V. G. Zolotukhin. Polynomial approximations and the Monte-Carlo method. *Theory of Probability & Its Applications*, 5(4):428–431, 1960. doi: 10.1137/1105046.

J. Fox. *Applied regression analysis and generalized linear models*. Sage Publications, 2015.

A. Gabrielli, M. Joyce, and F. S. Labini. Glass-like universe: Real-space correlation properties of standard cosmological models. *Physical Review D*, 65(8):083523, 2002.

C. F. Gauss. *Methodus Nova Integralium Valores per Approximationem Inveniendi*. H. Dieterich, 1815.

G. Gautier. *On sampling determinantal point processes.* PhD thesis, Centrale Lille Institut, 2020.

G. Gautier, R. Bardenet, and M. Valko. On two ways to use determinantal point processes for Monte Carlo integration. *Advances in Neural Information Processing Systems (NeurIPS)*, 32, 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.

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)*, 2019c. Code at http://github.com/guilgautier/DPPy/ Documentation at http://dppy.readthedocs.io/.

J. Ginibre. Statistical Ensembles of Complex, Quaternion and Real Matrices. *J. Math. Phys.*, 6:440–449, 1965. doi: 10.1063/1.1704292.

M. Guizar-Sicairos and J. C. Gutiérrez-Vega. Computation of quasi-discrete Hankel transforms of integer order for propagating optical wave fields. *Journal of the Optical Society of America A*, 21(1):53, jan 2004. ISSN 1084-7529. doi: 10.1364/JOSAA.21.000053.

D. Hawat, R. Bardenet, and R. Lachièze-Rey. Repelled point processes with application to numerical integration. 2023a.

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, 2023b. doi: 10.1007/s11222-023-10219-1.

E. Hlawka. Zur angenäherten berechnung mehrfacher integrale. *Monatshefte fur mathematik*, 66(2):140–151, 1962.

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.

J. Kim and S. Torquato. Effect of imperfections on the hyperuniformity of many-body systems. *Phys. Rev. B*, 97:054105, Feb 2018. doi: 10.1103/PhysRevB.97.054105.

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. doi: 10.1038/s41467-019-08360-5.

M. A. Klatt, G. Last, and D. Yogeshwaran. Hyperuniform and rigid stable matchings. *Random Structures & Algorithms*, 57(2):439–473, sep 2020. ISSN 1042-9832. doi: 10.1002/rsa.20923.

M. A. Klatt, G. Last, and N. Henze. A genuine test for hyperuniformity. arXiv, 2022. doi: 10.48550/ARXIV.2210.12790.

J. Koksma. Een algemeene stelling uit de theorie der gelijk matige verdeeling modulo 1. *Mathematica B (Zutphen)*, 11:7–11, 1942/1943.

L. J. Landau. Bessel Functions: Monotonicity and Bounds. *Journal of the London Mathematical Society*, 61(1):197–215, feb 2000. ISSN 00246107. doi: 10.1112/S0024610799008352.

G. Last and M. Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017. doi: 10.1017/9781316104477.

T. Leblé and S. Serfaty. Fluctuations of two-dimensional Coulomb gases. *Geometric and Functional Analysis*, 28, 2018. ISSN 1420-8970. doi: 10.1007/s00039-018-0443-1.

R. Leluc, F. Portier, J. Segers, and A. Zhuman. Speeding up Monte Carlo integration: Control neighbors for optimal convergence. 2023.

O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7(1):83–122, 1975. ISSN 00018678.

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.

J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, sep 2003. ISBN 9781135442286. doi: 10.1201/9780203496930.

S. Murray and F. Poulin. hankel: A Python library for performing simple and accurate Hankel transformations. *Journal of Open Source Software*, 4(37):1397, may 2019. ISSN 2475-9066. doi: 10.21105/joss.01397.

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. Neyman and E. L. Scott. Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(1):1–43, 1958. ISSN 00359246.

J. Nolan. *Univariate Stable Distributions, Models for Heavy Tailed Data*. Springer Cham, 01 2020. ISBN 978-3-030-52914-7.

E. Novak. Some results on the complexity of numerical integration. *In Monte Carlo and Quasi-Monte Carlo Methods*, page 161–183, 2016. doi: https://doi.org/10.1007/978-3-319-33507-0_6.

E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*. European Mathematical Society (EMS), Zŕich, 2008. ISBN 9783037190265. doi: 10.4171/026.

