The Dissertation Committee for Elizabeth Watson O'Reilly
certifies that this is the approved version of the following dissertation:

Repulsion of Determinantal Point Processes and Stationary Poisson Tessellations in High Dimensions

Committee:

François Baccelli, Supervisor

Rachel Ward

Gordan Žitković

Sanjay Shakkottai

Repulsion of Determinantal Point Processes and Stationary Poisson Tessellations in High Dimensions

by

Elizabeth Watson O'Reilly,

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2019

Dedicated to my family.

Acknowledgments

I would first like to thank my supervisor, François Baccelli, who has been an incredibly patient, encouraging, and inspiring mentor throughout my PhD. I would like to thank Gordan Žitković, Rachel Ward, and Sanjay Shakkottai for serving on my committee. I am also thankful to Jesper Møller for the collaboration leading to the first chapter of this dissertation. I would also like to thank Gustavo de Veciana for his mentorship and collaboration on my first research project in graduate school.

I wish to thank my parents, Anne and John, whose unconditional love and support got me to where I am today, as well as my two wonderful sisters Jane and Colleen. I also would like to thank my husband and best friend, Alfred, for being an amazing partner in this journey. This dissertation is dedicated to them and it is dedicated in memoriam of my stepmother Deborah.

Repulsion of Determinantal Point Processes and Stationary Poisson Tessellations in High Dimensions

Publication No. ____________________

Elizabeth Watson O'Reilly, Ph.D.
The University of Texas at Austin, 2019

Supervisor: François Baccelli

In this dissertation, new results on stochastic geometric models in high dimensional space are presented. We first concentrate on a particular class of repulsive point processes called determinantal point processes (DPPs). We establish a coupling of a DPP and its reduced Palm version showing the repulsive effect of a point of the point process. This is used for discussing the degree of repulsiveness in DPPs, including Ginibre point processes and other specific parametric models for DPPs.

We then study this repulsion for stationary DPPs in high dimensional Euclidean space. It is shown that for many families of DPPs, a typical point has no repulsive effect with high probability for large space dimension n. It is also proved that for some DPPs there exists an R^* such that the repulsive effect occurs at a distance of $\sqrt{n}R^*$ with high probability for large n. This

R^* is interpreted as the asymptotic reach of repulsion of the DPP. Examples of DPPs exhibiting this behavior are presented and an application to high dimensional Boolean models is given.

The second half of this dissertation examines zero cells of stationary Poisson tessellations. First, a stationary stochastic geometric model is proposed for analyzing one-bit data compression. The data is assumed to be an unconstrained stationary set, and each data point is compressed using one bit with respect to each hyperplane in a stationary and isotropic Poisson hyperplane tessellation. Size metrics of the zero cell of the tessellation are studied to determine how the intensity of hyperplanes must scale with dimension to ensure sufficient separation of different data by the hyperplanes or sufficient proximity of the data compressed together. The results have direct implications in compressive sensing and source coding.

We then study the concentration of the norm of a random vector Y uniformly sampled in the centered zero cell of a stationary random tessellation in high dimensions. It is shown that for a stationary and isotropic Poisson-Voronoi tessellation, $|Y|/\mathbb{E}(|Y|^2)^{\frac{1}{2}}$ approaches one as the dimension approaches infinity. For a stationary and isotropic Poisson hyperplane tessellation, we prove that $|Y|/\sqrt{n}$ will be within a fixed range (R_ℓ, R_u) with probability approaching one as dimension n tends to infinity.

Table of Contents

List of Tables

Chapter 1

Introduction

The field of stochastic geometry provides tools and models for the study of random geometric objects and patterns, and has found applications in many areas [21] including materials science, wireless networks, computational biology, machine learning, and information theory. Fundamental objects of study in stochastic geometry are point processes, which are random variables that take values in the space of counting measures, or equivalently, random objects taking values in the space of locally finite point configurations. There are many disciplines where phenomena can be modeled with point processes and random geometric models built from point processes such as tessellations and germ-grain models [75]. Recent work has used this theory to study probabilistic models for data analysis and transmission, for example in channel coding [4] and topological data analysis [81]. In many applications, the data is high dimensional, motivating the study of how the geometry of high dimensional space affects the models.

This thesis described new results on two different stochastic geometric models, with a particular interest in the behavior of these random objects in high dimensional space. The first models studied are determinantal point processes. The most commonly studied class of point processes are Poisson processes, where all points are stochastically independent from each other. However, models exhibiting interaction are often needed. Determinantal point processes (DPPs) are a useful class that model repulsion between particles [41, 52, 56]. They were initially introduced by Macchi in [57], and have since found uses in random matrix theory and machine learning [49]. An advantage of these point processes is that they have closed form expressions for their joint densities, but the nature of the repulsion is not clear from their definition.

In order to study the repulsion between points of a DPP, we turn to Palm theory. The Palm distribution of a point process X on a Polish space Λ with respect to some finite subset u of Λ is the conditional distribution of the point process, given that it contains points at the locations in u. The reduced Palm distribution with respect to u is the Palm distribution with the points at u removed. A useful property of DPPs is that the reduced Palm distribution has the distribution of another DPP.

To quantify repulsiveness of a DPP, we first note the following result proved in [35]: For any point u, let X be a DPP and X^u be a DPP with its reduced Palm distribution with respect to u. Then, there exists a coupling (X, X^u) such that X^u is obtained by removing a point process ξ_u from X. In joint work with Jesper Møller, this coupling has been made more precise by showing the existence of such a coupling where ξ_u has at most one point. This result says that the repulsive effect of a point at location u is to push out one other point with some probability, and the distribution of the point removed has been characterized and depends on the associated kernel of the DPP.

The initial motivation for studying DPPs was to extend threshold results on the Boolean model in high dimensions in [3]. The Boolean model studied in [3] consists of the union of balls with i.i.d radii and centers at the points of a Poisson point process in $\mathbb{R}^n$. Three different properties of the model are studied: the degree, the volume fraction, and percolation. The authors prove the existence of thresholds for the logarithmic intensity of the point process at which each property has a sharp transition as the dimension n tends to infinity. This work has applications to channel coding in information theory where the underlying Poisson point process models a codebook [4].

The question then became whether the thresholds related to degree, volume fraction, or percolation change when the underlying point process, or codebook, is not Poisson. In the application to channel coding, larger thresholds or faster rates of convergence are desirable, which leads to asking these same questions for Boolean models with underlying repulsive point processes. A popular class of models exhibiting repulsion is that of hard-core point processes [21]. These models have an intuitive repulsive nature, because no two points are allowed to be closer that some finite positive distance apart. However, in general, for these and other repulsive point processes such as pairwise interaction point processes [63], there is no closed form for their moment measures, and simulation can be difficult. DPPs provide a potentially desirable alternative since they do have a closed form for their moment measures and a relatively simple simulation procedure. The question of whether an underlying DPP would change the asymptotic behavior of the Boolean model led to a more general study of how the strength and reach of repulsion between points of DPPs interact with the geometry of high dimensional space.

Using the measure of repulsiveness related to the previously described coupling between a DPP and its reduced Palm distribution, we show that for some parametric classes of stationary and isotropic DPPs in $\mathbb{R}^n$, the effect of repulsion becomes very small in high dimensions, in the sense that placing a point at the origin has no effect with high probability. We also quantify an asymptotic reach of repulsion $R^* \in (0, \infty)$, such that with certain conditions on the associated kernel of the DPP, the repulsive effect occurs within a thin annulus around the sphere of radius $\sqrt{n}R^*$ with high probability for large dimension n. The conditions on the kernel are connected with the phenomenon of thin-shell concentration, where under certain conditions, high dimensional vectors have a norm that is concentrated near its expectation, see [14].

The second part of this thesis focuses on random tessellations, and in particular on the random polytope that contains the origin, called the zero cell. Important classes of random tessellations studied in stochastic geometry are generated from Poisson point processes, such as Poisson Voronoi and Poisson hyperplane tessellations. The main properties studied in high dimensions have been the volume and shape of particular cells, see [40] and [2]. In particular, the connection between high dimensional convex geometry and these models was studied in [40], where it is shown that there is a class of isotropic Poisson tessellations where the zero cell, that is, the cell containing the origin, satisfies the hyperplane conjecture asymptotically almost surely. If the tessellation is stationary, i.e., its distribution is invariant under translations, one can also study the distribution of the typical cell, obtained by averaging over all cells in a large bounded subset and then increasing this subset to the entire space.

This work was initially inspired by questions related to one-bit data compression. In this paradigm, a signal x is compressed into a sequence of one-bit measurements given by the measurement model $y_i = \text{sign}(\langle x, u_i \rangle - t_i)$, $i = 1, \ldots, m$ where each u_i is a random direction in the unit sphere $\mathbb{S}^{n-1}$ and t_i is a random displacement vector. Each pair (u_i, t_i) defines a hyperplane h_i in $\mathbb{R}^n$ and the measurement y_i gives the side of the hyperplane that x lies on. Thus, the measurements define a unique cell of the random hyperplane tessellation induced by the collection of hyperplanes $\{h_i\}_{i=1}^m$. Ensuring all of the data within a cell of the tessellation is close together ensures that the signal can be recovered with small error, and requires certain geometric constraints on the these random polytopes that can be guaranteed with high probability in high dimensions. This idea can be applied to recovering a codeword associated with the cell, as in source coding, or by reconstructing a high dimensional signal with a convex program, as in one-bit compressed sensing [13, 8, 7, 69].

We present a model for the compression using a stationary Poisson hyperplane process on all of $\mathbb{R}^n$. We consider the signal set to be all of $\mathbb{R}^n$ or stationary Poisson point process in $\mathbb{R}^n$, and we study the case of a typical signal at the origin, thus asking for geometric constraints on the zero cell that would ensure recovery of this typical signal with high probability. We prove results giving the scale at which the intensity of the hyperplane process must grow with dimension n so that a sufficient degree of separation or distortion is obtained with high probability for large dimension n. Additionally, we study a different metric of the zero cell inspired by this compression model for both the Poisson hyperplane as well as the Poisson-Voronoi tessellation. This metric is the norm of the random vector that is, conditionally on the tessellation, chosen uniformly at random from the zero cell. This is a measure of the distance of the mass of the cell from the origin. We show to what extent this norm concentrates as dimension tends to infinity.

1.1 Outline

Chapters 3 and 4 focus on determinantal point processes. Chapter 3 discusses the new result on the repulsive nature of DPPs by examining a coupling between a DPP and its reduced Palm distribution. In Chapter 4 this characterization of repulsion is used to study the strength and reach of the repulsive effect of a point as the space dimension tends to infinity. Chapters 5 and 6 describe the results on high dimensional stationary Poisson tessellations. Chapter 5 describes the model for one-bit compression and Chapter 6 presents the results on the concentration of the norm of the vector chosen uniformly from the zero cell in high dimensions.

1.2 Papers included

Chapter 3 is taken from [61], which is joint work with Jesper Møller. Chapter 4 is based on the paper [5] that is joint with François Baccelli, but as this paper was written before [61], the results have been reformulated to use the new coupling result. Chapter 5 is taken from the paper [6], which is also joint work with François Baccelli, and Chapter 6 is taken from [66].

Chapter 2

Preliminaries and Notation

In this chapter, we briefly cover some background material, definitions, and known results used throughout the remaining chapters. General references on point processes and the stochastic geometric models studied here include [26, 75, 21].

2.1 Point Processes

Let Λ be a a locally compact Polish space equipped with its Borel σ-algebra $\mathcal{B}$ and Radon measure ν. A point process X on Λ is a random locally finite subset of Λ. One can also view X as a random counting measure on Λ, having the form $X = \sum_{k \in \mathbb{N}} \delta_{T_k}$, where $\{T_k\}_{k \in \mathbb{N}}$ is a countable collection of points in Λ with no accumulation points.

We now recall some basic definitions related to point processes. A point process is called *simple* if almost surely $X(\{x\}) \leq 1$ for all $x \in \Lambda$. A point process is *stationary* if its distribution is invariant under translations. The *intensity measure* of a point process X is the measure on Λ defined by

$$\alpha(B) = \mathbb{E}[X(B)], \qquad B \in \mathcal{B}(\Lambda).$$

If X is stationary, $\alpha(B) = \rho\nu(B)$, and the constant ρ is called the *intensity* of the point process. The *k-th factorial moment measure* of a point process X is the measure $\alpha^{(k)}$ on $(\mathbb{R}^n)^k$ such that for all non-negative and measurable $f : (\mathbb{R}^n)^k \to \mathbb{R}$,

$$\int f(x_1, ..., x_k)\alpha^{(k)}(x_1, ..., x_k) = \mathbb{E}[\sum_{(x_1,...,x_k)\in X}^{\neq} f(x_1, ..., x_k)].$$

In particular, $\alpha^{(k)}(B_1 \times ... \times B_k) = \mathbb{E}[\prod_{i=1}^{k} X(B_i)]$, for $B_1, ..., B_k$ disjoint. If it exists, the density of $\alpha^{(k)}$ with respect to ν is called the k-th product density, $\rho^{(k)}$.

The most commonly studied point process is the Poisson point process. For this model, all points are stochastically independent, and the number of points in a bounded set follows a Poisson distribution. The formal definition is as follows.

Definition 2.1.1. A point process X on a Polish space Λ is Poisson with intensity measure α if for all disjoint subsets $B_1, \ldots, B_k \in \mathcal{B}(\Lambda)$ such that $\alpha(B_i) < \infty$ for all i,

$$\mathbb{P}(X(B_1) = m_1, ..., X(B_k) = m_k) = \prod_{i=1}^{k} \frac{\alpha(B_i)^{m_i}}{m_i!} e^{-\alpha(B_i)}.$$

If the measure α has a density $\rho(\cdot)$, then ρ is called the intensity function of the point process. In particular, the k-th factorial moment measure for a Poisson point process with intensity function $\rho(x)$ has density given by $\rho^{(k)}(x_1, \ldots, x_k) = \prod_{i=1}^{k} \rho(x_i)$ for all $k = 1, 2, \ldots$.

There are many non-Poisson point processes used to model random subsets with dependence between points. A large class of attractive point processes are called Cox processes. These models are also known as doubly stochastic Poisson processes because they consist of a Poisson point process with a random intensity measure, i.e. letting ξ be a random measure, a point process X is a Cox process if, conditional on ξ, X is a Poisson point process with intensity measure ξ. A large class of repulsive point processes are Gibbs point processes [21], which include hardcore and pairwise interaction models mentioned in the introduction. These are point processes that are absolutely continuous with respect to a Poisson point process. Determinantal point processes (DPPs) are another example of a repulsive model, and it was shown in [32] that DPPs are also Gibbs point processes. However, unlike many Gibbs point processes, they have a closed form for their moment measures and a relatively easy simulation procedure.

2.2 Palm Theory

Let X be a point process on Λ. For $u \in \Lambda$, the Palm distribution $\mathbb{P}^u$ of X can be interpreted as the conditional distribution of X given there is a point of X at location u. For a formal definition of this measure for a general point process X, see [26]. For a stationary point process on $\Lambda = \mathbb{R}^n$, without loss of generality we consider the Palm measure at the origin, which we call the 'typical point', and it can be formally defined as follows.

Let $(\Omega, \mathcal{A}, \{\theta_t\}_{t\in\mathbb{R}^n}, \mathbb{P})$ be a stationary framework and X a point process compatible with the flow $\{\theta_t\}_{t\in\mathbb{R}^n}$, implying X is stationary. Let ρ be the intensity of X. The Palm measure $\mathbb{P}^0$ associated with X is defined on $(\Omega, \mathcal{A})$

by

$$\mathbb{P}^0(A) := \frac{1}{\rho}\mathbb{E}\left[\int_B 1_A \circ \theta_t X(dt)\right],$$

for any bounded Borel set B with volume one. There is also the following ergodic interpretation of the Palm probability. If X is stationary and ergodic, then by Birkhoff's Pointwise Ergodic theorem, for all convex averaging sequences $\{K_m\}_{m\geq 1}$ in $\mathbb{R}^n$, and all $f : \Omega \to \mathbb{R}_+$ measurable and in $L_1(\mathbb{P}^0_N)$,

$$\frac{1}{V_n(K_m)}\int_{K_m} f \circ \theta_t X(dt) \to \lambda \mathbb{E}^0[f], \text{ as } m \to \infty, \qquad \mathbb{P} - a.s.$$

Thus, we can think of the Palm probability as the empirical average over all the points in a very large ball. The reduced Palm probability measure of X, denoted $\mathbb{P}^{0,!}$, is the Palm distribution with the point at 0 removed. An important result known as Slivnyak's theorem says that a Poisson point process has the same distribution as its reduced Palm distribution, i.e. $\mathbb{P}^{0,!} = \mathbb{P}$.

Finally, we recall that the nearest neighbor function of a stationary point process X in $\mathbb{R}^n$ is defined as

$$D(r) := \mathbb{P}^{0,!}(X(B_n(r)) > 0). \tag{2.1}$$

This gives the distribution of the distance to the nearest point to the typical point. If X is Poisson, Slivnyak's theorem implies that $D(r) = 1 - e^{-\mathbb{E}X(B_n(r))}$.

2.3 Poisson Tessellations

One of the stochastic geometric models studied here is the random mosaic, or tessellation. We first need the following notation. Denote by $\mathcal{F}, \mathcal{C}$

the sets of closed and convex subsets of $\mathbb{R}^n$, respectively. For $A \subset \mathbb{R}^n$, define

$$\mathcal{F}^A := \{F \in \mathcal{F} : F \cap A = \emptyset\} \text{ and } \mathcal{F}_A := \{F \in \mathcal{F} : F \cap A \neq \emptyset\}. \tag{2.2}$$

The σ-algebra $\mathcal{B}(\mathcal{F})$ of Borel sets of $\mathcal{F}$ is generated by either of the systems $\{\mathcal{F}_C : C \in \mathcal{C}\}$ and $\{\mathcal{F}^C : C \in \mathcal{C}\}$ (see Lemma 2.1.1 in [75]). Let $\mathcal{F}'$ and $\mathcal{C}'$ denote the sets of non-empty closed and compact sets in $\mathbb{R}^n$, respectively. Also, let $\mathcal{K}$ denote the set of convex bodies (non-empty compact convex sets).

A particle process is a point process in $\mathcal{C}'$. A mosaic, or tessellation, is defined to be a collection of convex polytopes in $\mathbb{R}^n$ such that the union is the entire space and no two polytopes in the collection share interior points. Letting $\mathbb{M}$ denote the set of all face-to-face mosaics (see [75]), a random mosaic in $\mathbb{R}^n$ is defined to be a particle process in $\mathbb{R}^n$ such that $X \in \mathbb{M}$ almost surely. The polytopes contained in the mosaic will be called the cells of the mosaic.

The intensity measure of a stationary particle process can be decomposed in the following way. Let $c : \mathcal{C}' \to \mathbb{R}^n$ be a *center function*, defined as a measurable map which is compatible with translations, i.e., $c(C+x) = c(C)+x$ for all $x \in \mathbb{R}^n$. Define the grain space

$$\mathcal{C}_0 := \{C \in \mathcal{C}' : c(C) = 0\},$$

and the homeomorphism (see [75] for more details)

$$\Phi : \mathbb{R}^n \times \mathcal{C}_0 \to \mathcal{C}'; \qquad (x, C) \to x + C.$$

Theorem 2.3.1. *(Theorem 4.1.1 in [75]) Let X be a stationary particle process in $\mathbb{R}^n$ with intensity measure $\Theta \neq 0$. Then there exist a number $\lambda \in (0, \infty)$*

and a probability measure $\mathbb{Q}$ *on* $\mathcal{C}_0$ *such that*

$$\Theta = \lambda\Phi(\nu \otimes \mathbb{Q}).$$

The number λ is called the intensity of the particle process and $\mathbb{Q}$ is called the grain distribution. The point process of centers of the cells of X is a stationary point process with intensity λ, and so λ will also be referred to as the cell intensity.

2.3.1 Zero cell and Typical cell

An important cell of the mosaic is the zero cell, defined as the cell that the origin is contained in. It will be denoted Z_0. Since larger cells are more likely to contain the origin, the zero cell is not a good measure of the average or "typical" cell. For a stationary random mosaic X with grain distribution $\mathbb{Q}$, a random set with distribution $\mathbb{Q}$ is called the typical cell of X. It can also be thought of as the zero cell of the tessellation under the Palm measure of the point process of cell centers. That is, its distribution is that of the cell containing the origin, conditioned on a cell of the tessellation having its center at the origin. This more accurately represents the average distribution of the cells in the random mosaic. Formally, we define this distribution as follows.

Definition 2.3.1. The typical cell Z of a random mosaic X with intensity λ is the random polytope with the following distribution. For all Borel sets $\mathcal{A} \in \mathcal{B}(\mathcal{K})$,

$$\mathbb{Q}(\mathcal{A}) = \frac{1}{\lambda|B|}\mathbb{E}\sum_{P\in X} 1_{\mathcal{A}}\{P - c(P)\}1_B(c(P)),$$

where $B \in \mathcal{B}(\mathbb{R}^n)$ is an arbitrary bounded Borel set. Also, in the case where X is ergodic, this distribution has the following interpretation:

$$\mathbb{Q}(\mathcal{A}) = \lim_{r\to\infty} \frac{1_{\mathcal{A}}\{P - c(P)\}1_{r[-1/2,1/2]^n}(c(P))}{\sum_{P\in X} 1_{r[-1/2,1/2]}(c(P))}, \qquad a.s.$$

By the ergodic interpretation of the distribution, we can think about the typical cell as taking a large compact set, picking a cell uniformly at random and translating it is some appropriate way so that it contains the origin.

It is known that (see, e.g., [75, (10.4) and (10.46)]), that the expected volume of the typical cell is given by the reciprocal of the cell intensity, i.e.,

$$\mathbb{E}[V(Z)] = \int V(K)\mathbb{Q}(K) = \frac{1}{\lambda}.$$

The following result provides an important relationship between the distribution of the zero cell and the typical cell of a stationary random mosaic, i.e. that the distribution of $Z_0 - c(Z_0)$ has a Radon-Nikodym derivative with respect to the distribution of Z given by $V(\cdot)/\mathbb{E}[V(Z)]$.