H. Ogata. A numerical integration formula based on the Bessel functions. *Publications of the Research Institute for Mathematical Sciences*, 41(4):949–970, 2005. ISSN 0034-5318. doi: 10.2977/prims/1145474602.

B. Osgood. *Lecture Notes for EE 261 the Fourier Transform and Its Applications*. CreateSpace Independent Publishing Platform, 2014. ISBN 9781505614497.

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.

D. B. Percival and A. T. Walden. *Spectral Analysis for Univariate Time Series*, volume 51. Cambridge University Press, mar 2020. ISBN 9781139235723. doi: 10.1017/9781139235723.

T. A. Rajala, S. C. Olhede, and D. J. Murrell. Spectral estimation for spatial point patterns. *ArXiv e-prints*, sep 2020. doi: 10.48550/arxiv.2009.01474.

T. A. Rajala, S. C. Olhede, J. P. Grainger, and D. J. Murrell. What is the Fourier transform of a spatial point process? *IEEE Transactions on Information Theory*, 69(8):5219–5252, 2023. doi: 10.1109/TIT.2023.3269514.

C.-h. Rhee and P. W. Glynn. Unbiased estimation with square root convergence for SDE models. *Operations Research*, 63(5):1026–1043, 2015. Publisher: INFORMS.

K. Riedel and A. Sidorenko. Minimum bias multiple taper spectral estimation. *IEEE Transactions on Signal Processing*, 43(1):188–195, mar 1995. ISSN 1053587X. doi: 10.1109/78.365298.

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.

M. T. Thomas. A generalization of Poisson's binomial limit for use in ecology. *Biometrika*, 36 Pt. 1-2:18–25, 1949.

D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982. ISSN 0018-9219. doi: 10.1109/PROC.1982.12433.

S. Torquato. Hyperuniform states of matter. *Physics Reports*, 745:1–95, jun 2018. ISSN 03701573. doi: 10.1016/j.physrep.2018.03.001.

S. Torquato and F. H. Stillinger. Local density fluctuations, hyperuniformity, and order metrics. *Physical Review E*, 68(4), oct 2003. doi: 10.1103/physreve.68.041113.

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, İ. 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.

A — The Fourier transform

The Fourier transform $\mathcal{F}$ of an integrable function $f : \mathbb{R}^d \to \mathbb{C}$ is the square integrable function

$$\mathcal{F}(f)(\mathbf{k}) = \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-\mathrm{i}\langle \mathbf{k}, \mathbf{x} \rangle} \mathrm{d}\mathbf{x}, \tag{A.1}$$

where $\langle \mathbf{k}, \mathbf{x} \rangle$ is the dot product of the *wavevector* $\mathbf{k}$ by $\mathbf{x}$. If f is furthermore a radial function, denoted abusively by $f(\mathbf{r}) = f(r)$, where $r = \|\mathbf{r}\|_2$, then the Fourier transform of f is the corresponding *symmetric Fourier transform* $\mathcal{F}_s$, namely

$$\mathcal{F}_s(f)(k) = \mathcal{F}(f)(\mathbf{k}) = (2\pi)^{d/2} \int_0^\infty r^{d/2} f(r) \frac{J_{d/2-1}(kr)}{k^{d/2-1}} \mathrm{d}r, \tag{A.2}$$

where $k = \|\mathbf{k}\|_2$ is called a *wavenumber*, and J_ν is the Bessel function of the first kind of order ν (Osgood, 2014). If we further define the *Hankel transform* of order $\nu \geq -1/2$ as

$$\mathcal{H}_\nu(f)(k) = \int_0^\infty f(r) J_\nu(kr) r \mathrm{d}r, \quad \text{for } k \geq 0, \tag{A.3}$$

then (A.2) rewrites as

$$\mathcal{F}_s(f)(k) = \frac{(2\pi)^{d/2}}{k^{d/2-1}} \mathcal{H}_{d/2-1}(\tilde{f})(k), \tag{A.4}$$

where $\tilde{f} : x \mapsto f(x) x^{d/2-1}$. Finally, note that the Fourier transform can be generalized to tempered distributions through duality (Coste, 2021, Appendix B).

Example A.1. Let $W \subset \mathbb{R}^d$ be a convex set, and denote by $|W|$ its Lebesgue measure, i.e., its volume. The Fourier transform of the scaled intersection volume $\alpha_0(., W)$, defined by Equation (3.8), is given by

$$\mathcal{F}(\alpha_0)(\mathbf{k}, W) = \frac{1}{|W|} \left(\mathcal{F}(\mathbb{1}_W)(\mathbf{k})\right)^2. \tag{A.5}$$

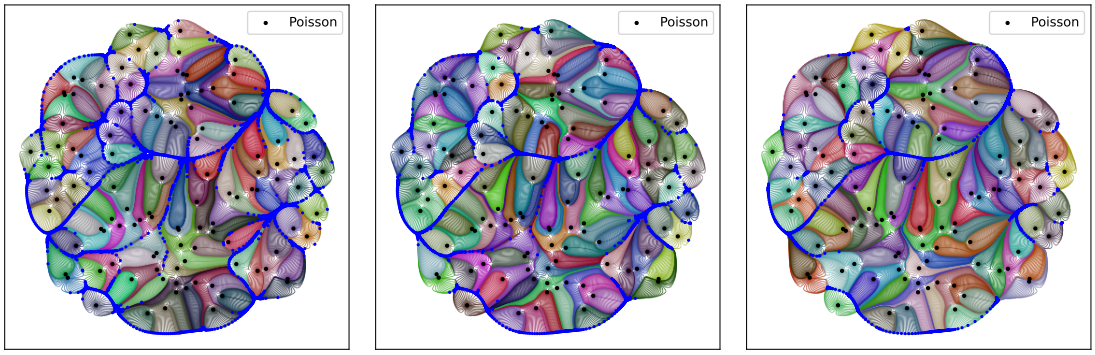
In particular, if $W = B(\mathbf{0}, R)$ is the Euclidean ball of radius R, $\alpha_0(., W)$ is a radial function (Torquato, 2018, Section 3.1.1). In this case, we abusively write $\alpha_0(\mathbf{r}, W) = \alpha_0(r, W)$ and

$$\mathcal{F}_s(\alpha_0)(k, W) = 2^d \pi^{d/2} \frac{\Gamma(1 + d/2)}{k^d} J_{d/2}^2(kR), \tag{A.6}$$

where Γ is Euler's Gamma function. On the other hand, if $W = \prod_{j=1}^{d}[-L_j/2, L_j/2]$, Equation (A.5) simplifies to

$$\mathcal{F}(\alpha_0)(\mathbf{k}, W) = \left(\prod_{j=1}^{d} \frac{\sin(k_j L_j/2)}{k_j \sqrt{L_j}/2} \right)^2 . \tag{A.7}$$

B — Simulation of a gravitational allocation



(a) $T_{\max} = 400$ (b) $T_{\max} = 600$ (c) $T_{\max} = 1000$

Figure B.1: Illustration of a gravitational allocation from Lebesgue to a realization (black dots) of a PPP in a disk. The region underlying the curves sharing the same color illustrates a basin, which collects the points of the space allocated to the point of the PPP that belongs to that particular colored region. The additional blue points are the positions where the trajectories of the opposite flow curves stop for $T_{\max} = 400$ (left), $T_{\max} = 600$ (middle), and $T_{\max} = 1000$ (right). Note that the provided image is purely illustrative, and no claims are being made regarding the existence of a gravitational allocation from the Lebesgue measure to a PPP in dimension $d = 2$. Plots obtained using the toolbox MCRPPy.

Consider a point process $\mathcal{X}$ in $\mathbb{R}^d$, and let's assume the existence of a gravitational allocation from Lebesgue to this point process. We've previously explored this concept in Example 2.5 and we emphasized that, in general, a gravitational allocation involves a differential equation that describes the motion of the points

of $\mathbb{R}^d \setminus \mathcal{X}$. This motion is driven by either a potential function U (as in (2.19)) or a force function F (as in (2.20)) exerted by the points in $\mathcal{X}$. Note that the choice between using U or F depends on the specific context and the presence of either of them. For completeness, we provide the algorithm commonly used for generating visual representations of gravitational allocations, as exemplified in Figure 2.5, and explain the underlying concept. For simplicity, we assume the existence of the force function F and that the gravitational allocation is based on F as in Chatterjee et al. (2010). The scenario involving a potential function can be obtained by substituting F with $-\nabla U$. We extend our gratitude to Manjunath Krishnapur for generously sharing his MATLAB code and providing valuable insights into gravitational allocations.