Theorem 2.3.2. *(Theorem 10.4.1 in [75]) Let X be a stationary random mosaic in $\mathbb{R}^n$. Denote its typical cell by Z and zero cell by Z_0. For any non-negative measurable and translation-invariant function $g : \mathcal{K} \to \mathbb{R}$,*

$$\mathbb{E}[g(Z_0)] = \frac{1}{\mathbb{E}[V(Z)]}\mathbb{E}[g(Z)V(Z)].$$

2.3.2 Poisson-Voronoi Mosaic

A special type of random mosaic comes from the Voronoi cells of a Poisson point process in $\mathbb{R}^n$. Let N be a stationary Poisson point process with intensity λ and for $x \in N$, define the Voronoi cell of N with center x by

$$C(x, N) := \{z \in R^n : \|z - x\| \leq \|z - y\| \text{ for all } y \in Z\}.$$

The collection $X := \{C(x, N) : x \in N\}$ is a stationary random mosaic and is called the Poisson-Voronoi mosaic induced by N. The intensity λ of the underlying Poisson point process is the cell intensity of the induced mosaic.

2.3.3 Poisson Hyperplane Mosaic

The second type of random mosaic we consider is the mosaic induced by a stationary and isotropic Poisson hyperplane process X in $\mathbb{R}^n$. Denote the set of $n - 1$ dimensional hyperplanes in $\mathbb{R}^n$ by $\mathcal{H}^n$ and the Grassmanian of $n-1$-dimensional linear subspaces of $\mathbb{R}^n$ by $G(n, n-1)$. The set $G(n, n-1)$ is the subset of hyperplanes in $\mathcal{H}^n$ that pass through the origin. A hyperplane process in $\mathbb{R}^n$ is a point process in the space $\mathcal{H}^n$.

The following theorem (see, e.g., [75]) provides a decomposition for the intensity measure for all stationary hyperplane processes. Note that elements of the space $\mathcal{H}^n$ are of the form

$$H(u, \tau) := \{x \in \mathbb{R}^n : \langle x, u \rangle = \tau\}, \tag{2.3}$$

where $u \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$.

Theorem 2.3.3. *(Theorem 4.4.2 and (4.33) in [75]) Let X be a stationary hyperplane process in $\mathbb{R}^n$ with intensity measure $\Theta \neq 0$. Then, there is a unique number $\gamma \in (0, \infty)$ and probability measure $\mathbb{Q}$ on $G(n, n-1)$ such that* *for all nonnegative measurable functions f on $\mathcal{H}^n$,*

$$\int_{\mathcal{H}^n} f d\Theta = 2\gamma \int_{S^{n-1}} \int_0^\infty f(H(u,\tau)) d\tau \phi(du),$$

where for $A \in \mathcal{B}(S^{n-1})$, $\phi(A) := \frac{1}{2}\mathbb{Q}(\{u^\perp : u \in A\})$. ϕ is called the spherical directional distribution. In particular, for $A \in \mathcal{B}(\mathcal{H}^n)$,

$$\Theta(A) = 2\gamma \int_{S^{n-1}} \int_0^\infty 1_{\{H(u,\tau)\in A\}} d\tau \phi(du).$$

The parameter γ is called the intensity and $\mathbb{Q}$ the directional distribution of X. If X is isotropic, i.e., if its distribution is invariant under rotations about the origin, then $\mathbb{Q}$ is rotationally invariant and thus is the Haar measure ν_{n-1} and $\phi = \sigma_{n-1}$, the normalized spherical Lebesgue measure on S^{n-1}.

The relation of the intensity γ to the cell intensity λ of the induced random mosaic is given by

$$\lambda = \kappa_n \left(\frac{\gamma\kappa_{n-1}}{n\kappa_n}\right)^n, \tag{2.4}$$

where κ_n is the volume of the unit ball in $\mathbb{R}^n$.

If the hyperplane process is stationary, then the induced random mosaic is stationary, and one can examine the typical cell of the tessellation. There are a number of metrics one can use to understand the size and shape of the typical cell. In the following, we describe some metrics for which the distribution is known in the case of a stationary Poisson hyperplane tessellation.

First, letting Z denote the typical cell of X, (2.4) implies that

$$\mathbb{E}[V(Z)] = \int V(K)\mathbb{Q}(K) = \frac{1}{\lambda} = \frac{1}{\kappa_n}\left(\frac{n\kappa_n}{\gamma\kappa_{n-1}}\right)^n. \tag{2.5}$$

The inradius r_{in} of a cell is the radius of the largest ball completely contained in the cell. The following result gives the distribution of the inradius of the typical cell of a stationary Poisson hyperplane process.

Theorem 2.3.4. *(Theorem 10.4.8 in [75]) Let X be a nondegenerate stationary Poisson hyperplane process in $\mathbb{R}^n$ with intensity γ. Let Z be the typical cell. Then,*

$$\mathbb{P}(r_{in}(Z) \leq a) = 1 - e^{-2\gamma a}, \qquad a \geq 0.$$

If we define the center function $c(C)$ to be the center of the largest ball included in the set cell C, Calka [15] showed that the distribution of the typical cell can be described in the following way. Let $R \in \mathbb{R}^+$ and $(U_0, ..., U_n) \in (\mathbb{S}^{d-1})^{(d+1)}$ be independent random variables such that R is exponentially distributed with parameter 2γ, as $r_{in}(Z)$ is in the above theorem, and $(U_0, ..., U_n)$ has density with respect to the uniform measure which is proportional to the volume of the simplex constructed with these $n+1$ vectors multiplied by the indicator that this simplex contains the origin. Then, let X_R be the hyperplane process X restricted to $\mathbb{R}^n \backslash B_r(0)$. Letting $\mathcal{C}_1$ be the polyhedron containing the origin obtained as the intersection of the $(n+1)$ half spaces bounded by the hyperplanes H_{RU_i}, $0 \leq i \leq n$ and $\mathcal{C}_2$ be the zero-cell of the hyperplane tessellation associated with X_R, the typical cell of the stationary and isotropic Poisson hyperplane tessellation is distributed as $\mathcal{C}_1 \cap \mathcal{C}_2$. In other words,

Theorem 2.3.5. *(Theorem 10.4.6 in [75]) Let X be a stationary and isotropic Poisson hyperplane process in $\mathbb{R}^n$ with intensity γ. If $\mathbb{Q}$ is the probability distribution of the typical cell Z with respect to the inball center as the center function, then for all borel sets $A \in \mathcal{B}(\mathcal{K})$,*

$$
\begin{aligned}
&\mathbb{Q}(A) \\
&= \frac{\mathbb{E}[V(Z)]\gamma^{n+1}}{(n+1)} \int_0^\infty \int_{(S^{n-1})^{n+1}} e^{-2\gamma r} \mathbb{P}\left(\bigcap_{H \in X \cap \mathcal{F}_{B^0(0,r)}} H_0^+ \cap \bigcap_{j=0}^{n} H^-(u_j, r) \in A \right) \\
&\quad \cdot \triangle_n(u_0, ..., u_n) 1_P(u_0, ..., u_n) \sigma_{n-1}(du_0) ... \sigma_{n-1}(du_n) dr.
\end{aligned}
$$

The triangle notation is the n-dimensional volume of the convex hull of the vectors, i.e.

$$\triangle_n(u_0, ..., u_n) = \frac{1}{n!} \nabla_n(u_1 - u_0, u_2 - u_0, ..., u_n - u_n),$$

where $\nabla_n(v_1, v_2, ..., v_n)$ is the volume of parallelpiped spanned by the vectors $v_1, \ldots, v_n$. The set $P \subset (S^{n-1})^{n+1}$ is the set of all $(n+1)$-tuples of unit vectors such that the origin is contained in their convex hull.

2.4 High Dimensional Space

This thesis will discuss the above random geometric models in high dimensional Euclidean space. The following notations and asymptotic formulas will be used throughout. Let $B_n(r)$ denote the ball or radius r centered at the origin in $\mathbb{R}^n$. The usual ℓ^2 norm of a vector is denoted by $|\cdot|$, and the L^2-norm on the space $L^2(\mathbb{R}^n)$ by $\|\cdot\|_2$. The $n-$dimensional volume of a set $K \subset \mathbb{R}^n$

is denoted by $V_n(K)$. The volume of the $n-$dimensional unit ball $B_n(1)$ is denoted by κ_n and the surface area of the $n-$dimensional unit sphere S^{n-1} is denoted by ω_n. They satisfy

$$\kappa_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}, \qquad \omega_n = n\kappa_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}.$$

Also recall the following special functions. The gamma function is defined as

$$\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}dt,$$

and the upper and lower regularized incomplete gamma functions are defined for all $R \geq 0$ by

$$\Gamma_u(x,R) := \frac{\int_R^\infty t^{x-1}e^{-t}dt}{\Gamma(x)}, \qquad \Gamma_\ell(x,R) := \frac{\int_0^R t^{x-1}e^{-t}dt}{\Gamma(x)},$$

respectively. Stirling's formula gives the following asymptotic expansion as $x \to \infty$:

$$\Gamma(x+1) \sim \sqrt{2\pi x}\left(\frac{x}{e}\right)^x. \tag{2.6}$$

The following asymptotic formulas will also be used: by (2.6), as $n \to \infty$,

$$\kappa_n \sim \frac{1}{\sqrt{n\pi}}\left(\frac{2\pi e}{n}\right)^{n/2} \qquad \text{and} \qquad \frac{\kappa_{n-1}}{n\kappa_n} \sim \frac{1}{\sqrt{2\pi n}}. \tag{2.7}$$

2.4.1 Log-concavity and Thin-shell estimate

For general random vectors Y_n in $\mathbb{R}^n$, the concentration of $|Y_n|$ for large n has been well-studied (see [31], [38], [47]). Indeed, in [31, Proposition 3], it is proved that Y_n is concentrated in a "thin shell", i.e., there exists a sequence $\{\varepsilon_n\}$ such that $\varepsilon_n \to 0$ as $n \to \infty$ and for each n,

$$\mathbb{P}\left(\left|\frac{|Y_n|}{\mathbb{E}[|Y_n|^2]^{\frac{1}{2}}} - 1\right| \geq \varepsilon_n\right) \leq \varepsilon_n, \tag{2.8}$$

if and only if $|Y_n|$ has a finite rth moment for $r > 2$, and for some $2 < p < r$,

$$\left|\frac{\mathbb{E}[|Y_n|^p]^{1/p}}{\mathbb{E}[|Y_n|^2]^{1/2}} - 1\right| \to 0 \text{ as } n \to \infty.$$

More can be said about the concentration under the assumption that the vector has a log-concave density. A function $f : \mathbb{R}^n \to \mathbb{R}$ is log-concave if its domain is a convex set and is $\log f$ is concave, i.e., if for all $s \in (0,1)$ and x, y in the domain of f,

$$\log f(sx + (1-s)y) \geq s \log f(x) + (1-s) \log f(y).$$

It is known that for a random vector Y with density $f(x) := g(|x|)$, where $g : \mathbb{R} \to \mathbb{R}$ is log-concave, the following thin-shell estimate holds:

$$\mathbb{E}\left(\frac{|Y|}{(\mathbb{E}|Y|^2)^{\frac{1}{2}}} - 1\right)^2 \leq \frac{C}{n}. \tag{2.9}$$

The best known estimate applies to general log-concave random vectors (not necessarily radial) and is given by the following theorem in [38].

Theorem 2.4.1. *(Guédon and Milman [38]) Let Y denote a random vector in $\mathbb{R}^n$ such that $\mathbb{E}Y = 0$ and $\mathbb{E}(Y \otimes Y) = I_n$. Assume Y has a log-concave density. Then, for some $C > 0$ and $c > 0$,*

$$\mathbb{P}\left(\left|\frac{|Y|}{\sqrt{n}} - 1\right| \geq t\right) \leq C \exp\left(-c\sqrt{n} \min(t^3, t)\right).$$

We refer to the monograph [14] for a summary of these and more related results.

Chapter 3

Measure of Repulsiveness of DPPs[1]

3.1 Introduction

Determinantal point processes (DPPs) have been of much interest over the last many years in mathematical physics and probability theory (see e.g. [12, 42, 57, 77, 79] and the references therein) and more recently in other areas, including statistics [52, 62], machine learning [49], signal processing [28], and neuroscience [78]. They are models for regularity/inhibition/repulsiveness, but there is a trade-off between repulsiveness and intensity [51, 52]. The results in this chapter shed further light on this issue by studying various couplings between a DPP and its reduced Palm distributions.

Section 3.2.1 provides our general setting for a DPP X defined on a locally compact Polish space Λ and specified by a so-called kernel $K : \Lambda \times \Lambda \to \mathbb{C}$ which satisfies certain mild conditions given in Section 3.2.2. Also, for any $u \in \Lambda$ with $K(u, u) > 0$, if X^u follows the reduced Palm distribution of X at u – intuitively, this is the conditional distribution of $X \setminus \{u\}$ given that $u \in X$ –

[1]This chapter is based on the following manuscript: *J. Møller and E. O'Reilly. Couplings for determinantal point processes and their reduced Palm distribution with a view to quantifying repulsiveness. arXiv:1806.07347, June 2018.* The author of this thesis performed substantial research that formed the results in this manuscript.

then X^u is another DPP; Section 3.2.3 provides further details. Furthermore, Section 3.2.4 discusses Goldman's [35] result that if for any compact set $S \subseteq \Lambda$, denoting K_S the restriction of K to $S \times S$, we have that the spectrum of K_S is < 1, then X stochastically dominates X^u and hence by Strassen's theorem there exists a coupling so that almost surely $X^u \subseteq X$. The difference $\kappa_u := X \setminus X^u$ is a finite point process with a known intensity function. In particular, for a standard Ginibre point process [33], which is a special case of a DPP in the complex plane, Goldman showed that κ_u consists of a single point which follows $N_{\mathbb{C}}(u, 1)$, the complex Gaussian distribution with mean u and unit variance. However, apart from this and other special cases, the distribution of κ_u is unknown.

Section 3.3 shows that more can be said: Under weaker conditions than in Goldman's paper, there is a coupling so that almost surely $X^u \subseteq X$, $\xi_u := X \setminus X^u$ consists of at most one point, and the distribution of ξ_u can be specified. Note that κ_u and ξ_u share the same intensity function. As in [35] we only verify the existence of our coupling result. We leave it as an open research problem to provide a specific coupling construction or simulation procedure for (X, X^u) (restricted to a compact subset of Λ); possibly this may provide a faster simulation algorithm than in [51, 52, 62].

Section 3.4 discusses how our coupling result can be used for describing the repulsiveness in a DPP. In particular, if for all $u \in \Lambda$ with $K(u, u) > 0$, almost surely ξ_u has one point, we call X a most repulsive DPP; we discuss this definition in connection to most repulsive stationary DPPs on $\mathbb{R}^d$ as specified in [52, 9]. For example, if X is a standard Ginibre point process, we obtain a similar result as in [35]: X is a most repulsive DPP and the point in ξ_u follows $N_{\mathbb{C}}(u, 1)$. Moreover, we consider the cases of a finite set Λ and when we have a stationary DPP defined on $\Lambda = \mathbb{R}^d$. Finally, we compare with most repulsive isotropic DPPs on $\mathbb{S}^d$, the d-dimensional unit sphere in $\mathbb{R}^{d+1}$, as studied in [60].

3.2 Background

Below we give the definition of a DPP, specify our assumptions, and recall that the reduced Palm distribution of a DPP is another DPP. We also describe a previous result on a coupling between a DPP and its reduced Palm version.

3.2.1 Definition of a DPP

Let X be a point process defined on a locally compact Polish space Λ equipped with its Borel σ-algebra $\mathcal{B}$ and a Radon measure ν which is used as a reference measure in the following. We assume that X is a DPP with kernel K which by definition means the following. First, X has no multiple points, so dependent on the context we view X as a random subset of Λ or as a random counting measure, and we let $X(B)$ denote the cardinality of $X_B := X \cap B$ for $B \in \mathcal{B}$. Second, K is a complex function defined on $K : \Lambda^2 \mapsto \mathbb{C}$. Third, for any $n = 1, 2, \ldots$ and any mutually disjoint bounded sets $B_1, \ldots, B_n \in \mathcal{B}$,

$$\mathrm{E}\left[X\left(B_1\right) \cdots X\left(B_n\right)\right] = \int_{B_1 \times \cdots \times B_n} \det\left\{K\left(u_i, u_j\right)\right\}_{i,j=1}^{n} \mathrm{d}\nu^n\left(u_1, \ldots, u_n\right)$$

is finite, where ν^n denotes the n-fold product measure of ν. This means that X has n-th order intensity function $\rho(u_1, \ldots, u_n)$ (also sometimes in the literature called n-th order correlation function) given by the determinant

$$\rho\left(u_1, \ldots, u_n\right) = \det\left\{K\left(u_i, u_j\right)\right\}_{i,j=1}^{n}, \qquad u_1, \ldots, u_n \in \Lambda, \tag{3.1}$$

and this function is locally integrable. In particular, $\rho(u) = K(u, u)$ is the intensity function of X, and when $B \in \mathcal{B}$ is bounded almost surely X_B is finite.

In the special case where $K(u, v) = 0$ whenever $u \neq v$, the DPP X is just a Poisson process with intensity function $\rho(u)$ conditioned on that there are no multiple points in X (if ν is diffuse, it is implicit that there are no multiple points). For other examples when Λ is a countable set and ν is the counting measure, see [49]; when $\Lambda = \mathbb{R}^d$ and ν is the Lebesgue measure, see [42, 52]; and when $\Lambda = \mathbb{S}^d$ (the d-dimensional unit sphere) and ν is the surface/Lebesgue measure, see [60]. Examples are also given in Section 3.4.2.

From (3.1) and the fact that the determinant of a complex covariance matrix is less than or equal to the product of its diagonal elements we obtain that

$$\rho\left(u_1, \ldots, u_n\right) \leq \prod_{i=1}^{n} \rho\left(u_i\right),$$

where the equality holds if and only if X is a Poisson process. Thus, apart from the case of a Poisson process, the counts $X(A)$ and $X(B)$ are negatively correlated whenever $A, B \in \mathcal{B}$ are disjoint.

3.2.2 Assumptions

We always make the following assumptions (a)–(c):

(a) K is Hermitian, that is, $K(u,v) = \overline{K(v,u)}$ for all $u, v \in \Lambda$;

(b) K is locally square integrable, that is, for any compact set $S \subseteq \Lambda$, the double integral $\int_S \int_S |K(u,v)|^2 \,\mathrm{d}\nu(u)\,\mathrm{d}\nu(v)$ is finite;

(c) K is of locally trace class, that is, for any compact set $S \subseteq \Lambda$, the integral $\int_S K(u,u)\,\mathrm{d}\nu(u)$ is finite.

By Mercer's theorem, excluding a ν^2-nullset, this ensures the existence of a spectral representation for the kernel restricted to any compact set $S \subseteq \Lambda$: Ignoring a ν^2-nullset, we can redefine K on $S \times S$ by

$$K(u,v) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(u) \overline{\phi_k^S(v)} \qquad u, v \in S, \tag{3.2}$$

where the eigenvalues λ_k^S are real numbers and the eigenfunctions ϕ_k^S constitute an orthonormal basis of $L^2(S)$, cf. Section 4.2.1 in [42]. Here, for any $B \in \mathcal{B}$, $L^2(B) = L^2(B, \nu)$ is the space of square integrable complex functions w.r.t. ν restricted to B. Note that (c) means $\mathrm{E}X(S) = \sum_{k=1}^{\infty} \lambda_k^S < \infty$. Thus, when $B \in \mathcal{B}$ is bounded, almost surely X_B is finite. When ν is diffuse, as we are redefining K by (3.2) we have effectively excluded the special case of the Poisson process (i.e. when K is 0 off the diagonal) because all the eigenvalues in (3.2) are then 0; however, as shown later, it will still make sense to consider the Poisson process when quantifying repulsiveness in DPPs.

We also always assume that

(d) for any compact set $S \subseteq \Lambda$, all eigenvalues satisfy $0 \leq \lambda_k^S \leq 1$.

In fact, under (a)–(c), the existence of the DPP with kernel K is equivalent to (d) (see e.g. Theorem 4.5.5 in [42]), and the DPP is then unique (Lemma 4.2.6 in [42]). If $\Lambda = \mathbb{R}^d$, ν is the Lebesgue measure, and $K(u, v) = K_0(u - v)$ is stationary, where $K_0 \in L^2(\mathbb{R}^d)$ and K_0 is continuous, we denote the Fourier transform of K_0 by $\hat{K}_0$. Then (d) is equivalent to $0 \leq \hat{K}_0 \leq 1$ (Proposition 3.1 in [42]).

Recalling that K_S is the restriction of K to $S \times S$, we sometimes consider one of the following conditions:

(e) For a given compact set $S \subseteq \Lambda$, K_S is a projection of finite rank n.

(f) For all compact $S \subseteq \Lambda$, all eigenvalues satisfy that $\lambda_k^S < 1$.

3.2.3 Reduced Palm distributions

Consider an arbitrary point $u \in \Lambda$ with $\rho(u) > 0$. Recall that the reduced Palm distribution of X at u is a point process X^u on Λ with n-th order intensity function

$$\rho^u(u_1, \ldots, u_n) = \rho(u, u_1, \ldots, u_n)/\rho(u).$$

This combined with (3.1) easily shows that X^u is a DPP with kernel

$$K^u(v, w) = K(v, w) - \frac{K(v, u)K(u, w)}{K(u, u)} \qquad v, w \in \Lambda, \tag{3.3}$$

see Theorem 6.5 in [77]. For any compact set $S \subseteq \Lambda$, it follows that the restriction $X_S^u := X^u \cap S$ follows the reduced Palm distribution of X_S at u.

3.2.4 Goldman's results

Goldman [35] made similar assumptions as in our assumptions (a)-(d), and in addition he assumed condition (f) throughout his paper. Two of his main results were the following.

(G1) For any $u \in \Lambda$ with $K(u,u) > 0$, there is a coupling of X and X^u so that almost surely $X^u \subseteq X$.

(G2) Suppose X is a standard Ginibre point process, that is, the DPP on $\Lambda = \mathbb{C} \equiv \mathbb{R}^2$, with ν being Lebesgue measure, and with kernel

$$K(v,w) = \frac{1}{\pi} \exp\left(v\overline{w} - \frac{|v|^2 + |w|^2}{2} \right), \qquad v, w \in \mathbb{C}. \tag{3.4}$$

Then, for the coupling in (G1) and any $u \in \mathbb{C}$, $X \setminus X^u$ consists of a single point which follows $N_{\mathbb{C}}(u,1)$.