The common approach to illustrate a gravitational involves discretization of the *opposite* gravitational flow curves of some points situated around the points of $\mathcal{X}$. Specifically, for each point $\mathbf{x}$ in $\mathcal{X}$, we select M points $\{\mathbf{y}_i^{\mathbf{x}}\}_{i=1}^M$ within the sphere $S(\mathbf{x}, \epsilon)$ centered at $\mathbf{x}$ of (small) radius $\epsilon > 0$. Then, we plot a discretization of the solutions to the differential equation

$$\frac{\partial}{\partial t} Y_{\mathbf{x}}(t) = -F(Y_{\mathbf{x}}(t)), \tag{B.1}$$

with starting points $\{\mathbf{y}_i^{\mathbf{x}}\}_{i=1}^M$, up to a maximum time $T_{\max} > 0$. In the same vein as Example 2.5, a solution of (B.1) with starting point $\mathbf{z}$ is called the *opposite flow curve* of $\mathbf{z}$, since (B.1) have the opposite sign of the differential equation (2.20). In other words, for each i, we draw $\mathbf{y}_i^{\mathbf{x}}(t)$ iteratively

$$\mathbf{y}_i^{\mathbf{x}}(t) = \mathbf{y}_i^{\mathbf{x}}(t-1) - hF(\mathbf{y}_i^{\mathbf{x}}(t-1)); \quad \mathbf{y}_i^{\mathbf{x}}(0) = \mathbf{y}_i^{\mathbf{x}},$$

where $h > 0$ is the discretization stepsize and $t \in \{0, \ldots, T_{\max}\}$. We present the pseudo-code in Algorithm 3.

Understanding Algorithm 3 requires grasping the formation of the boundaries of the basins. To elaborate, let $\mathcal{B}_{\mathbf{x}}$ be the basin of $\mathbf{x} \in \mathcal{X}$. $\mathcal{B}_{\mathbf{x}}$ is the collection of points of $\mathbb{R}^d \setminus \mathcal{X}$ that their gravitational flow curves terminate at $\mathbf{x}$. Consider a point $\mathbf{y} \in \mathbb{R}^d \setminus \mathcal{X}$. There are two possible scenarios for $\mathbf{y}$ to be located on the boundary $\partial\mathcal{B}_{\mathbf{x}}$ of $\mathcal{B}_{\mathbf{x}}$. The first scenario entails $F(\mathbf{y}) = 0$, indicating that $\mathbf{y}$ remains unaffected by the force F exerted by $\mathcal{X}$. The second scenario is characterized by a balanced force at $\mathbf{y}$, i.e. the vector $F(\mathbf{y})$ aligns with the boundary $\partial\mathcal{B}_{\mathbf{x}}$ and $F(\mathbf{y})$

is orthogonal to the unit normal vector of $\partial\mathcal{B}_{\mathbf{x}}$. In this case, the gravitational flow curve of $\mathbf{y}$ belongs to $\partial\mathcal{B}_{\mathbf{x}}$. Roughly, the collective gravitational flow curves of such points constitute $\partial\mathcal{B}_{\mathbf{x}}$. Therefore, the first approach for approximating the basins involves discretizing the gravitational flow curves of the boundary points, which necessitates locating the boundary points rendering this scenario hard in practice.

Alternatively, we can reasonably expect that the points $\{\mathbf{y}_i^{\mathbf{x}}\}_{i=1}^M$ in close proximity to $\mathbf{x}$ belong to $\mathcal{B}_{\mathbf{x}}$. Correspondingly, the opposite flow curves of $\{\mathbf{y}_i^{\mathbf{x}}\}_{i=1}^M$ terminate at $\partial\mathcal{B}\mathbf{x}$, so discretizing them with a small step size h offers a reasonable approximation of $\mathcal{B}_{\mathbf{x}}$. This is the basic idea behind Algorithm 3 which is implemented within MCRPPy[1]; see aslo the tutorial Jupyter Notebook[2].

Algorithm 3 Illustration of a gravitational allocation

sample a point pattern $\{\mathbf{x}_i\}_{i=1}^N$ from the point process $\mathcal{X}$
plot $\{\mathbf{x}_i\}_{i=1}^N$
fix M, h, $T_{\max}$, and ϵ positive
for $i : 1 \to N$ **do**
    choose $\{\mathbf{y}_j\}_{j=1}^M$ in $S(\mathbf{x}_i, \epsilon)$
    choose a color c
    plot $\{\mathbf{y}_j\}_{j=1}^M$ with $color = c$
    for $t : 1 \to T_{\max}$ **do**
        for $j : 1 \to M$ **do**
            $\mathbf{y}_j = \mathbf{y}_j - hF(\mathbf{y}_j)$
        end for
        plot $\{\mathbf{y}_j\}_{j=1}^M$ with $color = c$
    end for
end for