It follows from (G1) and (3.3) that $\kappa_u := X \setminus X^u$ is a finite point process with intensity function

$$\rho_{\kappa_u}(v) = |K(u,v)|^2 / K(u,u), \qquad v \in \Lambda. \tag{3.5}$$

Note that the standard Ginibre point process is stationary and isotropic with intensity $1/\pi$, but its kernel is only isotropic. In accordance with (G2), combining (3.4) and (3.5), ρ_{κ_u} is immediately seen to be the density of $N_{\mathbb{C}}(u,1)$.

3.3 Main result

The theorem below is our main result which is sharpening Goldman's result (G1) in two ways: We do not assume condition (f) and we establish a coupling so that X contains X^u, the difference is at most one point, and we can completely describe the distribution of this difference. In the proof of the theorem we use basic results and definitions for operators on the Hilbert space $\mathcal{L}^2(\Lambda)$, see e.g. [65, 68]. An outline of the proof is as follows. First, we dilate the operator associated to the DPP X to a projection operator on the union of two copies of Λ. Second, we use the existence of a coupling for projection operators in Lemma 3.3.1. Finally, we compress back down to Λ to obtain the desired coupling.

We use the following special result established under condition (e) and where ν_S denotes the restriction of the reference measure ν to a compact set $S \subseteq \Lambda$.

Lemma 3.3.1. *Assume $S \subseteq \Lambda$ is compact and let $\{\phi_k^S\}_{k=1}^n$ be an orthonormal set of functions in $L^2(S)$ with $1 \le n < \infty$. Let X and Y be DPPs with kernels K and L, respectively, so that*

$$K(v,w) = \sum_{k=1}^{n} \phi_k^S(v)\overline{\phi_k^S(w)}, \qquad L(v,w) = \sum_{k=1}^{n-1} \phi_k^S(v)\overline{\phi_k^S(w)}, \qquad v,w \in S$$

(setting $L(v,w) = 0$ if $n = 1$). Then there exists a monotone coupling of Y_S w.r.t. X_S such that almost surely $Y_S \subset X_S$, $\eta_S := X_S \setminus Y_S$ consists of one point, and the point in η_S has density $|\phi_n^S(\cdot)|^2$ w.r.t. ν_S.

Proof. Observe that K and L are the kernels of finite dimensional projections, a special case of trace-class positive contractions, and the difference,

$$K(v,w) - L(v,w) = \phi_n^S(v)\overline{\phi_n^S(w)}, \qquad v, w \in S,$$

is a positive definite kernel. Thus, by Theorem 3.8 in [56], X_S stochastically dominates Y_S. Therefore, there is a coupling such that almost surely $Y_S \subseteq X_S$. As Y_S has cardinality one less than X_S, almost surely $\eta_S := X_S \setminus Y_S$ consists of one point. Finally, for any Borel set $A \subseteq S$,

$$\mathrm{P}(\eta_S \cap A \neq \emptyset) = \mathrm{E}\left[1_{\{X(A)-Y(A)=1\}}\right] = \mathrm{E}[X(A)] - \mathrm{E}[Y(A)] = \int_A |\phi_n^S(v)|^2 \,\mathrm{d}\nu(v).$$

□

Denote $\|\cdot\|_2$ the usual norm on $\mathcal{L}^2(\Lambda)$ w.r.t. ν.

Theorem 3.3.2. *Let X be a DPP on Λ with kernel K satisfying conditions (a)–(d). For any $u \in \Lambda$ with $K(u,u) > 0$, there exists a coupling of X and X^u such that almost surely $X^u \subseteq X$ and $\xi_u := X \setminus X^u$ consists of at most one point. We have*

$$p_u := \mathrm{P}(\xi_u \neq \emptyset) = \frac{1}{K(u,u)} \int |K(u,v)|^2 \,\mathrm{d}\nu(v), \tag{3.6}$$

and conditioned on $\xi_u \neq \emptyset$ the point in ξ_u has density

$$f_u(v) := |K(u,v)|^2 / \|K(u,\cdot)\|_2^2, \qquad v \in \Lambda, \tag{3.7}$$

w.r.t. ν.

Compared to Goldman's result (G1), we also have $p_u = \mathrm{P}(\kappa_u \neq \emptyset)$ and f_u is the conditional density of a point in κ_u given that $\kappa_u \neq \emptyset$, cf. (3.5)–(3.7).

Proof. Denote $\mathcal{K}$ the locally trace class operator on $\mathcal{L}^2(\Lambda)$ with kernel K. As in section 3.3 in [56], consider the dilation of $\mathcal{K}$ given by

$$\mathcal{Q} := \begin{bmatrix} \mathcal{K} & \mathcal{L} \\ \mathcal{L} & \mathcal{I} - \mathcal{K} \end{bmatrix},$$

where $\mathcal{L} := \sqrt{\mathcal{K}(\mathcal{I} - \mathcal{K})}$. Then, since $Q = Q^2$, Q is an orthogonal projection on $L^2(\Lambda, \nu) \oplus L^2(\Lambda_0, \nu)$, where Λ_0 is a disjoint identical copy of Λ. If Λ is discrete, then Q is clearly locally trace class, since any compact set of a discrete space is finite. If Λ is not discrete, consider the operator

$$\mathcal{Q}' := \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix}^* \mathcal{Q} \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix} = \begin{bmatrix} \mathcal{K} & \mathcal{L}\mathcal{U} \\ \mathcal{U}^*\mathcal{L} & \mathcal{U}^*(\mathcal{I} - \mathcal{K})\mathcal{U} \end{bmatrix},$$

where $\mathcal{U}$ is a unitary operator from $\ell^2(\Lambda_0')$ to $L^2(\Lambda_0, \nu)$ for some countably infinite space Λ_0'. The operator $\mathcal{U}$ exists since any two infinite dimensional separable Hilbert spaces are unitarily equivalent. The operator $\mathcal{Q}'$ is an orthogonal projection on $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda_0')$, and $\mathcal{K}$ is the compression of $\mathcal{Q}'$ to Λ. Further, $\mathcal{Q}'$ is also locally trace class, because $\mathcal{K}$ is locally trace class on $L^2(\Lambda, \nu)$ by assumption, and all operators on $\ell^2(\Lambda_0')$ are locally of trace class since Λ_0' is discrete. Thus, $\mathcal{Q}'$ defines a projection DPP Y_Q on the union $\Lambda \cup \Lambda_0'$.

First, assume that Λ is compact. Then, the kernel of the operator $\mathcal{K}$ satisfies

$$K(v, w) = \sum_{k \geq 1} \lambda_k^\Lambda \phi_k^\Lambda(v) \overline{\phi_k^\Lambda(w)}, \qquad v, w \in \Lambda,$$

where $\{\phi_k^\Lambda\}$ is an orthonormal basis for $L^2(\Lambda)$, $\lambda_k^\Lambda \in [0,1]$ for all k, and $\sum_{k\geq 1} \lambda_k^\Lambda < \infty$. Also, the kernel for the operator $\mathcal{L}$ is then given by

$$L(v,w) = \sum_{k\geq 1} \sqrt{\lambda_k^\Lambda(1-\lambda_k^\Lambda)}\phi_k^\Lambda(v)\overline{\phi_k^\Lambda(w)}.$$

Note that

$$\mathcal{L}(L(\cdot,u))(w) = \int_\Lambda L(w,v)L(v,u)\,\mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda(1-\lambda_k^\Lambda)\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)},$$

and

$$\mathcal{K}(K(\cdot,u))(w) = \int_\Lambda K(w,v)K(v,u)\,\mathrm{d}\nu(v) = \sum_{k\geq 1} \left(\lambda_k^\Lambda\right)^2 \phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)}.$$

Hence, $\mathcal{K}(K(\cdot,u)) + \mathcal{L}(L(\cdot,u)) = K(\cdot,u)$. Also,

$$\mathcal{L}(K(\cdot,u))(w) = \int_\Lambda L(w,v)K(v,u)\,\mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda\sqrt{\lambda_k^\Lambda(1-\lambda_k^\Lambda)}\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)}$$

and

$$\mathcal{K}(L(\cdot,u))(w) = \int_\Lambda K(w,v)L(v,u)\,\mathrm{d}\nu(v) = \sum_{k\geq 1} \lambda_k^\Lambda\sqrt{\lambda_k^\Lambda(1-\lambda_k^\Lambda)}\phi_k^\Lambda(w)\overline{\phi_k^\Lambda(u)},$$

and so $\mathcal{L}(K(\cdot,u)) = \mathcal{K}(L(\cdot,u))$. Consequently, for fixed $u \in \Lambda$,

$$\psi_u(\cdot) := \begin{bmatrix} \frac{K(\cdot,u)}{\sqrt{K(u,u)}} \\ \mathcal{U}^*\left(\frac{L(\cdot,u)}{\sqrt{K(u,u)}}\right) \end{bmatrix}$$

is an eigenvector of the operator $\mathcal{Q}'$. Indeed, since $\mathcal{U}\mathcal{U}^* = \mathcal{I}$ by that fact that

$\mathcal{U}$ is unitary,

$$\mathcal{Q}'(\psi_u(\cdot)) = \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U} \end{bmatrix}^* Q \begin{bmatrix} \frac{K(\cdot,u)}{\sqrt{K(u,u)}} \\ (\mathcal{U}\mathcal{U}^*)\left(\frac{L(\cdot,u)}{\sqrt{K(u,u)}}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{U}^* \end{bmatrix} \begin{bmatrix} \frac{\mathcal{K}(K(\cdot,u))}{\sqrt{K(u,u)}} + \frac{\mathcal{L}(L(\cdot,u))}{\sqrt{K(u,u)}} \\ \frac{\mathcal{L}(K(\cdot,u))}{\sqrt{K(u,u)}} + \frac{(\mathcal{I}-\mathcal{K})(L(\cdot,u))}{\sqrt{K(u,u)}} \end{bmatrix} = \begin{bmatrix} \frac{K(\cdot,u)}{\sqrt{K(u,u)}} \\ \mathcal{U}^*\left(\frac{L(\cdot,u)}{\sqrt{K(u,u)}}\right) \end{bmatrix} = \psi_u(\cdot).$$

Then, we can define the projection

$$\mathcal{Q}'_u := \mathcal{Q}' - P_{\psi_u},$$

where P_{ψ_u} is the projection operator on $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda'_0)$ onto the span of ψ_u. This projection operator is also locally trace class since it is the difference of locally trace class operators. Then we can define the projection DPP Y_Q^u on $\Lambda \cup \Lambda'_0$ associated with $\mathcal{Q}'_u$. If $\mathcal{Q}'$ has finite rank, then $\mathcal{Q}'$ and $\mathcal{Q}'_u$ have corresponding kernels

$$Q' = \sum_{k=0}^{n} q_k q_k^T \quad \text{and} \quad Q'_u = \sum_{k=1}^{n} q_k q_k^T,$$

where $n < \infty$, $\{q_k\}_{k=1}^n$ is an orthonormal set, and $q_0 := \psi_u$. Applying Lemma 3.3.1 then gives the result.

Now, assume $\mathcal{Q}'$ projects onto an infinite dimensional subspace of $L^2(\Lambda, \nu) \oplus \ell^2(\Lambda'_0)$ and let $\{q_k\}_{k=0}^\infty$ be an orthonormal basis for the range of $\mathcal{Q}'$, where $q_0 := \psi_u$. For each positive integer M, define the finite dimensional projection kernels

$$Q'_M = \sum_{k=0}^{M} q_k q_k^T \quad \text{and} \quad Q'_{M,u} = \sum_{k=1}^{M} q_k q_k^T,$$

and let Y_{Q_M} and $Y_{Q_M}^u$ be the corresponding projection DPPs. By Lemma 3.3.1, there is a coupling of Y_{Q_M} and $Y_{Q_M}^u$ such that almost surely $Y_{Q_M} \supset Y_{Q_M}^u$, where $\xi_{Q_M}^u := Y_{Q_M} \backslash Y_{Q_M}^u$ consists of one point which has density $|\psi_u(\cdot)|^2$. By the same argument as in the proof of Lemma 20 in [35], the sequences Y_{Q_M} and $Y_{Q_M}^u$ are tight and converge in distribution to Y_Q and Y_Q^u, respectively, as $M \to \infty$. Also, the sequence $(Y_{Q_M}^u, \xi_{Q_M}^u)_M$ is tight, and thus a subsequence converges in distribution to (Y_Q^u, ξ_Q^u), where ξ_Q^u consists of one point with density $|\psi_u(\cdot)|^2$, and $Y_Q^u \cup \xi_Q^u$ is equal in distribution to Y_Q.

The projection operator P_{ψ_u} has kernel $\psi_u \psi_u^T$ and the compression of P_{ψ_u} to Λ is the integral operator with kernel

$$\frac{K(v,u)K(u,w)}{K(u,u)}.$$

Then, since the compression of $\mathcal{Q}'$ to Λ is the operator $\mathcal{K}$, the compression of $\mathcal{Q}'_u$ to Λ is the integral operator $\mathcal{K}^u$ with kernel

$$K^u(v,w) = K(v,w) - \frac{K(v,u)K(u,w)}{K(u,u)}.$$

This gives that $Y_Q \cap \Lambda$ has the same distribution as X and $Y_Q^u \cap \Lambda$ has the same distribution as X^u. Thus, almost surely

$$X = X^u \cup \xi_u,$$

where $\xi_u := \xi_Q^u \cap \Lambda$ and X^u are disjoint. Therefore, we have a coupling of X and X^u, where almost surely $X^u \subseteq X$ and the difference is at most one point. The probability of $\xi_u \neq \emptyset$ is the probability that ξ_Q^u is in Λ, and the density of

ξ_Q^u restricted to Λ is

$$f_{\xi_Q^u}(v) 1_{\{v\in\Lambda\}} = \frac{|K(v,u)|^2}{K(u,u)}$$

w.r.t. ν. Hence,

$$\mathrm{P}(\xi^u \neq \emptyset) = \mathrm{P}(\xi_Q^u \in \Lambda) = \int \frac{|K(v,u)|^2}{K(u,u)} \, \mathrm{d}\nu(v)$$

and the density of ξ_u conditioned on $\xi_u \neq \emptyset$ is $f_u(v) = |K(v,u)|^2/\|K(\cdot,u)\|_2^2$ w.r.t. ν.

Second, if Λ is not assumed to be compact, consider a sequence of compact sets $S_n \subset \Lambda$ such that $\cup_{n=1}^{\infty} S_n = \Lambda$ and $S_n \subseteq S_{n+1}$ for $n = 1, 2, \ldots$. For each n, using the result above with Λ replaced by S_n, there exists a coupling of $(X_{S_n}, X_{S_n}^u)$, where almost surely $X_{S_n} = X_{S_n}^u \cup \xi_{S_n}^u$, $\xi_{S_n}^u = X_{S_n} \setminus X_{S_n}^u$ consists of at most one point,

$$\mathrm{P}\left(\xi_{S_n}^u \neq \emptyset\right) = \int_{S_n} \frac{|K(v,u)|^2}{K(u,u)} \, \mathrm{d}\nu(v), \tag{3.8}$$

and conditioned on $\xi_{S_n}^u \neq \emptyset$ the density of the point in $\xi_{S_n}^u$ is

$$f_{u,S_n}(v) = |K(v,u)|^2 / \int_{S_n} |K(w,u)|^2 \, \mathrm{d}\nu(w) \tag{3.9}$$

w.r.t. ν_{S_n}. For consistency, let $T_1 = S_1$ and generate a realization $(y_{T_1}, y_{T_1}^u)$ of $(Y_{T_1}, Y_{T_1}^u) := (X_{S_1}, X_{S_1}^u)$, and for $n = 2, 3, \ldots$, let $T_n = S_n \setminus S_{n-1}$ and generate a realization $(y_{T_n}, y_{T_n}^u)$ of $(Y_{T_n}, Y_{T_n}^u)$ which follows the conditional distribution of $(X_{S_n} \setminus S_{n-1}, X_{S_n}^u \setminus S_{n-1})$ given that $(X_{S_n} \cap S_{n-1}, X_{S_n}^u \cap S_{n-1}) = (\cup_{i=1}^{n-1} y_{T_i}, \cup_{i=1}^{n-1} y_{T_i}^u)$. Then (X, X^u) is distributed as $(Y, Y^u) := (\cup_{n=1}^{\infty} Y_{T_n}, \cup_{n=1}^{\infty} Y_{T_n}^u)$, and almost surely, for $n = 2, 3, \ldots$, $Y_{T_{n-1}} \setminus Y_{T_{n-1}}^u \neq \emptyset$ implies that $Y_{T_n} \setminus Y_{T_n}^u =$

$Y_{T_{n+1}} \setminus Y^u_{T_{n+1}} = \ldots = \emptyset$, and so $\xi_u := Y \setminus Y^u$ consists of at most one point. The probability that ξ_u is non-empty is, by (3.8),

$$\mathrm{P}(\xi_u \neq \emptyset) = \lim_{n\to\infty} \int_{S_n} \frac{|K(v,u)|^2}{K(u,u)} \,\mathrm{d}\nu(v)$$

and hence by monotone convergence we obtain (3.6). Finally, (3.7) is obtained in a similar way using (3.9).

$\square$

3.4 Quantifying repulsiveness in DPPs

In this section we quantify how repulsive DPPs can be, using the probability p_u and the density f_u from Theorem 3.3.2 to describe the repulsive effect of a fixed point contained in a DPP. Note that X^u is the point process where there is a 'ghost point' at u that is affecting the remaining points. Using this coupling of X^u and X, it is clear that the repulsive effect of a point at location u is characterized by the difference between X^u and the original DPP X, where there is no repulsion coming from the location u. Further, $\xi_u = X \setminus X^u$ has intensity function

$$\rho_u(v) := |K(v,u)|^2/\|K(\cdot,u)\|_2^2, \qquad v \in \Lambda.$$

This is the intensity function for the points in X 'pushed out' by u under the Palm distribution. It makes also sense to consider ρ_u as the intensity function of $X \setminus X^u$ when ν is diffuse and X is a Poisson process because then $X = X^u$ and $\rho_u(v) = 0$ for $v \neq u$.

3.4.1 A measure of repulsiveness

Setting $0/0 = 0$, recall that the pair correlation function of X is defined by $g(v, w) = \rho(v, w)/(\rho(v)\rho(w))$ for $v, w \in \Lambda$, so it satisfies

$$1 - g(v, w) = |r(v, w)|^2, \qquad v, w \in \Lambda,$$

where $r(v, w) = K(v, w)/\sqrt{K(v, v)K(w, w)}$ is the correlation function obtained from K. Note that

$$\rho_u(v) = \rho(v)(1 - g(u, v)), \qquad v \in \Lambda. \tag{3.10}$$

As a global measure of repulsiveness in X when having a point of X at u, we suggest the probability of $\xi_u \neq \emptyset$, that is,

$$p_u = \int \rho_u(v)\,\mathrm{d}\nu(v) = \int |K(u, v)|^2/K(u, u)\,\mathrm{d}\nu(v).$$

By (3.10), there is a trade-off between intensity and repulsiveness: If p_u is fixed, we cannot both increase ρ and decrease g. Therefore, when using p_u as a measure to compare repulsiveness in two DPPs, they should share the same intensity function ρ. Then small/high values of p_u correspond to small/high degree of repulsiveness. For a stationary DPP X on $\mathbb{R}^d$, apart from a constant (given by the intensity of X), p_u is in agreement with the measure for repulsiveness in DPPs introduced in [52, 51]; see also [9, 5]. Indeed this measure is very specific for DPPs as discussed later in Section 3.4.2.5. Finally, note that when the intensity function ρ is constant, conditioned on $\xi_u \neq \emptyset$, the density $f_u(v) = \rho_u(v)/p_u$ of the removed point ξ_u is a characteristic of the DPP that is not dependent on the intensity function ρ.

If $p_u = 1$ for all $u \in \Lambda$ with $K(u,u) > 0$, we say that X is a globally most repulsive DPP. This is the case if K is a projection, that is, for all $v, w \in \Lambda$,

$$K(v,w) = \int K(v,y)K(y,w)\,\mathrm{d}\nu(y).$$

For short we then say that X is a projection DPP. The standard Ginibre point process given by (3.4) is globally most repulsive, and its kernel is indeed a projection; this follows from a straightforward calculation using that $(v,w) \to \exp(v\overline{w})$ is the reproducing kernel of the Bargmann-Fock space equipped with the standard complex Gaussian measure. At the other end, if ν is diffuse and X is a Poisson process with intensity function ρ, then $p_u = 0$ for all $u \in \Lambda$ with $\rho(u) > 0$, and so X is a globally least repulsive DPP.

If Λ is compact, then it follows from the spectral representation (3.2) and condition (d) that

$$\begin{aligned}\int_S |K(u,v)|^2\,\mathrm{d}\nu(v) &= \sum_k \sum_\ell \lambda_k^S \lambda_l^S \phi_k^S(u)\overline{\phi_\ell^S(u)} \int_S \overline{\phi_k^S(v)}\phi_\ell^S(v)\,\mathrm{d}\nu(v)\\ &= \sum_k \left(\lambda_k^S\right)^2 |\phi_k^S(u)|^2 \le \sum_k \lambda_k^S |\phi_k^S(u)|^2 = K(u,u),\end{aligned}$$

and so

$$p_u = \frac{\sum_k \left(\lambda_k^\Lambda\right)^2 |\phi_k^\Lambda(u)|^2}{\sum_k \lambda_k^\Lambda |\phi_k^\Lambda(u)|^2}. \tag{3.11}$$

Consequently, in this case, projection DPPs are the only globally most repulsive DPPs. Such a process has a fixed number of points which agrees with the rank of the kernel.

3.4.2 Examples

This section shows specific examples of our measure p_u and the distribution of a point in ξ_u.

3.4.2.1 DPPs defined on a finite set

Assume $\Lambda = \{1, \ldots, n\}$ is finite and ν is the counting measure; this is the simplest situation. Then $L^2(\Lambda) \equiv \mathbb{C}^n$, the class of possible kernels for DPPs corresponds to the class of $n \times n$ complex covariance matrices with all eigenvalues ≤ 1, and the eigenfunctions simply correspond to normalized eigenvectors for such matrices. For simplicity we only consider projection DPPs and Poisson processes below, but other examples of DPPs on finite sets include uniform spanning trees (Example 14 in [42]) and finite DPPs converging to the continuous Airy process on the complex plane [44].

The projection DPPs are given by complex projection matrices, ranging between the degenerated cases where $X = \emptyset$ and $X = \Lambda$. For example, consider the projection kernel of rank two given by $K(v, w) = \frac{1}{n} + t_v \overline{t_w}$, where $\sum_{i=1}^{n} t_i = 0$ and $\sum_{i=1}^{n} |t_i|^2 = 1$. For any $u \in \{1, \ldots, n\}$, we have $p_u = 1$ and

$$\rho_u(v) = \frac{|\frac{1}{n} + t_u \overline{t_v}|^2}{\frac{1}{n} + |t_u|^2}, \qquad v \in \{1, \ldots, n\},$$

is a probability mass function. This shows the repulsive effect of having a point of X at u; in particular, $\rho_u(v)$ has a global maximum point at $v = u$.

The kernel of a Poisson process with intensity function $\rho \leq 1$ and conditioned on having no multiple points is given by a diagonal covariance matrix with diagonal entries $\rho(1), \ldots, \rho(n)$. If $\rho(u) > 0$, then $p_u = \rho(u)$. This is a much different result as when we consider a Poisson process X on a space Λ where the reference measure ν is diffuse: If $\rho(u) > 0$, then $p_u = 0$ and almost surely $X = X^u$.

3.4.2.2 Ginibre point processes

From the standard Ginibre point process given by (3.4), other stationary point processes can be obtained. Independently thinning the process with a retention probability $\alpha\beta$, where $\beta > 0$ and $\alpha \in (0, 1/\beta]$, and multiplying each of the retained points by $\sqrt{\beta}$ gives a new stationary DPP with kernel

$$K(v, w) = \frac{\alpha}{\pi} \exp\left(\frac{v\overline{w}}{\beta} - \frac{|v|^2 + |w|^2}{2\beta}\right), \qquad v, w \in \mathbb{C}. \tag{3.12}$$

We have

$$\rho = \alpha/\pi, \qquad p_u = \alpha\beta, \qquad f_u(v) = \frac{\exp\left(-|v-u|^2/\beta\right)}{\pi\beta} \sim N_{\mathbb{C}}(u, \beta). \tag{3.13}$$

The case where $\alpha = 1$ and $0 < \beta \leq 1$ is mentioned in Goldman's paper [35], and the results in (3.13) match those in Remark 24 in [35]. [28] called the DPP with kernel (3.12) the scaled β-Ginibre point process but the bound $\alpha\beta \leq 1$ was not noticed. For any fixed value of $\rho > 0$, as the value of β increases to its maximum $\min\{1, 1/(\pi\rho)\}$, the more repulsive the process becomes, whilst as β decreases to 0, in the limit a Poisson process with intensity ρ is obtained.

3.4.2.3 DPPs on $\mathbb{R}^d$ with a stationary kernel

Suppose $\Lambda = \mathbb{R}^d$, ν is the Lebesgue measure, and $K(u,v) = K_0(u-v)$ is stationary, where $K_0 \in L^2(\mathbb{R}^d)$ and K_0 is continuous. Then it follows from Parseval's identity that $p_u = 1$ if and only if $\hat{K}_0$ is an indicator function whose integral agrees with the intensity of X, cf. Appendix J in [51]. A natural choice for the support of this indicator function is a ball centred at the origin in $\mathbb{R}^d$, and if (as in the standard Ginibre point process) we let the intensity be $1/\pi$, then the globally most repulsive DPP has a stationary and isotropic kernel given by

$$K(v,w) = \int_{|y|^d \leq d\Gamma(d/2)/(2\pi^{1+d/2})} \exp\left(2\pi i (v-w)\cdot y\right)\, \mathrm{d}y, \qquad v,w \in \mathbb{R}^d, \quad (3.14)$$

where $x \cdot y$ denotes the usual inner product for $x, y \in \mathbb{R}^d$ and $|y|$ is the usual Euclidean distance. For instance, for $d = 1$ this kernel is the sinc function and for $d = 2$ it is the jinc-like function

$$K(v,w) = J_1(2|v-w|)/(\pi|v-w|), \tag{3.15}$$

where J_1 is the Bessel function of order one. We straightforwardly obtain the following proposition, where the moments in (3.16) follow from Eq. 10.22.57 in [1].

Proposition 3.4.1. *For the globally most repulsive DPP on $\mathbb{R}^d$ with kernel given by* (3.14) *and for any $u \in \mathbb{C}$, we have that $\rho_u(v) = \pi|K(u,v)|^2$ is a probability density function. In particular, for $d = 2$,*

$$\rho_u(v) = J_1(2|v-u|)^2/\left(\pi|v-u|^2\right), \qquad v \in \mathbb{R}^2,$$

and the moments of $|Z_u - u|$ *with* $Z_u \sim \rho_u$ *satisfy*

$$\mathrm{E}\left(|Z_u - u|^k\right) = \frac{\Gamma(1+k/2)\Gamma(1-k)}{\Gamma(2-k/2)\Gamma(1-k/2)^2}, \qquad k \in (-2,1), \tag{3.16}$$

and are infinite for $k \geq 1$.

For comparison consider a standard Ginibre point process, where we can define Z_u in a similar way as in Proposition 3.4.1. In both cases, $|Z_u - u|$ is independent of $(Z_u - u)/|Z_u - u|$, which is uniformly distributed on the unit circle. However, the distribution of $|Z_u - u|$ is very different in the two cases: For the standard Ginibre point process, $|Z_u - u|^2$ is exponentially distributed and $|Z_u - u|$ has a finite k-th moment for all $k > -2$ given by $\Gamma(1+k/2)/(\pi\rho)^{k/2}$; whilst for the DPP on $\mathbb{R}^2$ with jinc-like kernel (3.15), $|Z_u - u|$ is heavy-tailed and has infinite k-th moments for all $k \geq 1$.

For any DPP X with kernel K and defined on $\mathbb{R}^d$, using independent thinning and scale transformation procedures similar to those in Section 3.4.2.2 (replacing $\sqrt{\beta}$ by $\beta^{1/d}$ when transforming the points in the thinned process), we obtain a new DPP with kernel

$$K_{\text{new}}(v,w) = \alpha K(v/\beta^{1/d}, w/\beta^{1/d}), \qquad v, w \in \mathbb{R}^d,$$

where $\beta \in (0,1]$ and $\alpha \in (0, 1/\beta]$. For instance, if K is the jinc-like kernel for the globally most repulsive DPP given by (3.15), the new DPP satisfies the same equations for its intensity ρ and its probability p_u as in (3.13). Hence, if ρ and β are the same for this new DPP and the scaled β-Ginibre point process, the two DPPs are equally repulsive in terms of p_u. However, the probability density function for the point in ξ_u conditioned on $\xi_u \neq \emptyset$ now becomes

$$f_u(v) = J_1\left(2|v-u|^2/\beta\right) / \left(\pi|v-u|^2/\beta\right). \tag{3.17}$$

The reach of the repulsive effect of the point at u is much different when comparing the densities in (3.13) and (3.17), in particular if β is large.

3.4.2.4 DPPs on $\mathbb{S}^d$ with an isotropic kernel

Suppose $\Lambda = \mathbb{S}^d$ is the d-dimensional unit sphere, ν is the Lebesgue measure, and $K(v,w) = K_0(v \cdot w)$ is isotropic for all $v, w \in \mathbb{S}^d$. Then the DPP with kernel K is isotropic, and $\rho = K_0(1)$ and p_u do not depend on the choice of $u \in \Lambda$. By a classical result of Schoenberg [76] and by Theorem 4.1 in [60], we have the following. The normalized eigenfunctions will be complex spherical harmonic functions, and K_0 will be real and of the form

$$K_0(t) = \rho \sum_{\ell=0}^{\infty} \beta_{\ell,d} \frac{C_\ell^{\left(\frac{d-1}{2}\right)}(t)}{C_\ell^{\left(\frac{d-1}{2}\right)}(1)}, \qquad -1 \leq t \leq 1,$$

where $C_\ell^{\left(\frac{d-1}{2}\right)}$ is a Gegenbauer polynomial of degree ℓ and the sequence $\beta_{0,d}, \beta_{1,d}, \ldots$ is a probability mass function. Further, letting $\sigma_d = \nu(\mathbb{S}^d) = 2\pi^{(d+1)/2}/\Gamma((d+1)/2)$, the eigenvalues of K are

$$\lambda_{\ell,d} = \rho\sigma_d\beta_{\ell,d}/m_{\ell,d}, \qquad \ell = 0, 1, \ldots,$$

with multiplicities

$$m_{0,1}, \qquad m_{\ell,1} = 2, \qquad \ell = 1, 2, \ldots, \qquad \text{if } d = 1,$$

and

$$m_{\ell,d} = \frac{2\ell + d - 1}{d - 1}\frac{(\ell + d - 2)!}{\ell!(d - 2)!}, \qquad \ell = 0, 1, \ldots, \qquad \text{if } d \in \{2, 3, \ldots\}.$$

So the DPP exists if and only if $\rho \le \inf_{\ell:\,\beta_{\ell,d}>0} m_{\ell,d}/(\sigma_d \beta_{\ell,d})$. Now, applying (3.11), we obtain

$$p_u = \rho\sigma_d \sum_{\ell=0}^{\infty} \beta_{\ell,d}^2 / m_{\ell,d}. \tag{3.18}$$

There is a lack of flexible parametric DPP models on the sphere where K_0 is expressible in closed form, see Section 4.3 in [60]. For instance, let $d = 2$ and consider the special case of the multiquadric model given by

$$K_0(t) = \rho\frac{1 - \delta}{\sqrt{1 + \delta^2 - 2\delta t}}, \qquad -1 \le t \le 1,$$

with $\delta \in (0, 1)$ a parameter and $0 < \rho \le 1/(4\pi(1 - \delta))$. Then, as shown in Section 4.3.2 in [60], the sequence

$$\beta_{\ell,2} = (1 - \delta)\delta^\ell, \qquad \ell = 0, 1, \ldots, \tag{3.19}$$

specifies a geometric distribution and

$$\lambda_{\ell,2} = 4\pi\rho\delta^\ell(1 - \delta)/(2\ell + 1) \le \delta^\ell/(2\ell + 1), \qquad \ell = 0, 1, \ldots.$$

As $\delta \to 0$, then $\lambda_{0,2} \to 4\pi\rho$ and $\lambda_{\ell,2} \to 0$ if $\ell \ge 1$, corresponding to the uninteresting case of a DPP with at most one point if $\rho < 1/(4\pi)$ and with exactly one point if $\rho = 1/(4\pi)$. From (3.18) and (3.19) we obtain

$$p_u = 4\pi\rho(1 - \delta)/(1 + \delta) \le 1/(1 + \delta),$$

with this upper bound obtained for the maximal value of $\rho = 1/(4\pi(1-\delta))$. Therefore the DPP with the multiquadric kernel is far from being globally most repulsive unless the expected number of points is very small.

Instead a flexible parametric model for the eigenvalues $\lambda_{\ell,d}$ is suggested in Section 4.3.4 in [60] so that globally most repulsive DPPs as well as Poisson processes are obtained as limiting cases. However, the disadvantage of that model is that we can only numerically calculate ρ and p_u.

3.4.2.5 Remark

The considerations in Sections 3.4.1 and 3.4.2.1-3.4.2.4 are strictly for DPPs. For example, the intensity function of a Gibbs point process can be both smaller and larger than the intensity function of its Palm distribution at a given point; whilst for a DPP, $\rho \geq \rho^u$. Furthermore, as a candidate for a 'globally most repulsive stationary Gibbs point process on $\mathbb{R}^2$', we may consider $Y = L_Z := \{x + Z : x \in L\}$, where L is the vertex set of a regular triangular lattice (the centres of a honeycomb structure) with one lattice point at the origin, and where Z is a uniformly distributed point in the hexagonal region given by the Voronoi cell of the lattice and centred at the origin (in other words, Y may be considered as the limit of a stationary Gibbs hard core process when the packing fraction of hard discs increases to the maximal value ≈ 0.907, see e.g. [29, 58]). However, the reduced Palm process at $u \in \mathbb{R}^2$ will be degenerated and given by $Y^u = L_u \setminus \{u\}$, which is a much different situation as compared to DPPs.

Chapter 4

Reach of Repulsion for Determinantal Point Processes in High Dimensions[1]

4.1 Introduction

Consider a sequence of point processes X_n indexed by dimension, that is, let X_n be a point process in $\mathbb{R}^n$ with constant intensity ρ_n for each n. If $\rho_n = e^{n\rho}$ and $R_n = \sqrt{n}R$, with $\rho \in \mathbb{R}$ and $R > 0$, then Stirling's formula gives

$$\mathrm{V}_n(B_n(R_n)) \sim \frac{1}{\sqrt{n\pi}} \left(\frac{2\pi e}{n}\right)^{\frac{n}{2}} R_n^n, \text{ as } n \to \infty.$$

This implies there exists a threshold $R^* = \frac{1}{\sqrt{2\pi e}e^{\rho}}$ such that as $n \to \infty$,

$$\mathbb{E}[X_n(B_n(R_n))] \sim e^{n(\rho + \frac{1}{2}\log 2\pi e + \log R) + o(n)} \to \begin{cases} 0, & R < R^* \\ \infty, & R > R^*. \end{cases} \tag{4.1}$$

This justifies the interest in this regime where the intensities grow exponentially with dimension and distances grow with the square root of the dimension. This regime also naturally arises in information theory, and following [4], it will be called the Shannon regime. In this chapter, the range and strength at

[1]This chapter is based on the following manuscript: *F. Baccelli and E. O'Reilly. Reach of repulsion for determinantal point processes in high dimensions. Journal of Applied Probability, 55(3), September 2018.* The author of this thesis performed substantial research that formed the results in this publication.

which DPPs asymptotically exhibit repulsion between points in this regime is quantified.

Mention of these issues appear in [80], where the authors characterize a certain class of DPPs by an effective “hard-core” diameter D that grows like $\sqrt{n}$, aligning with our observations. They observe that for $r < D$, the number of points in a ball of radius r around a typical point will be zero with probability approaching one, and for $r > D$, the number of points in a ball of radius r around a typical point is zero with probability approaching zero as dimension n goes to infinity. The behavior for $r < D$ is a result of the natural separation due to dimensionality as exhibited in (4.1). However, the observation that D is the maximal such separation is due to the ν-weakly sub-Poisson property of DPPs as defined in [10], and is a feature of all DPPs, not just those studied in [80]. This behavior is the same as a sequence of Poisson point processes in the same regime, and thus this separation of points in high dimensions is due to dimensionality and not the repulsion of the DPP model. Using the coupling from the last chapter, a more precise description of the repulsion in high dimensions is given that is specific to the associated kernel of the DPP.

Theorem 3.3.2 says that there exists a coupling of X and $X^{0,!}$ such that almost surely $X^{0,!} \subseteq X$ and $\eta := X \setminus X^{0,!}$ consists of at most one point, where $X^{0,!}$ denotes a point process with the reduced Palm distribution of X. Thus, when a point is “placed at” the origin, at most one point is “pushed out”. We consider the global measure of repulsiveness

$$p = \mathbb{P}(\eta \neq \emptyset),$$

as well as the point Y with the distribution of the point in η conditioned on $\eta \neq \emptyset$. This distribution of Y characterizes the location of the repulsive effect, while p characterizes the global strength of the repulsion.

Also, the repulsive effect of a typical point over a finite distance R is quantified by $\mathbb{P}[\eta(B_n(R)) > 0]$. Note also that

$$\begin{aligned}\mathbb{P}[\eta(B_n(R)) > 0] = \mathbb{E}[\eta(B_n(R))] &= \rho \mathrm{Vol}(B_n(R)) - \mathbb{E}[X^{0,!}(B_n(R))] \\ &= \rho\left[K_{Poi}(R) - K_{DPP}(R)\right],\end{aligned}$$

where K_{Poi} and K_{DPP} are Ripley's K-functions [72] for a Poisson point process and X, respectively.

Our main results describe the behavior of these measures of repulsiveness in the Shannon regime. Consider a sequence of stationary DPPs $\{X_n\}$, such that X_n lies in $\mathbb{R}^n$. For each n, let η_n be the point process, containing at most one point, such that $X_n = X_n^{0,!} \cup \eta_n$ in distribution and $X_n^{0,!} \cap \eta_n = \emptyset$. It is often the case that $\mathbb{P}[\eta_n \neq \emptyset] \to 0$ as $n \to \infty$. In this case, the coupling inequality implies that, in high dimensions, the total variation distance is small between X_n and $X_n^{0,!}$. Indeed,

$$\|X_n - X_n^{0,!}\|_{TV} \leq \mathbb{P}(\eta_n \neq \emptyset). \tag{4.2}$$

Since X_n and $X_n^{0,!}$ have the same distribution if and only if X_n is Poisson by Slivnyak's theorem [21], this says that such DPPs look increasingly like Poisson point processes as the space dimension increases.

However, the effect of the repulsion can still be observed by conditioning on $\eta_n \neq \emptyset$. Letting Y_n be the point in η_n conditioned on $\eta_n \neq \emptyset$, Theorem 3.3.2 says that Y_n has density

$$f_n(\cdot) := |K_n(\cdot)|^2/\|K_n\|_2^2.$$

Then, under certain conditions on the kernels K_n, $|Y_n|/\sqrt{n} \to R^* \in (0, \infty)$ in probability as dimension n tends to infinity. Here, R^* is interpreted as the *asymptotic reach of repulsion* in the Shannon regime for these DPPs. This result implies that in high dimensions, a typical point has its strongest repulsive effect on points that are at a distance of $\sqrt{n}R^*$ away, since it is at this distance that a point is "pushed out".

The parametric families of DPP kernels presented in [9] and [52] provide examples of DPPs exhibiting a reach of repulsion R^* and counterexamples where no finite R^* exists, as well as computational results on the rates of convergence when a threshold does occur. Four classes of DPPs are studied in Section 4.4: Laguerre-Gaussian DPPs, power exponential DPPs, Bessel-type DPPs, and normal-variance mixture DPPs. For Laguerre-Gaussian DPPs, the sequence $|Y_n|/\sqrt{n}$ satisfies a large deviations principle (established later in Lemma 4.4.1). As a consequence, the reach of repulsion R^* becomes a phase transition for the exponential rate at which $\mathbb{P}[\eta_n\left(B_n(R\sqrt{n})\right) > 0] \to 0$ as $n \to \infty$ (established later in Proposition 4.4.2). Power exponential DPPs are shown to have a finite reach of repulsion in the Shannon regime for certain parameters (established later in Proposition 4.4.4). Bessel-type DPPs are a more repulsive family that does not exhibit an R^*, but does not exhibit repulsion past the $\sqrt{n}$ scaling (established later in Proposition 4.4.5). Finally, normal-variance mixture DPPs provide additional examples of DPPs that exhibit an R^*, including the Cauchy and Whittle-Matérn models (established later in Propositions 4.4.7 and 4.4.6).

An application of these results is presented in Section 4.5. It is shown that some threshold results in [4] for Poisson Boolean models can be extended to generalized Laguerre-Gaussian DPP Boolean models in the Shannon regime. Finally, concluding remarks and open questions are stated in Section 4.6.

4.2 Setting

We restrict to the case of Section 3.4.2.3 in the previous chapter, and consider DPPs on $\mathbb{R}^n$ with a stationary kernel. Recall that these DPPs exist if and only if the Fourier transform of the stationary kernel is bounded between zero and one. Also, for a DPP X with stationary kernel K and intensity $K(0) = \rho$, (3.3) says that the reduced Palm distribution corresponds to a DPP with kernel

$$K_0^!(x, y) = K(x - y) - \frac{1}{\rho} K(x)K(y).$$

This implies by (2.1) that the nearest neighbor function of X is $D(r) = \mathbb{P}(X^{0,!}(B_n(r)) > 0)$, with $X^{0,!} \sim DPP(K_0^!)$. All of the examples in this chapter will also be real-valued and isotropic kernels, meaning $K(x) = R(|x|)$ for some $R : \mathbb{R}^n \to \mathbb{R}$. Throughout, when we state $X \sim DPP(K)$ is stationary, it is assumed that K is stationary and real-valued. For more on this class of DPPs, see [9, 52, 77]. There exist stationary DPPs with kernels that are not of this form (see [42, 4.3.7]), but they are complex-valued and not considered here.

4.3 Reach of Repulsion

When considering the reach of repulsion of a DPP, it is natural to first consider the nearest neighbor function (2.1). The following threshold behavior was observed for stationary DPPs in [80]. It is stated here for a sequence of DPPs in the Shannon regime. For each n, let $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$ be stationary with intensity $K_n(0) = e^{n\rho}$ for some $\rho \in \mathbb{R}$. Then, for $\tilde{R} := (2\pi e)^{-\frac{1}{2}} e^{-\rho}$,

$$\lim_{n\to\infty} \mathbb{P}(X_n^{0,!}(B_n(\sqrt{n}R)) > 0) = \begin{cases} 0, & R < \tilde{R} \\ 1, & R > \tilde{R}. \end{cases} \tag{4.3}$$

A proof of this fact is given in Appendix A.1.

This shows there is a separation of points as dimension tends to infinity for any stationary DPP. However, the same threshold behavior occurs if the elements of the sequence $\{X_n\}$ are stationary Poisson point processes, as a consequence of (4.1). This observation shows that this separation is due purely to dimensionality and is not a result of the repulsiveness of DPPs.

We instead study the point process η_n and the point Y_n with the distribution of the point in η_n conditioned on $\eta_n \neq \emptyset$, as defined in the introduction, to measure the consequence of repulsiveness in high dimensions that depends on the determinantal structure. Under certain limit conditions on the kernels of a sequence of DPPs, the point Y_n that is pushed out by the point at the origin is approximately at a distance of $\sqrt{n}R^*$ for some $R^* \in (0, \infty)$ as n goes to infinity.

Lemma 4.3.1. *For each n, let $X_n \sim DPP(K_n)$ be a stationary DPP in $\mathbb{R}^n$ and Y_n a random vector in $\mathbb{R}^n$ with probability density $K_n(\cdot)^2/\|K_n\|_2^2$. Assume that as $n \to \infty$,*

$$|Y_n|/\sqrt{n} \to R^* \text{ in probability.} \tag{4.4}$$

Then,

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(B(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = \begin{cases} 0, & R < R^* \\ 1, & R > R^*. \end{cases} \tag{4.5}$$

Remark 4.3.1. One way to show (4.4) is to show that

$$\lim_{n\to\infty} \frac{\mathrm{Var}(|Y_n|^2)}{n^2} = 0 \text{ and } \lim_{n\to\infty} \left(\frac{\mathbb{E}[|Y_n|^2]}{n}\right)^{1/2} = R^* \in (0, \infty),$$

and then apply Chebychev's inequality.