Figure B.1 illustrates a gravitational allocation from Lebesgue to a realization of a PPP is a disk using Algorithm 3. We used a stepsize $h = 0.005$ and a maximum time $T_{\max} = 400$ in the left panel, $T_{\max} = 600$ in the middle panel, and

[1] https://github.com/dhawat/MCRPPy/blob/main/src/mcrppy/gravitational_allocation.py

[2] https://github.com/dhawat/MCRPPy/blob/main/notebooks/tutorial_gravitational_allocation.ipynb

$T_{\max} = 1000$ in the right panel. The black dots represent the point pattern while the blue dots display the positions of the endpoints of the opposite flow curves, i.e. $\{\mathbf{y}_i^{\mathbf{x}}(T_{\max})\}_{1 \le i \le M, \mathbf{x} \in \mathcal{X}}$. When an opposite flow curve reaches the boundary of the corresponding basin, it may not halt immediately; it only ceases when it reaches a point where the force becomes zero. This is visible in Figure B.1 from the distinct positions of the blue dots when $T_{\max}$ varies from 400 to 1000.

C — Additional illustrations

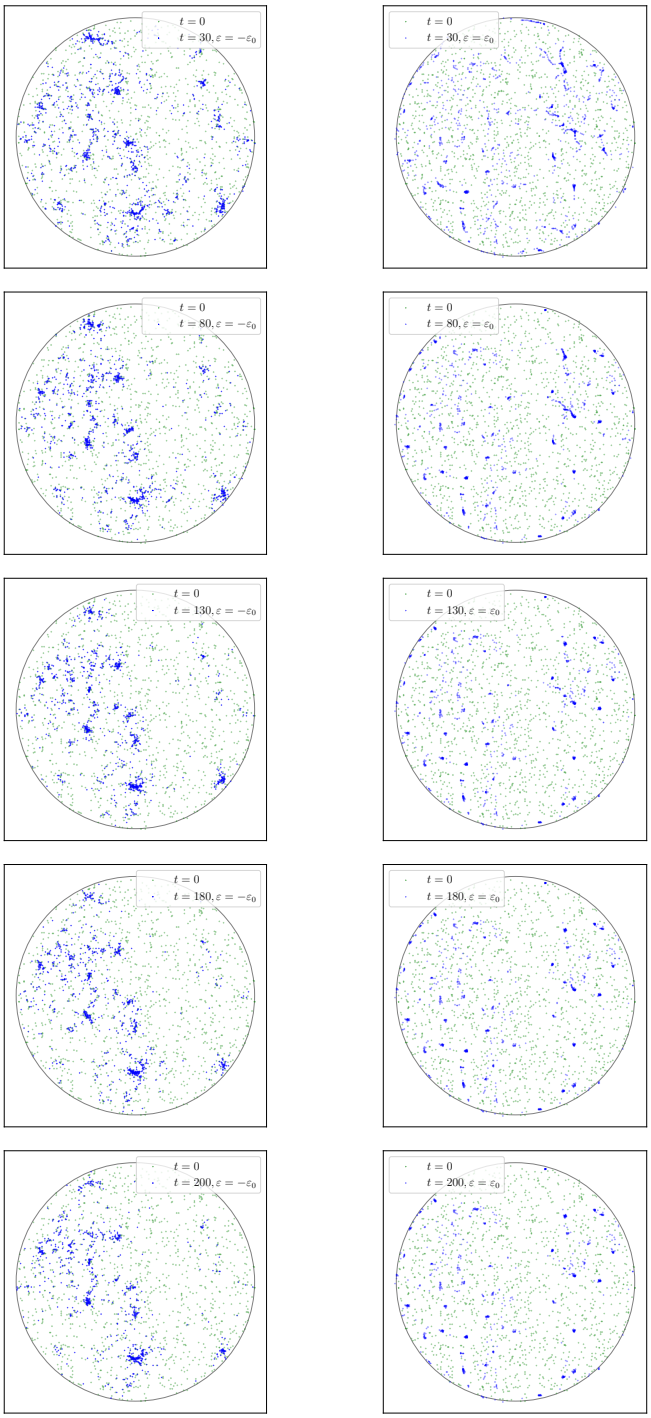


Figure C.1: The green points represent a sample from a PPP $\mathcal{P}$ of unit intensity, the blue points correspond to $\Pi_{-\varepsilon_0,t}\mathcal{P}$ (left), and $\Pi_{+\varepsilon_0,t}\mathcal{P}$ (right), for $t \in \{30, 80, 130, 180, 200\}$. Plots obtained using the toolbox MCRPPy.