For random vectors with log-concave distributions, the deviation estimate can be improved from the estimate obtained through Chebychev's inequality (see Remark 4.3.1) using Theorem 2.4.1.

Proposition 4.3.2. *Consider the setting of Lemma 4.3.1. Let* $\sigma_n^2 = \mathbb{E}|Y_n|^2$. *If* K_n^2 *is log-concave for all* n*, then there exist positive constants* C, c *such that for all* $\delta \in (0, 1)$,

$$\mathbb{P}[\eta_n(B_n(\sigma_n(1-\delta))) > 0|\eta_n \neq \emptyset] \leq Ce^{-c\sqrt{n}\delta^3},$$

and for all $\delta > 0$,

$$\mathbb{P}[\eta_n(B_n(\sigma_n(1+\delta))) > 0|\eta_n \neq \emptyset] \leq Ce^{-c\sqrt{n}\min(\delta^3,\delta)}.$$

If, in addition,

$$\lim_{n\to\infty} \sigma_n/\sqrt{n} = R^* \in (0, \infty), \tag{4.6}$$

then for this R^**, the threshold* (4.5) *occurs, and for all* $R < R^*$*, there exists a constant* $C(R) > 0$ *such that*

$$\liminf_{n\to\infty} -\frac{1}{\sqrt{n}} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0|\eta_n \neq \emptyset] \geq C(R).$$

Remark 4.3.2. The last conclusion of Proposition 4.3.2 about the rate also holds for $R > R^*$ if $B_n(\sqrt{n}R)$ is replaced by $\mathbb{R}^n \backslash B_n(\sqrt{n}R)$.

The assumption of large deviation principle (LDP) concentration leads to an estimate of the exponential rate of convergence with speed n and an exact computation of the reach of repulsion R^*.

Proposition 4.3.3. *Consider the setting of Lemma 4.3.1. Suppose* $|Y_n|/\sqrt{n}$ *satisfies a LDP with strictly convex rate function* I*. Then, for* R^* *such that*

$I(R^*) = 0$, the threshold (4.5) *occurs. Also, for $R < R^*$,*

$$-\inf_{r<R} I(r) \leq \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset]$$
$$\leq \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] \leq -\inf_{r\leq R} I(r),$$

and if the rate function I is continuous at R,

$$\lim_{n\to\infty} -\frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = I(R).$$

Remark 4.3.3. The second conclusion of Proposition 4.3.3 about the rate also holds for $R > R^*$ if $B_n(\sqrt{n}R)$ is replaced by $\mathbb{R}^n \backslash B_n(\sqrt{n}R)$.

If a sequence of DPPs in increasing dimensions exhibits a reach of repulsion R^*, this says that the point in η_n, conditioned on $\eta_n \neq \emptyset$, is most likely to be near distance $\sqrt{n}R^*$ away from the origin in high dimensions. If R^* is less than $\tilde{R}$ from (4.3), this point is most likely to be removed at a distance where points of X_n appear with probability decreasing to zero as n increases due to dimensionality. If R^* is larger than $\tilde{R}$, the point "pushed out" by repulsion is most likely to lie at a distance where points of X_n appear with high probability. Thus it is of interest to check whether there exist DPP models such that R^* is greater than or equal to $\tilde{R}$, i.e., if $\mathbb{P}(X_n^{0,!}(B_n(\sqrt{n}R^*)) = 0) \to 0$ as $n \to \infty$. In Sections 4.4.1 and 4.4.2 examples of DPP models with this reach are provided.

The above results have strong assumptions, and open up additional questions. The first question is whether Y_n tends to lie at a distance scaling with $\sqrt{n}$, i.e., is the Shannon regime the right one to examine the repulsiveness between points of a family of DPPs in high dimensions? By the radial symmetry of the density of each Y_n, the coordinates $\{Y_{n,k}\}_{k=1}^n$ are identically distributed, and the sequence $|Y_n|^2$ is the sequence of row sums of a triangular array of random variables with identically distributed rows. If the coordinate distributions depend on dimension in such a way that $\mathbb{E}\left(|Y_n|^2\right) \neq O(n)$, then a different scaling is needed.

4.4 Examples

In the following, specific families that were presented in [9] and [52] are examined that illustrate both examples of DPP models satisfying the above results, as well as examples that do not. These examples provide a window into the wide scope of repulsive behavior that can be described using this framework.

The first task will be to determine the behavior of $\mathbb{P}[\eta_n \neq \emptyset]$ as n increases. For almost all of the examples provided in this section, $\lim_{n\to\infty} \mathbb{P}[\eta_n \neq \emptyset] = 0$, but each class exhibits this convergence at different speeds. Then the goal is to determine if the DPP models satisfy the conditions of Propositions 4.3.1, 4.3.2, or 4.3.3.

4.4.1 Laguerre-Gaussian Models

For each n, let $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$ be a Laguerre-Gaussian DPP as described in [9] with intensity $K_n(0) = e^{n\rho}$, i.e., for some $m \in \mathbb{N}$, $\alpha \in \mathbb{R}^+$, let

$$K_n(x) = \frac{e^{n\rho}}{\binom{m-1+\frac{n}{2}}{m-1}} L_{m-1}^{n/2}\left(\frac{1}{m}\left|\frac{x}{\alpha}\right|^2\right) e^{-\frac{|x/\alpha|^2}{m}}, \qquad x \in \mathbb{R}^n, \tag{4.7}$$

where $L_m^\beta(r) = \sum_{k=0}^m \binom{m+\beta}{m-k} \frac{(-r)^k}{k!}$, for all $r \in \mathbb{R}$, denote the Laguerre polynomials. From [9], the condition $0 \leq \hat{K}_n \leq 1$ translates to a bound on α_n,

$$\alpha \leq \frac{1}{e^\rho (m\pi)^{1/2}} \binom{m-1+n/2}{m-1}^{\frac{1}{n}}. \tag{4.8}$$

Direct calculations give that the global measure of repulsiveness is

$$\mathbb{P}[\eta_n \neq \emptyset] = \frac{e^{n\rho}\alpha_n^n}{\binom{m-1+\frac{n}{2}}{m-1}^2} \left(\frac{m\pi}{2}\right)^{\frac{n}{2}} \sum_{k,j=0}^{m-1} \binom{m-1+\frac{n}{2}}{m-1-k} \binom{m-1+\frac{n}{2}}{m-1-j} \frac{(-1)^{k+j}}{k!j!} \frac{\Gamma\left(\frac{n}{2}+k+j\right)}{2^{k+j}\Gamma\left(\frac{n}{2}\right)}. \tag{4.9}$$

By (4.8), $\mathbb{P}[\eta_n \neq \emptyset] \leq 2^{-\frac{n}{2}} f(n,m)$, where

$$f(n,m) = \sum_{k,j=0}^{m-1} \frac{\binom{m-1+n/2}{m-1-k}\binom{m-1+n/2}{m-1-j}}{\binom{m-1+n/2}{m-1}} \frac{(-1)^{k+j}}{k!j!} \frac{\Gamma\left(\frac{n}{2}+k+j\right)}{2^{k+j}\Gamma\left(\frac{n}{2}\right)} = O(n^{m-1}).$$

It follows from [9, (5.7)] that for fixed n, $\lim_{m\to\infty} 2^{-\frac{n}{2}} f(n,m) = 1$, and as $\alpha \to 0$, K_n approaches the Poisson kernel. Thus, this class of DPPs covers a wide range of repulsiveness for fixed dimension n. However, for any fixed m, the dominant behavior as n increases is $2^{-\frac{n}{2}}$.

Since $\binom{m-1+n/2}{m-1}^{\frac{1}{n}}$ decreases to one as n goes to infinity, a sufficient condition for (4.8) to hold for all n is $0 < \alpha \leq e^{-\rho}(m\pi)^{-\frac{1}{2}}$. Note that this scaling for the intensity is the right one for observing interactions between the parameters of the model because it provides a trade-off between how large the parameter α can be and the magnitude of ρ. If the intensity did not grow as quickly with dimension, the upper bound on α would depend less and less on changes in ρ as dimension increased, and if the intensity grew more quickly, α would tend to zero as n went to infinity.

Proposition 4.3.3 holds for this sequence of DPPs. Indeed, the next lemma shows that the sequence $|Y_n|/\sqrt{n}$ satisfies a LDP.

Lemma 4.4.1. *Fix $m \in \mathbb{N}$, $\rho \in \mathbb{R}$, and let $\alpha \in (0, e^{-\rho}(m\pi)^{-1/2}]$. For each n, let Y_n be a random vector in $\mathbb{R}^n$ with probability density $K_n(\cdot)^2/\|K_n\|_2^2$, where K_n is given by (4.7). Then, the sequence $\{|Y_n|/\sqrt{n}\}_n$ satisfies an LDP with rate function*

$$\Lambda^*(x) = \frac{2x^2}{\alpha^2 m} - \frac{1}{2} + \frac{1}{2}\log\left(\frac{\alpha^2 m}{4x^2}\right).$$

Using this lemma, Proposition 4.3.3 implies that an R^* exists, and the exponential rates can be determined. In addition, using (4.9), the exponential rate of decay of $\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0]$ can be computed.

Proposition 4.4.2. *Fix $m \in \mathbb{N}$, $\rho \in \mathbb{R}$, and let $\alpha \in (0, e^{-\rho}(m\pi)^{-1/2})$. For each n, let $X_n \sim DPP(K_n)$ where K_n is given by (4.7). Then, for $R^* := \sqrt{m}\frac{\alpha}{2}$,*

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n}\log\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0] \\
&\qquad = \begin{cases} -\rho - \frac{1}{2}\log 2\pi e + \frac{2R^2}{\alpha^2 m} - \log R, & 0 < R < R^* \\ -\rho - \log\alpha - \frac{1}{2}\log\frac{m\pi}{2}, & R > R^*. \end{cases}
\end{aligned}$$

The rate decays as R increases to $R^* := \sqrt{m}\frac{\alpha}{2}$ and then for $R > R^*$, the rate no longer depends on R. This coincides with our interpretation of R^* as the asymptotic reach of repulsion of the sequence of DPPs.

For a fixed α, a larger m will give farther reach, and for a fixed m, a larger α will provide a farther reach. However, by the bound $\alpha \le e^{-\rho}(m\pi)^{-1/2}$, the following upper bound on the reach holds uniformly for all m:

$$R^* := \frac{\alpha\sqrt{m}}{2} \leq \frac{1}{2e^{\rho}\pi^{1/2}}.$$

Note that the larger ρ is, the smaller the upper bound on R^* can be. This follows from the relationship between α and ρ: the higher the intensity, the smaller α must be for the DPP to exist. Since a larger α implies a larger value of $\mathbb{P}[\eta_n(\mathbb{R}^n) > 0]$, the parameter α is associated with the strength of the repulsiveness. The relationship with ρ showcases the following tradeoff observed in [52]: the higher the intensity of the DPP, the less repulsive it can be.

As mentioned in the previous section, it is of interest to know whether there is a range of parameters such that R^* is greater than $\tilde{R}$, the threshold for the convergence of the nearest-neighbor function of X (4.3). For Laguerre-Gaussian models, $R^* := \frac{\sqrt{m}\alpha}{2}$ is larger than $\tilde{R}$ and α satisfies the condition of Lemma 4.4.1 if

$$\left(\frac{2}{e}\right)^{1/2} < e^{\rho}\sqrt{m\pi}\alpha \leq 1.$$

Since the lower bound is strictly less than one, there is a non-empty range for α such that the reach of repulsion reaches past $\tilde{R}$.

4.4.2 Power Exponential Spectral Models

The power exponential spectral models, introduced in [52], are defined through the Fourier transform of the kernel. For almost all of these models, there is no closed form for the kernel K. Using properties of the Fourier transform, a similar analysis of the repulsive behavior can still be performed.

For each n, let $X_n \sim DPP(K_n)$ be a power exponential DPP with intensity $K_n(0) = e^{n\rho}$ and parameters $\nu > 0$ and $\alpha_n > 0$, i.e., let

$$\hat{K}_n(x) = e^{n\rho} \frac{\Gamma(\frac{n}{2}+1)\alpha_n^n}{\pi^{n/2}\Gamma(\frac{n}{\nu}+1)} e^{-|\alpha_n x|^\nu}, \qquad x \in \mathbb{R}^n. \tag{4.10}$$

When $\nu = 2$, a closed form expression for K_n exists and is called the Gaussian kernel. The condition $0 \le \hat{K}_n < 1$ implies the following upper bound on α_n:

$$\alpha_n < \frac{\Gamma(\frac{n}{\nu}+1)^{\frac{1}{n}}\pi^{1/2}}{e^\rho \Gamma\left(\frac{n}{2}+1\right)^{\frac{1}{n}}}, \tag{4.11}$$

and the asymptotic expansion for the upper bound on α_n as $n \to \infty$ is

$$\left(\frac{\Gamma(\frac{n}{\nu}+1)\pi^{n/2}}{e^{n\rho}\Gamma(\frac{n}{2}+1)}\right)^{1/n} \sim \left(\frac{\sqrt{\frac{2\pi n}{\nu}}\left(\frac{n}{\nu e}\right)^{n/\nu}\pi^{n/2}}{e^{n\rho}\sqrt{\frac{2\pi n}{2}}\left(\frac{n}{2e}\right)^{n/2}}\right)^{1/n} \sim e^{-\rho} n^{\frac{1}{\nu}-\frac{1}{2}} \frac{(2\pi e)^{1/2}}{(\nu e)^{1/\nu}} = O(n^{\frac{1}{\nu}-\frac{1}{2}}).$$

By Parseval's theorem and a change of variables,

$$\begin{aligned}
\mathbb{P}[\eta_n \neq \emptyset] &= \frac{1}{e^{n\rho}}\|K_n\|_2^2 = \frac{1}{e^{n\rho}}\|\hat{K}_n\|_2^2 = \frac{1}{e^{n\rho}}\left(e^{n\rho}\frac{\Gamma\left(\frac{n}{2}+1\right)\alpha_n^n}{\pi^{n/2}\Gamma\left(\frac{n}{\nu}+1\right)}\right)^2 \int_{\mathbb{R}^n} e^{-2|\alpha_n x|^\nu} \mathrm{d}x \\
&= e^{n\rho}\left(\frac{\Gamma(\frac{n}{2}+1)\alpha_n^n}{\pi^{n/2}\Gamma(\frac{n}{\nu}+1)}\right)^2 \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \int_0^\infty r^{n-1} e^{-2(\alpha r)^\nu} \mathrm{d}r \\
&= e^{n\rho}\frac{\Gamma(\frac{n}{2}+1)\alpha_n^{2n}}{\pi^{\frac{n}{2}}\Gamma(\frac{n}{\nu}+1)^2}\frac{n}{2^{\frac{n}{\nu}}\nu\alpha_n^n}\int_0^\infty t^{\frac{n}{\nu}-1}e^{-t}\mathrm{d}t = 2^{-\frac{n}{\nu}}\alpha_n^n\frac{e^{n\rho}\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\Gamma(\frac{n}{\nu}+1)}.
\end{aligned} \tag{4.12}$$

By the bound on α_n (4.11),

$$\mathbb{P}[\eta_n \neq \emptyset] \le 2^{-\frac{n}{\nu}}.$$

For fixed dimension n, the global measure of repulsion approaches its upper bound of one for large ν. Thus, this class covers a wide range of repulsiveness similar to the Laguerre-Gaussian DPPs. However, for fixed ν, the measure decays exponentially as n goes to infinity. Note that for $\nu > 2$, the rate is smaller than for the Laguerre-Gaussian models, i.e., the decay is slower.

The following results show that if the parameters α_n grow appropriately with n, this sequence satisfies the assumptions of Proposition 4.3.1.

Lemma 4.4.3. *For each n, let Y_n be a vector in $\mathbb{R}^n$ with density $\frac{K_n^2}{\|K_n\|_2^2}$ such that $\hat{K}_n$ is given by (4.10). Assume $\alpha_n \sim \alpha n^{\frac{1}{\nu}-\frac{1}{2}}$ as $n \to \infty$ for $\alpha \in (0,\infty)$, and $\alpha_n \leq \left(\frac{\Gamma(\frac{n}{\nu}+1)\pi^{n/2}}{e^{n\rho}\Gamma(\frac{n}{2}+1)}\right)^{1/n}$ for all n. Then, as $n \to \infty$,*

$$\frac{|Y_n|}{\sqrt{n}} \to \alpha\frac{(2\nu)^{1/\nu}}{4\pi} \text{ in probability.}$$

Lemma 4.3.1 then implies the following.

Proposition 4.4.4. *For each n, let $X_n \sim DPP(K_n)$ where $\hat{K}_n$ satisfies the assumptions in Lemma 4.4.3. Then, for $R^* := \alpha\frac{(2\nu)^{1/\nu}}{4\pi}$,*

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0|\eta_n \neq \emptyset] = \begin{cases} 0, & R < R^*, \\ 1, & R > R^*. \end{cases}$$

For $\nu > 1$, the reach of repulsion R^* for the power exponential models can also reach past the nearest neighbor threshold $\tilde{R}$. Indeed, for $\alpha_n \sim \alpha n^{\frac{1}{\nu}-\frac{1}{2}}$, $R^* := \alpha\frac{(2\nu)^{1/\nu}}{4\pi}$ satisfies $\mathbb{P}[X_n(B_n(0,\sqrt{n}R^*)) = 0] \to 0$ as $n \to \infty$ if

$$\alpha\frac{(2\nu)^{1/\nu}}{4\pi} > \frac{1}{\sqrt{2\pi e e^{\rho}}}.$$

By the asymptotic formula (4.11) for the upper bound of α_n, $\alpha < \frac{\sqrt{2\pi e}}{e^{\rho}(\nu e)^{1/\nu}}$. Thus, R^* reaches past $\tilde{R}$ when $\alpha_n \sim \alpha n^{\frac{1}{\nu}-\frac{1}{2}}$ and

$$\frac{4\pi}{(2\nu)^{1/\nu}e^{\rho}\sqrt{2\pi e}} < \alpha < \frac{\sqrt{2\pi e}}{e^{\rho}(\nu e)^{1/\nu}}.$$

The interval is non-empty since the upper bound is strictly greater than the lower bound for $\nu > 1$.

4.4.3 Bessel-type Models

Another class of DPP models presented in [9] is the Bessel-type. This class is more repulsive than the previous two families of models. It is shown that while the Shannon regime is the right scaling to examine the repulsiveness of this class in high dimensions, a sequence of these DPPs does not satisfy the conditions of Proposition 4.3.1.

For each n, let $X_n \sim DPP(K_n)$ be a Bessel-type DPP with parameters $\sigma \geq 0$, $\alpha > 0$, and intensity $K_n(0) = e^{n\rho}$, for $\rho \in \mathbb{R}$. That is, let

$$K_n(x) = e^{n\rho}2^{(\sigma+n)/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)\frac{J_{(\sigma+n)/2}(2|x/\alpha|\sqrt{(\sigma+n)/2})}{(2|x/\alpha|\sqrt{(\sigma+n)/2})^{(\sigma+n)/2)}}. \tag{4.13}$$

From [9], the bound $0 \leq \hat{K}_n \leq 1$ implies that

$$\alpha_n^n \leq \frac{(\sigma+n)^{n/2}\Gamma\left(\frac{\sigma}{2}+1\right)}{e^{n\rho}(2\pi)^{n/2}\Gamma\left(\frac{\sigma+n}{2}+1\right)}. \tag{4.14}$$

Similarly to the previous examples, this family contains DPPs covering a wide range of repulsiveness measured by η_n, and as $n \to \infty$, they are more repulsive in the sense that $\mathbb{P}[\eta_n \neq \emptyset]$ decays slower. Indeed,

$$\mathbb{P}[\eta_n \neq \emptyset] = \frac{1}{e^{n\rho}} \int_{\mathbb{R}^n} K_n(x)^2 dx = e^{n\rho} \frac{(2\pi)^{n/2}\alpha^n}{(\sigma+n)^{n/2}\Gamma\left(\frac{n}{2}\right)} \frac{\Gamma\left(\frac{\sigma+n+2}{2}\right)^2 \Gamma\left(\frac{n}{2}\right)\Gamma(\sigma+1)}{\Gamma\left(\frac{\sigma}{2}+1\right)^2 \Gamma\left(\sigma+\frac{n}{2}+1\right)}$$
$$= e^{n\rho} \frac{(2\pi)^{n/2}\alpha^n}{(\sigma+n)^{n/2}} \frac{\Gamma(\sigma+1)\Gamma\left(\frac{\sigma}{2}+\frac{n}{2}+1\right)^2}{\Gamma\left(\frac{\sigma}{2}+1\right)^2 \Gamma\left(\sigma+\frac{n}{2}+1\right)},$$

and by the upper bound (4.14),

$$\mathbb{P}[\eta_n \neq \emptyset] \leq \frac{\Gamma(\sigma+1)\Gamma\left(\frac{\sigma}{2}+\frac{n}{2}+1\right)}{\Gamma\left(\frac{\sigma}{2}+1\right)\Gamma\left(\sigma+\frac{n}{2}+1\right)}.$$

By Stirling's formula, as $n \to \infty$, $\frac{\Gamma(\sigma+1)\Gamma\left(\frac{\sigma}{2}+\frac{n}{2}+1\right)}{\Gamma\left(\frac{\sigma}{2}+1\right)\Gamma\left(\sigma+\frac{n}{2}+1\right)} = O(n^{-\sigma/2})$.

These DPPs do not satisfy the conditions of Proposition 4.3.1, and so the concentration of the first moment measure does not occur, contrary to the first two families presented. However, the repulsive measure does not reach past the $\sqrt{n}$ scale in the sense of the following proposition.

Proposition 4.4.5. *Let $\rho \in \mathbb{R}$, $\alpha > 0$, and $\sigma \geq 0$. For each n, let $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$ with K_n given by* (4.13)*. Then, for any $\beta > \frac{1}{2}$ and $R > 0$,*

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(\mathbb{R}^n \backslash B_n(Rn^\beta)) > 0 | \eta_n \neq \emptyset] = 0.$$

4.4.4 Normal Variance Mixture Models

Another class of DPPs described in [52] are those with normal-variance mixture kernels. Let $X_n \sim DPP(K_n)$ be a normal-variance mixture DPP in $\mathbb{R}^n$ with intensity $e^{n\rho}$ for $\rho \in \mathbb{R}$, i.e., let

$$K_n(x) = e^{n\rho} \frac{\mathbb{E}[W^{-n/2} e^{-|x|^2/(2W)}]}{\mathbb{E}[W^{-n/2}]}, \qquad x \in \mathbb{R}^n,$$

for some non-negative real-valued random variable W such that $\mathbb{E}[W^{-n/2}] < \infty$. From [52], the bound $0 \leq \hat{K} \leq 1$ translates to the following bound on the intensity:

$$e^{n\rho} \leq \mathbb{E}[W^{-n/2}]/(2\pi)^{n/2}. \tag{4.15}$$

If $\sqrt{2W} = \alpha$, this is known as the Gaussian DPP model. If $W \sim \text{Gamma}(\nu + \frac{n}{2}, 2\alpha^2)$, this is called the Whittle-Matérn model. The Cauchy model is given when $\frac{1}{W} \sim \text{Gamma}(\nu, 2\alpha^{-2})$. In both cases $\nu > 0$ and $\alpha > 0$ are parameters.

This family of DPPs does not cover a wide range of repulsiveness like the previous families. Indeed, for any random variable W in $\mathbb{R}^+$ such that $\mathbb{E}[W^{-\frac{n}{2}}] < \infty$, Parseval's theorem, Jensen's inequality, (4.15), and Fubini's theorem imply

$$\begin{aligned}
\mathbb{P}[\eta_n \neq \emptyset] &= \frac{1}{e^{n\rho}} \int_{\mathbb{R}^n} \hat{K}_n(x)^2 \mathrm{d}x = \frac{1}{e^{n\rho}} \int_{\mathbb{R}^n} \left(e^{n\rho} \frac{(2\pi)^{\frac{n}{2}}}{\mathbb{E}\left[W^{-\frac{n}{2}}\right]} \mathbb{E}\left[e^{-2\pi^2|x|^2 W} \right] \right)^2 \mathrm{d}x \\
&\leq \frac{(2\pi)^{\frac{n}{2}}}{\mathbb{E}\left[W^{-\frac{n}{2}}\right]} \int_{\mathbb{R}^n} \mathbb{E}\left[e^{-4\pi^2|x|^2 W} \right] \mathrm{d}x \\
&= \frac{(2\pi)^{\frac{n}{2}}}{\mathbb{E}[W^{-\frac{n}{2}}]} \mathbb{E}\left((4\pi W)^{-\frac{n}{2}} \mathbb{E}\left[(4\pi W)^{\frac{n}{2}} \int_{\mathbb{R}^n} e^{-4\pi^2|x|^2 W} \mathrm{d}x \,\middle|\, W \right] \right) = 2^{-\frac{n}{2}}.
\end{aligned}$$

Is it difficult to make further general statements about this class because the behavior of the sequence $|Y_n|/\sqrt{n}$ depends greatly on the distribution of the $\mathbb{R}^+$-valued random variable W. The rest of the section will describe results for specific models in this class.

Consider a sequence of normal-variance mixture DPPs all associated with the same random variable W. If W is a constant α, the random variables X_n become multivariate Gaussian vectors with mean zero and variance depending on α. The scaled norms of these vectors are well-known to satisfy a LDP since the coordinates are independent. This also corresponds to a Laguerre-Gaussian DPP with parameter $m = 2$.

There is also a subclass of the normal-variance mixture models that satisfy Proposition 4.3.2. In [82], it is proved that if W has a log-concave density, then the normal-variance mixture distribution is log-concave. This implies that K_n^2 is log-concave, and thus if condition (4.6) holds, the conclusion of Proposition 4.3.2 holds. Since the Gamma distribution for shape parameter ν greater than 1 is log-concave and $\nu + \frac{n}{2} \geq 1$ for large n, Whittle-Matérn DPPs are an example from this subclass and exhibit an R^* as shown in the following proposition.

Proposition 4.4.6. *For each n, let $X_n \sim DPP(K_n)$ be a Whittle-Matérn model in $\mathbb{R}^n$ with intensity $e^{n\rho}$ and parameters $\nu > 0$ and $\alpha > 0$, i.e., let*

$$K_n(x) = e^{n\rho} \frac{2^{1-\nu}}{\Gamma(\nu)} \frac{|x|^\nu}{\alpha^\nu} \mathbb{K}_\nu \left(\frac{|x|}{\alpha} \right), \qquad x \in \mathbb{R}^n, \tag{4.16}$$

where $\alpha \leq \frac{\Gamma(\nu)^{\frac{1}{n}}}{\Gamma\left(\nu + \frac{n}{2}\right)^{\frac{1}{n}} 2\sqrt{\pi} e^\rho}$ and $\mathbb{K}_\nu$ is the modified Bessel kernel of the second kind. Then, for $R^* := \frac{\alpha}{2}$,

$$\lim_{n \to \infty} \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = \begin{cases} 0, & R < R^* \\ 1, & R > R^*. \end{cases}$$

Remark 4.4.1. The upper bound on α implies that for all ν,

$$R^* := \frac{\alpha}{2} \leq \frac{\Gamma(\nu)^{\frac{1}{n}}}{\Gamma\left(\nu + \frac{n}{2}\right)^{\frac{1}{n}} 4\sqrt{\pi} e^\rho} < \frac{1}{\sqrt{2\pi e} e^\rho} := \tilde{R},$$

since $\left(\frac{\Gamma(\nu)}{\Gamma\left(\nu+\frac{n}{2}\right)}\right)^{\frac{1}{n}} \leq 1$ and $4 > \sqrt{2e}$. Thus, for these models, R^* never reaches past the nearest neighbor threshold $\tilde{R}$.

Finally, the following proposition shows that the Cauchy models satisfy the conditions of Lemma 4.3.1 if the α parameter grows appropriately with n.

Proposition 4.4.7. *For each n, let $X_n \sim DPP(K_n)$ be a Cauchy model in $\mathbb{R}^n$ with intensity $e^{n\rho}$ and parameters $\nu > 0$ and $\alpha_n > 0$, i.e., let*

$$K_n(x) = \frac{e^{n\rho}}{(1 + |\frac{x}{\alpha_n}|^2)^{\nu+\frac{n}{2}}}, \qquad x \in \mathbb{R}^n.$$

Assume $\alpha_n \sim \alpha n^{1/2}$ as $n \to \infty$ for some $\alpha > 0$ such that $\alpha_n \leq \frac{\Gamma(\nu+\frac{n}{2})^{\frac{1}{n}}}{\sqrt{\pi}e^{\rho}\Gamma(\nu)^{\frac{1}{n}}}$ for each n. Then, for $R^* := \alpha$,

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = \begin{cases} 0, & R < R^* \\ 1, & R > R^*. \end{cases}$$

Remark 4.4.2. The upper bound on α_n has the following asymptotic expansion as $n \to \infty$:

$$\alpha_n \leq \frac{\Gamma(\nu + \frac{n}{2})^{\frac{1}{n}}}{\sqrt{\pi}e^{\rho}\Gamma(\nu)^{\frac{1}{n}}} \sim \frac{n^{1/2}}{\sqrt{2\pi e}e^{\rho}}.$$

Thus, if $\alpha_n \sim \alpha n^{\frac{1}{2}}$, the reach of repulsion has the upper bound

$$R^* := \alpha \leq \frac{1}{\sqrt{2\pi e}e^{\rho}}.$$

This upper bound is precisely the threshold $\tilde{R}$ for the nearest neighbor function, and so unlike in the case of Laguerre-Gaussian DPPs and power exponential DPPs, the reach of repulsion R^* for a sequence of Cauchy models with fixed parameter ν cannot reach past $\tilde{R}$.

4.5 Application to determinantal Boolean models in the Shannon regime

Poisson Boolean models in the Shannon regime were studied in [3], and the degree threshold results can be extended to Laguerre-Gaussian DPPs using Proposition 4.4.2.

The setting is the following: Consider a sequence of stationary DPPs X_n, indexed by dimension, where $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$. Assume that for each n, K_n is continuous, symmetric, and $0 \leq \hat{K}_n < 1$. Let the intensity of X_n be $K_n(0) = e^{n\rho}$. Let $X_n = \sum_k \delta_{T_n^{(k)}}$ and $R > 0$. Then, consider the sequence of particle processes [75], called determinantal Boolean models,

$$\mathcal{C}_n = \bigcup_k B_n\left(T_n^{(k)}, \frac{\sqrt{n}R}{2}\right).$$

The degree of each model is the expected number of balls that intersect the ball centered at zero under the reduced Palm distribution, i.e., $\mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))]$. In the case when X_n is Poisson, $\mathbb{E}^{0,!}[X_n(B(\sqrt{n}R))] = \mathbb{E}[X_n(B(\sqrt{n}R))]$ by Slivnyak's theorem, and

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}^{0,!}[X_n(B_n(\sqrt{n}R))] = \rho + \frac{1}{2}\log 2\pi e + \log R.$$

To extend this result to DPPs, it is sufficient to show that as $n \to \infty$,

$$\mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))] \sim \mathbb{E}[X_n(B_n(\sqrt{n}R))].$$

Note that it is impossible to extend this result to a repulsive point process like the Matérn hardcore process, since $\mathbb{E}[X_n^{0,!}(B_n(R_n))] = 0$ for all R_n less than the hardcore radius.

However, for DPPs, notice that

$$\frac{\mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]} = 1 - \frac{\mathbb{E}[\eta_n(B_n(\sqrt{n}R))]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]} = 1 - \frac{\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]}.$$

Thus, if $\frac{\mathbb{P}[\eta_n(B_n(\sqrt{n}R))>0]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]} \to 0$ as $n \to \infty$, then the degree of the determinantal Boolean model has the same asymptotic behavior as the Poisson Boolean model.

In the case of Laguerre-Gaussian kernels, this is the case, and the earlier results even provide the rate at which the quantity goes to zero, which exhibits a threshold at R^* as is expected.

Proposition 4.5.1. *Let $m \in \mathbb{N}$ and $\rho \in \mathbb{R}$. For each n, let $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$ where*

$$K_n(x) = \frac{e^{n\rho}}{\binom{m-1+n/2}{m-1}} L_{m-1}^{n/2}\left(\frac{1}{m}\left|\frac{x}{\alpha}\right|^2\right) e^{-\frac{|x/\alpha|^2}{m}},$$

and α is a parameter such that $0 < \alpha < \frac{1}{\sqrt{m\pi e^{\rho}}}$. Then,

$$\lim_{n\to\infty} -\frac{1}{n} \ln \frac{\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]} = \begin{cases} \frac{2R^2}{\alpha^2 m}, & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ \frac{1}{2} + \log 2 - \log\alpha - \frac{1}{2}\log m + \log R, & R > \sqrt{m}\frac{\alpha}{2}. \end{cases}$$

4.6 Conclusion

By examining a measure of repulsiveness of DPPs, this paper provides insight into the high dimensional behavior of different families of DPP models. Most of the families of DPPs presented in this paper have a global measure of repulsion decreasing to zero as dimension increases, indicating that they become more and more similar to Poisson point processes in high dimensions by (4.2). However, the reach of the small repulsive effect can still be quantified. By making a connection between the kernel of the DPP and the concentration in high dimensions of the norm of a random vector, we have shown under certain conditions that there exists a distance on the $\sqrt{n}$ scale at which the repulsive effect of a point of the DPP model is strongest as $n \to \infty$. It has been illustrated that some families of DPPs exhibit this reach of repulsion and some do not. The results are summarized in Table 4.1.

Some questions remain concerning the range of possible repulsive behavior of DPPs in high dimensions. First, the results can be extended to scalings other than the Shannon regime in the following way. Assumption (4.4) in Lemma 4.3.1 can be generalized to the assumption that for some sequence b_n, $\frac{|Y_n|}{b_n} \to R^*$ as $n \to \infty$. If $b_n \neq O(n^{\frac{1}{2}})$, the result holds for a different scaling than the Shannon regime, and the repulsiveness is strongest near $R^* b_n$ in high dimensions. While this is precisely what is shown not to happen for the Bessel-type DPPs if $\sigma > 0$, examples of this generalization for $b_n = o(n)$ can be obtained from the power exponential DPPs when $\alpha_n = o(n^{\frac{1}{\nu} - \frac{1}{2}})$. However, as noted in the introduction, any distance scaling smaller than $\sqrt{n}$ will not reach the regime where the expected number of points goes to infinity as dimension grows. Thus, this scaling appears less interesting. It would be interesting to find a family of DPPs that exhibits the concentration for $b_n \gg \sqrt{n}$.

For almost all of the DPPs studied in this paper, $\mathbb{P}[\eta_n \neq \emptyset] \to 0$ as

$n \to \infty$. The exception is given by the case of the Bessel-type kernel when $\sigma = 0$. For any fixed $c \in (0, 1]$, if you let $\alpha_n^n = c\frac{n^{n/2}}{e^{n\rho}(2\pi)^{n/2}\Gamma(\frac{n}{2}+1)}$, then

$$\mathbb{P}(\eta_n \neq \emptyset) = c.$$

for all n. When $c = 1$, this class of DPPs are the most repulsive DPPs on $\mathbb{R}^n$ with a stationary kernel described in Section 3.4.2.3. For these DPPs, the point Y_n has a distribution such that $|Y_n|$ has infinite k-the moment for all $k \geq 1$. It would be interesting to find a necessary and sufficient condition for $\mathbb{P}[\eta_n \neq \emptyset]$ to converge to zero, but for any applications where one would like a sequence of DPPs to return a different result that a sequence of Poisson point processes, this class of most repulsive DPPs, where $\eta_n \neq \emptyset$ a.s., is certainly the class of DPPs to focus on.

Table 4.1: Summary of Results

DPP Class	$\mathbb{P}[\eta_n \neq \emptyset] \leq$	R^*	Rate type	$R^* > \tilde{R}$
Laguerre-Gaussian	$2^{-\frac{n}{2}}O(n^{m-1})$	$\sqrt{m}\frac{\alpha}{2}$	LDP	$\left(\frac{2}{e}\right)^{\frac{1}{2}} < e^{\rho}\sqrt{m\pi}\alpha < 1$
Power-Exponential	$2^{-\frac{n}{\nu}}$	$\alpha\frac{(2\nu)^{\frac{1}{\nu}}}{4\pi}$	Chebychev	$\frac{2}{2^{\frac{1}{\nu}}e} < \frac{e^{\rho}\nu^{\frac{1}{\nu}}}{\sqrt{2\pi e}}\alpha < \frac{1}{e^{\frac{1}{\nu}}}$
Bessel-type	$O(n^{-\sigma/2})$	N/A	N/A	N/A
Whittle-Matérn	$2^{-\frac{n}{2}}$	$\frac{\alpha}{2}$	Log-concave	N/A
Cauchy	$2^{-\frac{n}{2}}$	α	Chebychev	N/A

Chapter 5

The Stochastic Geometry of Unconstrained One-bit Data Compression[1]

5.1 Introduction and Motivations

One-bit compressed sensing is a method of signal recovery from a sequence of measurements contained in $\{-1, 1\}$. More specifically, one aims to recover the signal $x \in \mathbb{R}^n$ from measurements of the form

$$y_i = \text{sign}(\langle u_i, x \rangle - t_i),$$

where the u_i are independent vectors in $\mathbb{R}^n$ and t_i random displacements in $\mathbb{R}$. One can interpret this problem geometrically, by the fact that each pair (u_i, t_i) defines a unique affine hyperplane in $\mathbb{R}^n$ with normal unit vector u_i at distance t_i from the origin. The measurement $y_i \in \{-1, 1\}$ then indicates which side of the hyperplane the signal x lies on. This collection of hyperplanes tessellates the space of signals into convex cells. Two signals contained in the same cell will have the same set of one-bit measurements $\{y_i\}$. The quality of this compression can be measured in a few different ways. For instance, one can

[1] This chapter is based on the following manuscript: *F. Baccelli and E. O'Reilly. The stochastic geometry of unconstrained one-bit data compression. arXiv:1810.06095, 2018.* The author of this thesis performed substantial research that formed the results in this manuscript.

measure how likely it is that two different signals are compressed differently, i.e., lie in different cells of the tessellation. As in one-bit compressed sensing, the quality can also be determined by having a small error in signal recovery, which can be guaranteed if the collection of hyperplanes tessellate the signal space into cells small enough to ensure all signals within a single cell are close in Euclidean distance.

Previous work ([8], [48], [70]) has examined this problem when it is known that the signal lies in some bounded set $K \subset \mathbb{R}^n$. Here, we consider the data set to be either all of $\mathbb{R}^n$ or an uncountable discrete subset of $\mathbb{R}^n$ modeled with a stationary Poisson point process. The assumption that the data is Poisson provides a worse-case scenario, since any dependence between the underlying points increases one's ability to compress the data in such a way that the signals can be recovered with small error. The set of random hyperplanes used to obtain the one-bit measurements is given by a stationary and isotropic Poisson hyperplane process. The reasons for this choice are discussed at the end of the paper (see Subsection 5.6.3), the key reason being that it leads to the least volume of data compressed with a typical data point among a wide collection of hyperplane models.

The aim is to find the minimum intensity of the hyperplane process at some scaling with the space dimension n such that different data will be separated by hyperplanes with high probability, and also for data compressed in the same way to be close with high probability. Under the assumption of stationarity, we can ask for, in some sense, a "typical" instance to satisfy the desired property. To address the "typicality", there are two viewpoints to take. One is from the view of a typical data point, and in the stationary regime, we can consider its location to be at the origin. The cell of the tessellation that the typical data point is contained in is then the so-called zero cell [21], also referred to as the Crofton cell. The other viewpoint is to ask that a typical cell satisfy some property, e.g., to have small diameter. The typical cell of a stationary Poisson hyperplane tessellation can be interpreted as the distribution of the cell obtained when taking a large ball centered at the origin, and picking a cell intersecting that ball uniformly at random. The zero cell is larger in mean than the typical cell, as there is bias towards larger cells when asking that it contain the origin. The viewpoint of a typical signal and its cell, the zero cell, seems a more natural viewpoint to take here, and will be the main focus, although some results are also derived on the typical cell for comparison.

To summarize the results, consider a sequence of compressions indexed by dimension, i.e., for each n, let X_n be a stationary and isotropic Poisson hyperplane tessellation in $\mathbb{R}^n$ with intensity γ_n that is used to compress the underlying data. We let $\gamma_n \sim \rho n^\alpha$ as $n \to \infty$ and discuss the values of α for which a good separation or low distortion of the data can be achieved with high probability by the hyperplanes when n is large. Several criteria of good separation and low distortion are discussed. By good separation, we mean a property that connects differences between data and differences between their encodings. By low distortion, we mean a property than connects closeness of data and similarity of their encodings. The results on the matter are summarized below when data are the whole of $\mathbb{R}^n$.

The first separation criterion discussed is that the distance to the nearest data that is compressed differently from the typical data (i.e., the closest point of the Euclidean space which is not in the zero cell) be small. It is shown that as long as $\alpha > 0$, this distance tends to zero in distribution as n tends to infinity.

The second separation criterion considered is that some transformation of the typical signal is compressed differently than the typical signal with high probability. We discuss two types of transformations: (i) a Gaussian displacement with fixed variance σ per dimension (which is the least demanding of the criteria discussed here), and (ii) a displacement at a fixed distance σ away and in a random direction. For case (i), we show that, for $\alpha = 0$, the typical signal is compressed in the same way as the typical signal with a probability decreasing exponentially with ρ. We also show that the same holds in case (ii) provided $\alpha = \frac{1}{2}$.

The first low distortion criterion is the requirement that the volume of other data compressed with a typical data be small. The hyperplane intensities discussed above are not large enough for this to hold. While data in most directions will be separated from the typical data, there is a set of directions of decreasing measure in which the compression will remain identical, and in high dimension, this is where most of the volume of data compressed like the typical signal lies. Considering this low distortion criterion, we show that, for

$\alpha = 1$, there is a threshold for ρ above which the expected value of the volume in question goes to zero and below which it approaches infinity.

A small volume still does not ensure that all data compressed together is close in Euclidean distance. This motivates the discussion of a second low distortion criterion. In the case where data is the whole Euclidean space, the requirement is that the point which is the farthest away from the typical data and encoded in the same way be within some distance R. It is shown that if we increase α to $\frac{3}{2}$, then there exists a finite value for ρ above which this probability approaches one as dimension n tends to infinity. A similar criterion for the case when the data is modeled with a Poisson point process is also discussed.

Some of these scalings can be significantly decreased if it is known that the data are 'sparse', namely lie within a lower dimensional subspace of $\mathbb{R}^n$. In Section 5.5, we show how this affects the intensity of hyperplanes needed for the above low distortion criteria.

The results have several implications in compressed sensing and in source coding. These are discussed in Subsections 5.6.1 and 5.6.2 at the end of the paper.

5.2 Preliminaries

5.2.1 Poisson Hyperplane Tessellations

A hyperplane process X in $\mathbb{R}^n$ is a random counting measure on the space $\mathcal{H}^n$. The process X is stationary if its distribution is invariant under translations and it is isotropic if its distribution is invariant under rotations about the origin. The hyperplane process X with intensity measure Θ is Poisson if for all disjoint $A_1, ..., A_k \in \mathcal{B}(\mathcal{H}^n)$ such that $\Theta(A_i) < \infty$ for all i,

$$\mathbb{P}(X(A_1) = m_1, ..., X(A_k) = m_k) = \prod_{i=1}^{k} \frac{\Theta(A)^{m_i}}{m_i!} e^{-\Theta(A)}.$$

Remark 5.2.1. Recall from (2.4) that the cell intensity λ of the induced random mosaic of a hyperplane process X in $\mathbb{R}^n$ is related to the intensity γ of X in the following way:

$$\lambda = \kappa_n \left(\frac{\gamma \kappa_{n-1}}{n \kappa_n} \right)^n.$$

Consider a sequence of hyperplane tessellations X_n in increasing dimensions $\mathbb{R}^n$ with intensity γ_n and cell intensity λ_n. If $\lambda_n \sim e^{n\lambda}$ as $n \to \infty$, then $\gamma_n \sim \rho n$ as $n \to \infty$. This exponential scaling with dimension for the point process of cell centroids matches the so-called Shannon regime studied in [3] and Chapter 4, and corresponds to a linear scaling of the hyperplane intensity with dimension.