D — Résumé en Français (Summary in French)

L'intégration numérique consiste à approximer l'intégrale

$$I(f) = \int_{\mathbb{R}^d} f(\mathbf{x}) \mathrm{d}\mathbf{x}$$

d'une fonction intégrable $f : \mathbb{R}^d \to \mathbb{R}$ en utilisant des évaluations ponctuelles de f. L'algorithme d'approximation le plus courant consiste à sommer les évaluations de f en N points $(\mathbf{x}_i)_{i=1}^N$, appelés *nœuds*, multiplié par des *poids* appropriés $(w_i)_{i=1}^N$, c'est-à-dire

$$I(f) \approx \sum_{i=1}^{N} w_i f(\mathbf{x}_i). \tag{D.1}$$

Outre l'analogie de l'approximation dans (D.1) avec une somme de Riemann, Bakhvalov (1971) a prouvé que l'erreur minimale dans une classe de fonctions F peut être obtenue par un algorithme de cette forme (D.1), à condition que F satisfasse certaines propriétés. Plus précisément, pour trouver la borne inférieure

$$\inf_{\widehat{I} \in \mathcal{A}} \sup_{f \in F} |I(f) - \widehat{I}(f)|,$$

sur l'ensemble $\mathcal{A}$ de tous les algorithmes approximant $I(f)$, il suffit de chercher parmi les algorithmes de la forme (D.1) si F est convexe, c'est-à-dire, si $f_1, f_2 \in F$, alors leur combinaison convexe $tf_1 + (1-t)f_2 \in F$ pour $t \in [0,1]$ et centrée symétriquement, c'est-à-dire, si $f \in F$, alors $-f \in F$; voir (Novak and Woźniakowski, 2008, Section 4.2.2). Ainsi, (D.1) semble un choix raisonnable d'approximation et le point focal restant réside dans la détermination de la configuration des nœuds $(\mathbf{x}_i)_{i=1}^N$ (et les poids correspondants) qui contrôle l'erreur, pour une large classe de fonctions, tout en garantissant un délai de calcul raisonnable.

Parmis les algorithmes de la forme (D.1), on peut distinguer deux types : les constructions déterministes de nœuds, appelées *règles de quadrature*, et les configurations aléatoires de nœuds, appelées *méthodes de Monte Carlo*. En notant que l'intégrale d'une fonction lisse f est l'aire engendrée par le graphe de f, les règles de quadrature les plus basiques estiment cette aire en utilisant la somme des aires de fenêtres rectangulaires/trapèzoidales adjacentes. En faible dimension d, la littérature regorge de règles de quadrature notables. Par exemple, la quadrature de Gauss est exacte, c'est-à-dire que l'erreur d'approximation est égale à zéro lorsque f est un polynôme (unidimensionnel) de degré jusqu'à $2N-1$ (Gauss, 1815) ; voir aussi (Owen, 2013, Chapitre 7). Cependant, l'avantage des règles de quadrature diminue rapidement lorsque la dimension augmente, et les méthodes de Monte Carlo deviennent plus appropriées. Ce phénomène montre que les constructions déterministes souffrent de la malédiction de la dimension et a été formulé dans un théorème de Bakhvalov (1959) ; voir aussi (Novak, 2016), (Owen, 2013, Chapitre 7) et (Dimov, 2008, Chapitre 3). Le théorème de Bakhvalov fournit une borne inférieure de l'erreur des règles de quadrature et de la racine de l'erreur quadratique moyenne (RMSE) des méthodes de Monte Carlo, pour la classe de fonctions C^k à support compact ayant des dérivées partielles bornées jusqu'à l'ordre $k > 0$. Plus précisément, le théorème indique que le taux de convergence de l'erreur dans le pire cas d'une règle de quadrature est borné par $N^{-k/d}$ comparée à $N^{-1/2-k/d}$ pour la RMSE au pire cas d'une méthode de Monte Carlo. Cela met en évidence une possible amélioration d'ordre $N^{-1/2}$ d'une méthode de Monte Carlo par rapport à une règle de quadrature dans n'importe quelle dimension d'espace et motive les méthodes de Monte Carlo pour l'estimation des intégrales en grande dimension.