5.2.2 Zero cell

The following result (see Theorem 10.4.9 in [75]) states that for stationary Poisson hyperplane processes, isotropic hyperplanes minimize the expected area of the zero cell over all spherical distributions. This result helps to justify considering the class of isotropic Poisson hyperplanes to tessellate the space, since cells of smaller volume may lead to a more efficient compression.

Theorem 5.2.1. *Let X be a nondegenerate stationary Poisson hyperplane* *process in* $\mathbb{R}^n$ *of intensity* γ*, and let* Z_0 *be the zero cell of the induced hyperplane tessellation. Then,*

$$\mathbb{E}V_n(Z_0) \geq n!\kappa_n \left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}} \right)^n ,$$

with equality if and only if X *is isotropic.*

As mentioned in the introduction, a small volume is not sufficient to ensure that two data points that have the same compression are close together. This requires the cell the points are contained in to have small diameter, but this is a difficult quantity to study. A related quantity is the radius of the smallest ball centered at the origin that contains the cell $\mathcal{C}$, i.e., the quantity

$$R_M(\mathcal{C}) = \inf\{r > 0 : \mathcal{C} \subset B(r)\}.$$

The distribution of $R_M(Z_0)$ is described in [15]. It is based on the observation that if $R_M \geq r$, then the sphere of radius r centered at the origin will not be covered by the random arcs generated by the hyperplanes that compose the faces of Z_0, i.e., $rS^{n-1} \cap \mathrm{int}(Z_0) \neq \emptyset$. Since the directional distribution of X is just the Haar measure on S^{n-1}, the probability that $R_M \geq r$ is the probability that S^{n-1} can be covered by a Poisson number N of independent spherical caps, with angular radii divided by π distributed as $d\nu(\theta) = \pi \sin(\pi\theta) 1_{[0,1/2]}(\theta) d\theta$. Unfortunately, no explicit formula for this probability is known beyond dimension two.

5.2.3 Palm Distribution

Throughout this paper, when the underlying data is assumed to be discrete, it is modeled by a stationary Poisson point process N with intensity λ. Since this is an unbounded collection of data, we need some way of examining a typical data point and the cell of the tessellation that contains it. To do this, we use the Palm probability measure of N as defined in Section 2.2.

5.3 Results

In this section, for each n, let X_n be a stationary and isotropic Poisson hyperplane process in $\mathbb{R}^n$ with intensity γ_n representing the compression scheme (note that the Poisson assumption implies that the compression scheme is characterized by the single parameter $\gamma_n > 0$, for all dimensions n). The zero cell of the tessellation is denoted $Z_{0,n}$ and the typical cell is denoted Z_n. In the case where the underlying data is discrete, N_n is a stationary Poisson point process with intensity λ_n lying in $\mathbb{R}^n$ and independent of X_n, representing the data. The Palm probability of N_n is denoted by $\mathbb{P}^0_n$.

As explained in the introduction, the goal is to find the minimum intensity γ_n needed to separate or minimize the distortion of the data $\mathbb{R}^n$ or N_n with high probability according to various criteria listed there.

5.3.1 Distance from typical data to nearest data compressed differently

Given a typical data point, we first ask how far away the closest data is that is compressed differently in any direction. When the data is all of $\mathbb{R}^n$, this is the distance to the nearest separating hyperplane in any direction. To find the distribution of this distance, notice that if no hyperplane hits the ball of radius r centered on the typical data, then this distance is greater than r. This is the spherical contact distribution [75]:

$$D_n(r) := \mathbb{P}\left(X_n\left(\mathcal{F}_{B_n(r)}\right) = 0\right).$$

Proposition 5.3.1. *Assume $\gamma_n \to \infty$ as $n \to \infty$, for example $\gamma_n \sim \rho n^\alpha$ as $n \to \infty$ for any $\alpha > 0$. Then, for fixed $r > 0$,*

$$\lim_{n\to\infty} D_n(r) = 0.$$

Proof. By the fact that X is Poisson,

$$\lim_{n\to\infty} D_n(r) = \lim_{n\to\infty} \mathbb{P}(X_n(\mathcal{F}_{B_n(r)}) = 0) = \lim_{n\to\infty} e^{-\Theta_n(\mathcal{F}_{B_n(r)})} = \lim_{n\to\infty} e^{-2\gamma_n r} = 0.$$

□

Another viewpoint to take is the distance to the nearest data compressed differently from the center of a typical cell of the tessellation, where the center is considered to be the center of the largest ball completely contained in the cell. This is equivalent to asking for the distribution of the inradius of the typical cell. Theorem 2.3.4 implies the following.

Proposition 5.3.2. *Assume $\gamma_n \to \infty$ as $n \to \infty$, for example $\gamma_n \sim \rho n^\alpha$ for any $\alpha > 0$. Then, for fixed $r > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(r_{in}(Z_n) > r) = 0.$$

5.3.2 Separation of two different data

The next criterion for separation is the probability that two different data points, one obtained by some given transformation of the other, are compressed differently, i.e., the probability that there is at least one hyperplane separating them.

First, consider the case where the transformation is a random displacement by an i.i.d. Gaussian with mean zero and variance σ^2 per dimension.

Proposition 5.3.3. *For each n, let $Y_n \sim \mathcal{N}(0, \sigma^2 I_n)$ be a Gaussian random vector in $\mathbb{R}^n$. Assume $\gamma_n \sim \rho n^\alpha$ for some $\rho > 0$ as $n \to \infty$. Then,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n \in Z_{0,n}) = \begin{cases} 0, & \alpha > 0 \\ e^{-\sqrt{\frac{2}{\pi}}\rho\sigma}, & \alpha = 0 \\ 1, & \alpha < 0. \end{cases}$$

Proof. First, by the decomposition of the spherical Lebesgue measure (Equation (1.41) in [64]), for all $x \in \mathbb{R}^n$,

$$\begin{aligned}
\Theta(\mathcal{F}_{[0,x]}) &= 2\gamma_n \int_{S^{n-1}} \int_0^\infty 1_{\{H(u,t)\cap[0,x]\neq 0\}} dt\sigma(du) \\
&= 2\gamma_n \int_{S^{n-1}} \int_0^\infty 1_{\{0\leq t\leq \langle x,u\rangle_+\}} dt\sigma(du) \\
&= 2\gamma_n \int_{S^{n-1}} \langle x,u\rangle_+ \sigma(du) = 2\gamma_n |x| \int_{S^{n-1}} \left\langle \frac{x}{|x|}, u \right\rangle_+ \sigma(du) \\
&= 2\gamma_n \frac{\kappa_{n-1}}{n\kappa_n} |x|, \qquad (5.1)
\end{aligned}$$

where $a_+ = \max(a, 0)$. Then, since X is Poisson, by (5.1),

$$\mathbb{P}(x \in Z_{0,n}) = \mathbb{P}\left(X\left(\mathcal{F}_{[0,x]}\right) = 0\right) = e^{-\Theta(\mathcal{F}_{[0,x]})} = e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}|x|}. \qquad (5.2)$$

By (5.2),

$$\mathbb{P}(Y_n \in Z_{0,n}) = \mathbb{E}\left[\mathbb{P}(Y_n \in Z_{0,n}|Y_n)\right] = \mathbb{E}\left[e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}|Y_n|}\right].$$

By the strong law of large numbers, $|Y_n|^2/n \to \sigma^2$ a.s., and by (2.7), as $n \to \infty$,

$$\frac{2\gamma\kappa_{n-1}}{n\kappa_n} \sim \frac{2\rho n^\alpha \kappa_{n-1}}{n\kappa_n} \sim \frac{2\rho n^\alpha}{\sqrt{2\pi n}} = \sqrt{\frac{2}{\pi}}\rho n^{\alpha-\frac{1}{2}}. \qquad (5.3)$$

Then, as $n \to \infty$,

$$\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}|Y_n| \sim \sqrt{\frac{2}{\pi}}\rho n^\alpha \frac{|Y_n|}{\sqrt{n}} \to \sqrt{\frac{2}{\pi}}\rho\sigma n^\alpha, \text{ a.s.}$$

Thus,

$$\lim_{n\to\infty} \mathbb{E}\left[e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}|Y_n|}\right] = \begin{cases} 0, & \alpha > 0 \\ e^{-\sqrt{\frac{2}{\pi}}\rho\sigma}, & \alpha = 0 \\ 1, & \alpha < 0. \end{cases}$$

□

Next, consider the case where the displacement is uniformly chosen on the sphere of fixed radius δ. By the fact that the tessellation is isotropic, this is equivalent to looking at the linear contact distribution for any fixed direction $u \in S^{n-1}$ at distance δ:

$$L_u(\delta) := \mathbb{P}\left(X\left(\mathcal{F}_{[0,\delta u]}\right) = 0\right).$$

Proposition 5.3.4. *For each n, let $Y_{n,\delta}$ be a uniformly chosen random point on the sphere of radius δ in $\mathbb{R}^n$. Under the same assumptions as in Proposition 5.3.3,*

$$\lim_{n\to\infty} \mathbb{P}(Y_{n,\delta} \in Z_{0,n}) = \begin{cases} 0, & \alpha > \frac{1}{2} \\ e^{-\sqrt{\frac{2}{\pi}}\rho\delta}, & \alpha = \frac{1}{2} \\ 1, & \alpha < \frac{1}{2}. \end{cases}$$

Proof. By (5.2),

$$\mathbb{P}(Y_{n,\delta} \in Z_{0,n}) = \mathbb{E}\left[e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}|Y_{n,\delta}|}\right] = e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}\delta}.$$

Then, by the asymptotic formula (5.3), as $n \to \infty$,

$$\frac{2\gamma\kappa_{n-1}}{n\kappa_n}\delta \sim \sqrt{\frac{2}{\pi}}\rho n^{\alpha-\frac{1}{2}}\delta.$$

By continuity, the conclusion holds. □

Note that a scaling of γ_n greater than $n^{\frac{1}{2}}$ (resp. more than a constant) is needed for this last separation criterion (resp. that of the Gaussian displacement) to hold as dimension increases. This is less than what is needed for the expected volume of $V_n(Z_{0,n})$ to be small as seen in the next section. This indicates that in high dimensions, most of the volume of the cell is concentrated in a set of directions with very small measure.

5.3.3 Volume of data compressed together

This section is focused on the asymptotic behavior as n goes to infinity of the volume of the data that is compressed together in a cell of the tessellation. The requirement that this volume tends to zero is a first low distortion criterion. One viewpoint is to examine the volume of data in the cell containing a typical data point. When the data is all of $\mathbb{R}^n$, this is the just the volume of $Z_{0,n}$. This quantity has been studied in [43] and [39]. The expected value is

$$\mathbb{E}[V_n(Z_{0,n})] = n!\kappa_n \left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^n = \left((n!\kappa_n)^{1/n}\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^n. \tag{5.4}$$

From [39], the following bounds on higher moments of $V_n(Z_{0,n})$ are obtained:

$$\Gamma(n+1)\kappa_n^k \left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^{kn} \le \mathbb{E}[V_n(Z_{0,n})^k] \le \Gamma(kn+1)\kappa_n^k \left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^{kn}. \tag{5.5}$$

A corollary in [39] shows there exist constants c and C, not depending on n or γ, such that

$$c\sqrt{n}\left(\frac{\pi}{e}\frac{n}{\gamma}\left(1+\frac{1}{n}\right)^{\frac{n}{2}}\right)^{2n} \le \mathrm{Var}[V_n(Z_{0,n})] \le C\sqrt{n}\left(\frac{\pi}{e}\frac{n}{\gamma}\left(1+\frac{1}{n}\right)^{\frac{n}{2}}\right)^{2n}. \tag{5.6}$$

The authors note that if γ scales with n in such a way that $\mathbb{E}[V_n(Z_{0,n})] = 1$ for all n, the lower bound implies that the variance of $V_n(Z_{0,n})$ approaches infinity as the dimension n increases, which contrasts with the behavior seen in the typical cell of the Poisson-Voronoi tessellation, where the variance converges to zero, see [2].

By the asymptotic formulas (2.7) and the above results, we obtain the following limiting behavior as dimension goes to infinity.

Proposition 5.3.5. *Let $\gamma_n \sim \rho n$ as $n \to \infty$ for some $\rho > 0$. Then,*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[V_n(Z_{0,n})] = -\ln\rho + \ln\pi - \frac{1}{2}.$$

In addition,

$$\lim_{n\to\infty} \mathbb{E}[V_n(Z_{0,n})] = \begin{cases} 0, & \rho > \frac{\pi}{\sqrt{e}} \\ \infty, & \rho < \frac{\pi}{\sqrt{e}}. \end{cases}$$

Proof. By (2.7), as $n \to \infty$,

$$\begin{aligned}(n!\kappa_n)^{1/n} \frac{n\kappa_n}{2\gamma\kappa_{n-1}} &\sim \left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \frac{1}{\sqrt{n\pi}}\left(\frac{2\pi e}{n}\right)^{n/2}\right)^{1/n} \frac{\sqrt{2\pi n}}{2\gamma} \\ &\sim \frac{n}{e}\frac{\sqrt{2\pi e}}{\sqrt{n}}\frac{\sqrt{\pi n}}{\sqrt{2}\gamma} = \frac{\pi}{\sqrt{e}}\frac{n}{\gamma}.\end{aligned}$$

Thus, by (5.4), under the assumption $\gamma_n \sim \rho n$, we have the following limiting behavior:

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[V_n(Z_{0,n})] &= \lim_{n\to\infty} \ln\left[(n!\kappa_n)^{1/n} \frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right] \\ &= \ln\frac{\pi}{\sqrt{e}\rho} = -\ln\rho + \ln\pi - \frac{1}{2}.\end{aligned}$$

This implies the last statement. □

Another viewpoint is to consider the volume of the typical cell Z_n of the tessellation. This measures the volume of a typical collection of data that is compressed together.

Proposition 5.3.6. *If $\gamma_n \sim \rho n$ for some $\rho > 0$ as $n \to \infty$, then*

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[V_n(Z_n)] = -\ln\rho - \frac{1}{2}.$$

In addition,

$$\lim_{n\to\infty} \mathbb{E}[V_n(Z_n)] = \begin{cases} 0, & \rho > \frac{1}{\sqrt{e}} \\ \infty, & \rho < \frac{1}{\sqrt{e}}. \end{cases}$$

Proof. By (2.5), the expected value of the volume is

$$\mathbb{E}[V_n(Z_n)] = \frac{1}{\kappa_n} \left(\frac{n\kappa_n}{\gamma \kappa_{n-1}} \right)^n.$$

Then, by (2.7), as $n \to \infty$,

$$\frac{1}{\kappa_n^{1/n}} \frac{n\kappa_n}{\gamma_n \kappa_{n-1}} \sim (n\pi)^{1/n} \left(\frac{n}{2\pi e} \right)^{1/2} \frac{\sqrt{2\pi n}}{\gamma_n} \sim \frac{n}{\gamma_n \sqrt{e}}.$$

Thus, assuming $\gamma_n \sim \rho n$ as $n \to \infty$ for $\rho > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[V(Z_n)] = -\log \rho - \frac{1}{2}.$$

The right hand side is positive if $\rho < e^{-1/2}$ and negative if $\rho > e^{-1/2}$, which implies the last statement of the proposition. □

When the data set is (the support of) a stationary Poisson point process, the volume of the zero cell has to be replaced by the number of points of N_n that lie in $Z_{0,n}$. A similar threshold exists for the expected amount of data in $Z_{0,n}$, but it depends on the intensity of N_n. This then implies that for ρ big enough, the probability that there is another data point in the cell of a typical data is small, meaning that with high probability, the cell of the tessellation determines the data uniquely.

Proposition 5.3.7. *For each n, assume N_n is a Poisson point process in $\mathbb{R}^n$ with intensity $\lambda_n = n^{n(\alpha-1)} e^{n\lambda}$ for some $\lambda \in \mathbb{R}$ and $\alpha \in \mathbb{R}$. Let $\gamma_n \sim \rho n^\alpha$ as*

$n \to \infty$ for some $\rho > 0$. Then,

$$\lim_{n\to\infty} \mathbb{E}_n^{0,!}[N_n(Z_{0,n})] = \begin{cases} 0, & \rho > e^{\lambda}\pi/\sqrt{e} \\ \infty, & \rho < e^{\lambda}\pi/\sqrt{e}. \end{cases}$$

Thus, for $\rho > \frac{e^{\lambda}\pi}{\sqrt{e}}$,

$$\lim_{n\to\infty} \mathbb{P}_n^0(N_n(Z_{0,n}) = 1) \to 1.$$

Proof. By Slivnyak's theorem,

$$\mathbb{E}_n^{0,!}[N_n(Z_{0,n})] = \mathbb{E}[N_n(Z_{0,n})] = \mathbb{E}\left[\mathbb{E}[N_n(Z_{0,n})|Z_{0,n}]\right] = \lambda_n \mathbb{E}[V_n(Z_{0,n})]. \quad (5.7)$$

By the assumption on γ_n and (2.7),

$$\frac{2\gamma_n \kappa_{n-1}}{n\kappa_n} \sim \frac{\sqrt{2}\rho}{\sqrt{\pi}} n^{\alpha - \frac{1}{2}}, \text{ as } n \to \infty. \quad (5.8)$$

Then, by (2.7) and (5.8), as $n \to \infty$,

$$\frac{1}{n} \log \mathbb{E}[V_n(Z_{0,n})] \sim \log\left(\frac{n}{e}\right) + \frac{1}{2}\log\frac{2\pi e}{n} + \log\frac{\sqrt{\pi}}{\sqrt{2}\rho n^{\alpha-\frac{1}{2}}} = (1-\alpha)\log n + \log\frac{\pi}{\rho\sqrt{e}}.$$

By the assumption on λ_n and (5.7), as $n \to \infty$,

$$\frac{1}{n}\log \mathbb{E}_n^{0,!}[N_n(Z_{0,n})] \sim \lambda + \log\frac{\pi}{\rho\sqrt{e}}. \quad (5.9)$$

The threshold follows. Then, by Slivnyak's theorem and Jensen's inequality,

$$\begin{aligned} \mathbb{P}_n^0(N_n(Z_{0,n}) = 1) &= \mathbb{P}(N_n(Z_{0,n}) = 0) = \mathbb{E}[e^{-\lambda_n V_n(Z_{0,n})}] \\ &\geq e^{-\lambda_n \mathbb{E}[V_n(Z_{0,n})]} = e^{-\mathbb{E}_n^{0,!}[N_n(Z_{0,n})]}. \end{aligned}$$

Thus, for $\rho > e^{\lambda}\pi/\sqrt{e}$,

$$\lim_{n\to\infty} \mathbb{P}_n^0(N_n(Z_{0,n}) = 1) = 1.$$

□

5.3.4 Farthest distance between two data points compressed together

Another and more demanding low distortion criterion is that all the data compressed together be close in Euclidean distance. Consider first the case when the data is all of $\mathbb{R}^n$. We want to find the scaling necessary for γ_n to ensure that all data points in the zero cell are within some distance from the typical data point at the origin. This is equivalent to showing that the radius of the smallest ball centered at the origin that contains all of the zero cell is small. As mentioned in Section 5.2.2, a closed form for the distribution of this radius R_M is only known in dimension two, but we can obtain bounds that give the following asymptotic behavior.

Theorem 5.3.8. *Assume $\gamma_n \sim \rho n^\alpha$ as $n \to \infty$ and let $R > 0$. Then, there exist constants x_ℓ in the interval $(0, 1)$ and x_u in the interval $(1, \infty)$, independent of R, such that for all $\rho > \rho_u := x_u \frac{\sqrt{\pi}}{R\sqrt{2}}$,*

$$\lim_{n\to\infty} \mathbb{P}(R_M(Z_{0,n}) \geq n^{3/2-\alpha} R) = 0.$$

and for all $\rho < \rho_\ell := x_\ell \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \mathbb{P}(R_M(Z_{0,n}) \leq n^{3/2-\alpha} R) = 0.$$

Before proving the Theorem, we need the following. Define the beta prime density with parameters $n \in \mathbb{N}$ and $\sigma > 0$ as follows:

$$f_{n,\sigma}(x) = c_{n,\sigma} \left(1 + \frac{|x|^2}{\sigma^2}\right)^{-\frac{n+1}{2}} \text{ for } x \in \mathbb{R}^n, \text{ with } c_{n,\sigma} = \frac{\Gamma(\frac{n+1}{2})}{\sigma^n \pi^{n/2} \Gamma(\frac{1}{2})}.$$

Let $X_1, \ldots, X_m$ be i.i.d random vectors in $\mathbb{R}^n$ with density $f_{n,\sigma}$ and let $P^\sigma_{m,n}$ denote the convex hull of these points. Also, define $A := A(X_1, ..., X_n)$ to be the $d-1$ dimensional affine subspace containing the points $X_1, \ldots, X_n$, and let $h(A)$ be the signed distance from the origin to the subspace A. The following lemma gives the probability that the points $X_1, \ldots, X_n$ form a face of $P^\sigma_{m,n}$.

Lemma 5.3.9.

$$\mathbb{P}\left([X_1, \ldots, X_n] \textit{ is a facet in } P^\sigma_{m,n} \textit{ such that } |h(A)| \leq r\right)$$
$$= \frac{2\Gamma(\frac{d+1}{2})}{\sigma\Gamma(\frac{d}{2})\sqrt{\pi}} \int_{-r}^{r} \left(1+\frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} \left(\frac{1}{\sigma\pi}\int_{-\infty}^{t}\left(1+\frac{s^2}{\sigma^2}\right)^{-1} \mathrm{d}s\right)^{m-n} \mathrm{d}t.$$

Proof. Let $\pi_{A^\perp}$ be the projection from $\mathbb{R}^n$ to the 1-dimensional subspace $A^\perp$ and define the isometry $I_{A^\perp} : A^\perp \mapsto \mathbb{R}$ such that $I_{A^\perp}(0) = 0$.

By Lemma 3.1 in [46], if X has density $f_{n,\sigma}$, then $I_{A^\perp}(\pi_{A^\perp}(X))$ has density

$$f_{1,\sigma}(s) = \frac{1}{\sigma\pi}\left(1+\frac{s^2}{\sigma^2}\right)^{-1}.$$

This was stated with $\sigma = 1$ in the reference, but if X has density $f_{n,\sigma}$, then X/σ has density $\tilde{f}_{n,1}$, and the more general statement follows from a change a variables, since $I_{A^\perp}(\pi_{A^\perp}(X/\sigma)) = I_{A^\perp}(\pi_{A^\perp}(X))/\sigma$.