Dans sa forme la plus élémentaire, la méthode Monte Carlo brute consiste à représenter l'intégrale d'une fonction à support compact K comme une espérance sous une distribution uniforme, et à l'estimer par la moyenne empirique basée sur N échantillons uniformes i.i.d. dans K. La Figure 1.1a montre 1000 points uniformes i.i.d. observés dans le cube $[-1,1]^2$. Des calcules classiques en probabilité montrent que la RMSE de la méthode Monte Carlo brute a la forme $\sigma(d)N^{-1/2}$. Le taux de convergence $N^{-1/2}$ de la RMSE est indépendant de la dimension d et l'échantillonnage des nœuds est peu coûteux, favorisant la méthode Monte Carlo

brute en grande dimension. Cependant, mettant de côté la dépendance de $\sigma(d)$ en d, $N^{-1/2}$ est un taux de convergence lent : diviser l'erreur par 10 nécessite de multiplier le nombre de nœuds N par un facteur de 100. Cela devient défavorable dans des scénarios exigeant une estimation de haute précision, en particulier lorsque l'évaluation de f coûte cher. Certaines méthodes visent à améliorer ce taux de convergence en introduisant une dépendance sophistiquée à travers les poids (Delyon and Portier, 2016; Azaïs et al., 2018; Leluc et al., 2023). D'autres méthodes, appelées méthodes de réduction de la variance, visent à réduire $\sigma(d)$, par exemple en utilisant une fonction auxiliaire ayant un intégrale connue ; voir (Owen, 2013, Chapitres 8 et 10) pour des résultats classiques et (South et al., 2022) ainsi que les références qui y sont mentionnées pour des travaux plus récents. Alternativement, on peut envisager une configuration de nœuds plus régulièrement répartie dans K qu'un tirage i.i.d., comme dans la méthode quasi-Monte Carlo randomisée (RQMC) qui accélère le taux de convergence (Owen, 2008). La Figure 1.1b illustre 1000 nœuds utilisés en RQMC. Nous observons une disposition de points plus régulière que les 1000 points i.i.d. affichés dans la Figure 1.1a.

En statistiques spatiales, nous appelons une configuration aléatoire de points un *processus ponctuel* (Chiu et al., 2013; Møller and Waagepetersen, 2003; Daley and Vere-Jones, 2003). En particulier, la configuration de points uniformes i.i.d. correspond au *processus ponctuel binomial* (BPP). Le BPP converge vers le célèbre *processus ponctuel de Poisson homogène* (PPP) lorsque le nombre de points approche correctement l'infini (Chiu et al., 2013; Coeurjolly et al., 2017; Last and Penrose, 2017). Cela a positionné le PPP comme le processus de référence associé à la méthode Monte Carlo brute. Cependant, lorsqu'on travaille avec le PPP, ainsi qu'avec de nombreux autres processus ponctuels, le nombre de nœuds N est lui-même une variable aléatoire. Dans de tels scénarios, nous exprimons le taux de convergence de la RMSE en utilisant l'espérance du nombre de nœuds $\mathbb{E}[N]$ au lieu de la valeur conventionnelle N. Des processus ponctuels plus sophistiqués peuvent également être utilisés en intégration numérique. En effet, la première meusure d'intensité d'un processus ponctuel $\mathcal{X}$ ayant une intensité constante $\rho > 0$ (espé-

rance du nombre de points par unité de volume) relie l'intégrale d'une fonction à support compact f à l'espérance d'une statistique linéaire

$$\mathbb{E}\left[\frac{1}{\rho}\sum_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})\right] = \int_{\mathbb{R}^d} f(\mathbf{x})\mathrm{d}\mathbf{x},$$

offrant ainsi une méthode de Monte Carlo *non biaisée* basée sur $\mathcal{X}$. De manière générale, nous appelons un processus ponctuel *régulier* lorsqu'il est plus régulièrement réparti que le PPP, comme dans la Figure 1.1b. Les processus ponctuels réguliers semblent plus prometteurs que le PPP en intégration numérique comme ils présentent une meilleure représentation de l'espace. Une classe intéressante de processus ponctuels réguliers est la classe des *processus ponctuels déterminantaux* (DPPs) (Macchi, 1975; Hough et al., 2009). Les méthodes de Monte Carlo avec des DPPs améliorent le taux de convergence de la RMSE au-delà du $O(N^{-1/2})$ (Ermakov and Zolotukhin, 1960; Bardenet and Hardy, 2020; Coeurjolly et al., 2021; Belhadji et al., 2019). Le principal goulot d'étranglement des méthodes basées sur des DPPs est la complexité d'échantillonnage élevée des DPPs.

L'objectif de cette thèse est d'étudier des processus ponctuels réguliers qui sont des candidats potentiels pour accélérer la convergence de la méthode Monte Carlo brute. Une classe importante de processus ponctuels réguliers est la classe des processus ponctuels *hyperuniformes* (HUPPs) (Torquato, 2018; Torquato and Stillinger, 2003; Kim and Torquato, 2018; Klatt et al., 2019, 2022, 2020; Coste, 2021; Dereudre and Flimmel, 2023). Un processus ponctuel est HUPP si la variance du nombre de points dans une grande fenêtre décroît plus lentement que le volume de cette fenêtre. Formellement, $\mathcal{X}$ est un HUPP si

$$\lim_{R\to\infty} \frac{\mathbb{V}\mathrm{ar}\left[\mathcal{X}(B(\mathbf{0},R))\right]}{|B(\mathbf{0},R)|} = 0,$$

où $B(\mathbf{0},R)$ est une boule de rayon R centrée à l'origine. En particulier, tout HUPP permet d'obtenir une méthode de Monte Carlo avec une décroissance plus rapide de la RMSE par rapport à la méthode Monte Carlo brute, lorsque f est une fonction indicatrice. La Figure 1.1c montre un sample de 1000 points d'un HUPP. En dehors de l'intégration numérique, cette propriété statistique implique des propriétés physiques souhaitables pour les matériaux (Torquato, 2018, Section 14). Cependant, prouver rigoureusement qu'un processus ponctuel est HUPP est généralement difficile. La pratique courante est d'utiliser quelques échantillons

du processus ponctuel pour estimer une mesure spectrale appelée la fonction de structure. La décroissance de la fonction de structure autour de zéro fournit un indicateur de l'hyperuniformité. La pratique courante est d'estimer la fonction de structure du processus ponctuel à partir d'un ou plusieurs échantillons, puis à analyser son comportement près de zéro pour détecter l'hyperuniformité en se référant à des règles empiriques (Torquato, 2018; Klatt et al., 2019). Cependant, cette évaluation graphique n'est pas normalisée et n'est souvent pas décrite en détail de manière reproductible dans la littérature, les choix d'implémentation et les propriétés statistiques faisant souvent partie du folklore de chaque domaine. Notre première contribution est une revue et dérivation des estimateurs existant de la fonction de structure (Hawat et al., 2023b). Nous utilisons également ces estimateurs pour proposer un test statistique de l'hyperuniformité asymptotiquement valide. Récemment un autre test statistique d'hyperuniformité a été proposé par Klatt et al. (2022), adapté à un type particulier (Classe I) de HUPPs. Nous avons également construit une librairie de Python modulaire et open-source appelé structure_factor, qui implémente tous les estimateurs, diagnostics et tests discutés.

Revenant à l'intégration numérique, dans le but d'imiter la régularité d'un DPPs avec un moindre coût computationnel, nous proposons un operatuer nommé l'opérateur de répulsion. Cet opérateur réduit le regroupement de points dans une configuration en écartant légèrement les points les uns des autres (Hawat et al., 2023a). Le choix de l'opérateur est motivé par les allocations gravitationnelles des processus ponctuels vers la mesure de Lebesgue (Chatterjee et al., 2010; Nazarov et al., 2007) et par les gaz de Coulomb (Serfaty, 2019; Leblé and Serfaty, 2018). Notre principal résultat théorique est que l'application de l'opérateur de répulsion à un PPP donne lieu à une méthode de Monte Carlo admettant une variance plus petite que la méthode Monte Carlo brute. En termes de complexité computationnelle, comparativement à une complexité (au moins) cubique pour la méthode Monte Carlo avec DPPs de Bardenet and Hardy (2020), la complexité de notre méthode de Monte Carlo est seulement quadratique et peut être facilement parallélisée sans communication entre les tâches. Les expériences montrent que la méthode de Monte Carlo que nous proposons est compétitive avec la RQMC. De plus, les résultats des expériences laissent entendre que la réduction de la va-

riance persiste lorsque l'opérateur est appliqué à des processus ponctuels invariants par translation qui sont plus reguliers que le PPP. Cet observation suggère la possibilité que l'opérateur de répulsion est universel, dans le sens où il entraîne systématiquement une réduction de la variance, quel que soit le processus ponctuel auquel il est appliqué.