Also, by Corollary 3.6 in [46], if $X_1, \ldots X_n$ have the beta prime density $f_{n,1}$, then $h^2(A)/\sigma^2$ has density

$$g(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi}} t^{-\frac{1}{2}}(1+t)^{-(\frac{n+1}{2})} 1_{\{t\geq 0\}}.$$

By a changes of variables,

$$\mathbb{P}\left(|h(A)| \leq r\right) = \frac{2\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi}} \int_0^{r/\sigma} (1+y^2)^{-\frac{n+1}{2}} dy = \frac{2\Gamma(\frac{n+1}{2})}{\sigma\Gamma(\frac{n}{2})\sqrt{\pi}} \int_0^{r} \left(1+\frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} dt.$$

Hence, the distribution of $|h(A)|$ has density

$$\tilde{h}(t) = \frac{2\Gamma(\frac{n+1}{2})}{\sigma\Gamma(\frac{n}{2})\sqrt{\pi}} \left(1 + \frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} 1_{\{t\geq 0\}}.$$

Then, by the fact that $[X_1, \ldots, X_n]$ is a facet of $P^{\sigma}_{m,n}$ if and only if $I_{A^\perp}(\pi_{A^\perp}(X_i)) \leq h(A)$ for all $i = n+1, \ldots, m$, or $I_{A^\perp}(\pi_{A^\perp}(X_i)) \geq h(A)$ for all $i = n+1, \ldots, m$. This gives

$$\begin{aligned}
&\mathbb{P}\left([X_1, \ldots, X_n] \text{ is a facet in } P^{\sigma}_{m,n} \text{ such that } |h(A)| \leq r\right)\\
&= \int_0^r \mathbb{P}\left([X_1, \ldots, X_n] \text{ is a facet } \text{ in } P^{\sigma}_{m,n} \,\middle|\, |h(A)| = t\right) \tilde{h}(t)dt\\
&= \int_0^r (\mathbb{P}\left(I_{A^\perp}(\pi_{A^\perp}(X_i)) \leq t \text{ for each } i = n+1, \ldots, m\right)\\
&\qquad\qquad + \mathbb{P}\left(I_{A^\perp}(\pi_{A^\perp}(X_i)) \geq t \text{ for each } i = n+1, \ldots, m\right))\tilde{h}(t)dt\\
&= \int_0^r \left(\int_{-\infty}^t f_{1,\sigma}(s)ds\right)^{m-n} \tilde{h}(t)dt + \int_0^r \left(\int_t^{\infty} \tilde{f}_{1,\sigma}(s)ds\right)^{m-n} \tilde{h}(t)dt\\
&= \int_{-r}^r \left(\int_{-\infty}^t f_{1,\sigma}(s)ds\right)^{m-n} \tilde{h}(t)dt,
\end{aligned}$$

where the last inequality follows from the fact that the densities are symmetric. Hence,

$$\begin{aligned}
&\mathbb{P}\left([X_1, \ldots, X_n] \text{ is a facet in } P^{\sigma}_{m,n} \text{ such that } |h(A)| \leq r\right)\\
&\quad = \frac{2\Gamma(\frac{n+1}{2})}{\sigma\Gamma(\frac{n}{2})\sqrt{\pi}} \int_{-h}^{h} \left(1 + \frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} \left((\sigma\pi)^{-1} \int_{-\infty}^{t} \left(1 + \frac{s^2}{\sigma^2}\right)^{-1} \mathrm{d}s\right)^{m-n} \mathrm{d}t.
\end{aligned}$$

□

We can now prove Theorem 5.3.8.

Proof. (of Theorem 5.3.8)

Let X be a random vector in $\mathbb{R}^n$ with density $f_{n,\sigma}$. By a generalization of Lemma 7.7 in [45], we have the vague convergence

$$m\mathbb{P}(m^{-1}X \in \cdot) \to \nu(\cdot), \tag{5.10}$$

as $m \to \infty$, where ν is a measure on $\mathbb{R}^n \backslash \{0\}$ with density

$$x \mapsto \frac{2\sigma}{\omega_{n+1}} |x|^{-n-1}. \tag{5.11}$$

Let $\Pi_n(\sigma)$ be a Poisson point process on $\mathbb{R}^n \backslash \{0\}$ with intensity measure ν. Then, (5.10) implies the following generalization of (4.6) in [45]: As $m \to \infty$,

$$\sum_{i=1}^{m} \delta_{X_i/m} \to \Pi_n(\sigma) \text{ in distribution,} \tag{5.12}$$

where $X_1, \ldots, X_m$ are i.i.d random vectors in $\mathbb{R}^n$ with density $f_{n,\sigma}$. Now, let $P_{m,n}^{\sigma}$ be the convex hull of $X_1, \ldots, X_m$. The convergence (5.12) implies that

$$\begin{aligned} &\lim_{m \to \infty} \mathbb{E}[\# \text{ of faces within distance } mh \text{ in } P_{m,n}^{\sigma}] \\ &\quad = \mathbb{E}[\# \text{ of faces within distance } h \text{ in } \mathcal{C}(\Pi_n(\sigma))], \end{aligned}$$

with $\mathcal{C}(P)$ denoting the convex hull of the points in set P. Now, by the same argument as in the proof of Theorem 1.21 of [46], the convex dual of $\mathcal{C}(\Pi_n(\sigma))$ has the same distribution as the zero cell $Z_{0,n}$ of a stationary and isotropic hyperplane tessellation with intensity $\gamma_n = \frac{\sigma\omega_n}{\omega_{n+1}}$. Hence, the distances to the facets of the convex hull of $\Pi_n(\sigma)$ are the reciprocal of the distances to the vertices of $Z_{0,n}$. This gives

$$
\begin{aligned}
&\mathbb{E}[\#\text{ of vertices at distance greater than } r \text{ in } Z_{0,n}] \\
&\qquad = \mathbb{E}[\#\text{ of facets at distance less than } r^{-1} \text{ in } \mathcal{C}(\Pi_n(\sigma))] \\
&\qquad = \lim_{m\to\infty} \mathbb{E}[\#\text{ of facets at distance less than } mr^{-1} \text{ in } P^{\sigma}_{m,n}] \\
&\qquad = \lim_{m\to\infty} \binom{m}{n} \mathbb{P}\left([X_1,\dots,X_n] \text{ is a facet of } P^{\sigma}_{m,n} \text{ such that } |h(A)| \le mr^{-1}\right) \\
&\qquad = \lim_{m\to\infty} \binom{m}{n} \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} \int_{-m/r}^{m/r} \left(1+\frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} \left(\frac{1}{\pi\sigma}\int_{-\infty}^{t}\left(1+\frac{s^2}{\sigma^2}\right)^{-1}\mathrm{d}s\right)^{m-n} \mathrm{d}t,
\end{aligned}
$$

where the last equality follows from Lemma 5.3.9. By the same arguments as in Lemma 4.9 in [46], as $m \to \infty$,

$$
\begin{aligned}
\int_{-m/r}^{m/r} \left(1+\frac{t^2}{\sigma^2}\right)^{-\frac{n+1}{2}} \left(\frac{1}{\pi\sigma}\int_{-\infty}^{t}\left(1+\frac{s^2}{\sigma^2}\right)^{-1}\mathrm{d}s\right)^{m-n} \mathrm{d}t \\
\sim m^{-n}\sigma\pi^n\Gamma(n)\Gamma_u\left(n,\pi^{-1}\sigma r\right).
\end{aligned}
$$

Then, since $\binom{m}{n} \sim \frac{m^n}{n!}$ as $m \to \infty$,

$$
\begin{aligned}
&\binom{m}{n} \frac{2\Gamma(\frac{n+1}{2})}{\sigma\sqrt{\pi}\Gamma(\frac{n}{2})} \int_{-m/r}^{m/r} (1+t^2)^{-\frac{n+1}{2}} \left(\tilde{c}_{1,\frac{n+1}{2}}\int_{-\infty}^{t}(1+s^2)^{-1}\mathrm{d}s\right)^{m-n} \mathrm{d}t \\
&\sim \frac{2}{n}\frac{\Gamma(\frac{n+1}{2})\pi^n}{\sqrt{\pi}\Gamma(\frac{n}{2})}\Gamma_u\left(n,\sigma\pi^{-1}r\right) = \pi^{n-\frac{1}{2}}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}+1)}\Gamma_u\left(n,\sigma\pi^{-1}r\right).
\end{aligned}
$$

Let $\gamma_n = \frac{\sigma\omega_n}{\omega_{n+1}}$, i.e., let $\sigma = \gamma_n\frac{\omega_{n+1}}{\omega_n}$. Then,

$$
\mathbb{E}[\#\text{ of vertices farther than } r \text{ in } Z_{0,n}] = \frac{\Gamma(\frac{n+1}{2})\pi^n}{\sqrt{\pi}\Gamma(\frac{n}{2}+1)}\Gamma_u\left(n,\gamma_n\frac{\omega_{n+1}}{\pi\omega_n}r\right).
$$

Similar computations give

$$
\mathbb{E}[\#\text{ of vertices closer than } r \text{ in } Z_{0,n}] = \frac{\Gamma(\frac{n+1}{2})\pi^n}{\sqrt{\pi}\Gamma(\frac{n}{2}+1)}\Gamma_\ell\left(n,\gamma_n\frac{\omega_{n+1}}{\pi\omega_n}r\right).
$$

Now, by Markov's inequality,

$$\begin{aligned}\mathbb{P}(R_M(Z_{0,n}) \geq n^{3/2-\alpha}R) &= \mathbb{P}(\#\text{ of vertices farther than } n^{3/2-\alpha}R \text{ in } Z_{0,n} > 0) \\ &\leq \mathbb{E}[\#\text{ of vertices farther than } n^{3/2-\alpha}R \text{ in } Z_{0,n}] \\ &= \frac{\Gamma(\frac{n+1}{2})\pi^n}{\sqrt{\pi}\Gamma(\frac{n}{2}+1)}\Gamma_u\left(n, \gamma_n \frac{\omega_{n+1}}{\pi\omega_n} n^{3/2-\alpha}R\right).\end{aligned}$$

Also,

$$\begin{aligned}\mathbb{P}(R_M(Z_{0,n}) \leq n^{3/2-\alpha}R) &\leq \mathbb{P}(\#\text{ of vertices closer than } n^{3/2-\alpha}R \text{ in } Z_{0,n} > 0) \\ &\leq \mathbb{E}[\#\text{ of vertices closer than } n^{3/2-\alpha}R \text{ in } Z_{0,n}] \\ &= \frac{\Gamma(\frac{n+1}{2})\pi^n}{\sqrt{\pi}\Gamma(\frac{n}{2}+1)}\Gamma_\ell\left(n, \gamma_n \frac{\omega_{n+1}}{\pi\omega_n} n^{3/2-\alpha}R\right).\end{aligned}$$

By the assumption on γ_n and (2.6), as $n \to \infty$,

$$\gamma_n \frac{\omega_{n+1}}{\pi\omega_n} n^{3/2-\alpha}R \sim \rho n^\alpha \frac{2\pi^{\frac{n+1}{2}}\Gamma(n/2)}{\pi\Gamma(\frac{n+1}{2})2\pi^{n/2}} n^{3/2-\alpha}R \sim \rho n^{3/2}\left(\frac{2}{\pi n}\right)^{1/2} R = \frac{\rho R\sqrt{2}}{\sqrt{\pi}} n.$$

Then, by Laplace's method (see Lemma A.2 in [66]), for $\rho > \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \Gamma_u\left(n, \gamma_n \frac{\omega_{n+1}}{\pi\omega_n} n^{3/2-\alpha}R\right) = \ln \frac{\rho R\sqrt{2}}{\sqrt{\pi}} - \frac{\rho R\sqrt{2}}{\sqrt{\pi}} + 1.$$

and similarly, for $\rho < \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \Gamma_\ell\left(n, \gamma_n \frac{\omega_{n+1}}{\pi\omega_n} n^{3/2-\alpha}R\right) = \ln \frac{\rho R\sqrt{2}}{\sqrt{\pi}} - \frac{\rho R\sqrt{2}}{\sqrt{\pi}} + 1.$$

Since $\frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2}+1)} = O(n^{-1/2})$, for $\rho > \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(R_M(Z_{0,n}) \geq n^{3/2-\alpha}R) \leq \ln \rho R\sqrt{2\pi} - \frac{\rho R\sqrt{2}}{\sqrt{\pi}} + 1,$$

and for $\rho < \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(R_M(Z_{0,n}) \leq n^{3/2-\alpha} R) \leq \ln \rho R \sqrt{2\pi} - \frac{\rho R\sqrt{2}}{\sqrt{\pi}} + 1,$$

The function $\ln \pi + \ln x - x + 1$ is concave, and has two zeros, one $0 < x_\ell < 1$ and one where $x_u > 1$. These zeros determine the values of

$$\rho_\ell := x_\ell \frac{\sqrt{\pi}}{R\sqrt{2}} \quad \text{and} \quad \rho_u := x_u \frac{\sqrt{\pi}}{R\sqrt{2}},$$

respectively.

□

Next consider the case where the underlying data is a Poisson point process, and more precisely the regime where the expected number of points in the zero cell goes to infinity. Theorem 5.3.10 below gives a sufficient condition for all points of the point process which are contained in the zero cell (the cell of the typical data) to be within distance R_n from the point at the origin (the typical data). The result also shows that the same scaling that is sufficient for the criterion to be satisfied is also necessary.

Theorem 5.3.10. *Consider the setting of the Proposition 5.3.7, with λ fixed, and assume that $\rho < \rho^* := \frac{e^\lambda \pi}{\sqrt{e}}$.*

(i) If $R > \frac{\sqrt{e}}{e^\lambda \sqrt{2\pi}}$, then $\frac{\sqrt{\pi}}{R\sqrt{2}} < \rho^*$ and for all ρ in the interval $(\frac{\sqrt{\pi}}{R\sqrt{2}}, \rho^*)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \geq R n^{\frac{3}{2}-\alpha} \,\middle|\, N_n(Z_{0,n}) > 0 \right)$$
$$\leq \lambda + \frac{1}{2} \log 2\pi e + \log R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} + \log 4.$$

(ii) Let

$$a(R,\lambda) = \max\left(\left(\lambda + \frac{1}{2}\log 2\pi e + \log R + \log 4\right), 1\right) \geq 1.$$

If R is such that $\rho_u := \frac{\sqrt{\pi}}{R\sqrt{2}}\, a(R,\lambda) < \rho^*$, which holds for R large enough, then for all ρ in the interval (ρ_u, ρ^*),

$$\lim_{n\to\infty} \mathbb{P}_n^0\left(\max_{x_i\in N_n\cap Z_0} |x_i| \geq R n^{\frac{3}{2}-\alpha} \,\middle|\, N_n(Z_{0,n}) > 0\right) = 0, \tag{5.13}$$

where the convergence is at least exponential of rate $\lambda + \frac{1}{2}\log 32\pi e R^2 - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} < 0$.

(iii) For all $\rho < \min\left(\frac{\sqrt{\pi}}{R\sqrt{2}}, \rho^*\right)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n^0\left(\max_{x_i\in N_n\cap Z_{0,n}} |x_i| \leq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0\right)$$
$$\leq \lambda + \frac{1}{2}\log 2\pi e + \log R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} + \log 4.$$

(iv) If $R < (4e^{\lambda}\sqrt{2\pi e})^{-1}$, then for all ρ in $\left(0, \min\left(\frac{\sqrt{\pi}}{R\sqrt{2}}, \rho^*\right)\right)$,

$$\lim_{n\to\infty} \mathbb{P}_n^0\left(\max_{x_i\in N_n\cap Z_0} |x_i| \leq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0\right) = 0, \tag{5.14}$$

where the convergence is at least exponential of rate $\lambda + \frac{1}{2}\log 32\pi e R^2 - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} < 0$.

Proof. First, by (5.2) and two changes of variable,

$$\begin{aligned}
\mathbb{E}[N_n(Z_{0,n} \cap B_n(R)^c)] &= \lambda_n \mathbb{E}[V_n(Z_{0,n} \cap B_n(R)^c)] \\
&= \lambda_n \int_{B_n(R)^c} \mathbb{P}(x \in Z_{0,n}) dx = \int_{B_n(R)^c} e^{-\frac{2\gamma_n \kappa_{n-1}}{n\kappa_n}|x|} dx \\
&= \lambda_n n\kappa_n \int_R^\infty r^{n-1} e^{-\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} r} dr \\
&= n\kappa_n \left(\frac{2\gamma_n \kappa_{n-1}}{n\kappa_n}\right)^{-n} \int_{\frac{2\gamma_n\kappa_{n-1}}{n\kappa_n}R}^\infty y^{n-1} e^{-y} dy \\
&= \lambda_n n! \kappa_n \left(\frac{2\gamma_n \kappa_{n-1}}{n\kappa_n}\right)^{-n} \Gamma_u\left(n, \frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} R\right) \\
&= \lambda_n \mathbb{E}[V_n(Z_{0,n})] \Gamma_u\left(n, \frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} R\right) \\
&= \mathbb{E}[N_n(Z_{0,n})] \Gamma_u\left(n, \frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} R\right),
\end{aligned}$$

and similarly,

$$\mathbb{E}[N_n(Z_{0,n} \cap B_n(R))] = \mathbb{E}[N_n(Z_{0,n})] \Gamma_\ell\left(n, \frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} R\right).$$

By Laplace's method (see Lemma A.2 in [66]) and (5.8), if $\frac{\sqrt{2}\rho R}{\sqrt{\pi}} > 1$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \Gamma_u\left(n, \frac{2\gamma_n\kappa_{n-1}}{n\kappa_n} n^{\frac{3}{2}-\alpha} R\right) = \log \frac{\sqrt{2}\rho R}{\sqrt{\pi}} - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} + 1,$$

and the limit is 1 otherwise. Then, by (5.9),

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R)^c)] = \begin{cases} \lambda + \ln \sqrt{2\pi e} R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}}, & \frac{\sqrt{2}\rho R}{\sqrt{\pi}} > 1 \\ \lambda + \ln \frac{\pi}{\rho\sqrt{e}}, & \frac{\sqrt{2}\rho R}{\sqrt{\pi}} < 1. \end{cases} \tag{5.15}$$

Similarly,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R))] = \begin{cases} \lambda + \ln \sqrt{2\pi e} R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}}, & \frac{\sqrt{2}\rho R}{\sqrt{\pi}} < 1 \\ \lambda + \ln \frac{\pi}{\rho\sqrt{e}}, & \frac{\sqrt{2}\rho R}{\sqrt{\pi}} > 1. \end{cases} \tag{5.16}$$

Next, by the fact that N_n is Poisson,

$$\mathbb{E}[N_n(Z_{0,n})^2] = \mathbb{E}[\mathbb{E}[N_n(Z_{0,n})^2 | Z_{0,n}]] = \lambda_n^2 \mathbb{E}[V_n(Z_{0,n})^2] + \lambda_n \mathbb{E}[V_n(Z_{0,n})],$$

and by the second moment inequality, we have

$$\begin{aligned}
\mathbb{P}(N_n(Z_{0,n}) > 0) &\geq \frac{\mathbb{E}[N_n(Z_{0,n})]^2}{\mathbb{E}[N_n(Z_{0,n})^2]} = \frac{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})]^2}{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})^2] + \lambda_n \mathbb{E}[V_n(Z_{0,n})]} \\
&= \frac{\mathbb{E}[V_n(Z_{0,n})]^2}{\mathbb{E}[V_n(Z_{0,n})^2]} \left(\frac{1}{1 + \frac{\lambda_n \mathbb{E}[V_n(Z_{0,n})]}{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})^2]}} \right). \qquad (5.17)
\end{aligned}$$

Then, by Jensen's inequality,

$$\frac{\lambda_n \mathbb{E}[V_n(Z_{0,n})]}{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})^2]} \leq \frac{\lambda_n \mathbb{E}[V_n(Z_{0,n})]}{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})]^2} = \frac{1}{\lambda_n \mathbb{E}[V_n(Z_{0,n})]},$$

and by the assumption on ρ, $\lim_{n\to\infty} \lambda_n \mathbb{E}[V_n(Z_{0,n})] = \infty$ by Proposition 5.3.7, and so

$$\lim_{n\to\infty} \frac{\lambda_n \mathbb{E}[V_n(Z_{0,n})]}{\lambda_n^2 \mathbb{E}[V_n(Z_{0,n})^2]} = 0.$$

Then, by (5.4) and (5.5), as $n \to \infty$,

$$\mathbb{P}(N_n(Z_{0,n}) > 0) \gtrsim \frac{\mathbb{E}[V_n(Z_{0,n})]^2}{\mathbb{E}[V_n(Z_{0,n})^2]} \sim \frac{\Gamma(n+1)^2}{\Gamma(2n+1)}.$$

Now, by Markov's inequality and (5.17),

$$\begin{aligned}
&\mathbb{P}_n^0 \left(\max_{x_i \in Z_{0,n} \cap N_n} |x_i| \geq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0 \right) \\
&\quad = \mathbb{P}_n^{0,!} \left(N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R)^c) > 0 \,\middle|\, N_n(Z_{0,n}) > 0 \right) \\
&\quad = \frac{\mathbb{P}_n^{0,!} \left(N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R)^c) > 0 \right)}{\mathbb{P}_n^{0,!}(N_n(Z_{0,n}) > 0)} \\
&\quad \lesssim \mathbb{E}[N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R)^c)] \frac{\Gamma(2n+1)}{\Gamma(n+1)2}.
\end{aligned}$$

Thus, by (2.7) and (5.15), for $\rho > \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \geq R \,\middle|\, N_n(Z_{0,n}) > 0 \right)$$
$$\leq \lambda + \frac{1}{2} \log 2\pi e + \log R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} + \log 4.$$

Thus, for all $\rho > \rho_u := \max\{\frac{\sqrt{\pi}}{R\sqrt{2}}, \frac{\sqrt{\pi}}{R\sqrt{2}}(\lambda + \frac{1}{2}\log 2\pi e + \log R + \log 4)\}$,

$$\lim_{n\to\infty} \mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \geq R \,\middle|\, N_n(Z_{0,n}) > 0 \right) = 0.$$

This completes the proofs of (i) and (ii).

Now, again by Markov's inequality and (5.17),

$$\mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \leq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0 \right)$$
$$\lesssim \mathbb{E}[N_n(Z_{0,n} \cap B_n(n^{\frac{3}{2}-\alpha} R))] \frac{\Gamma(2n+1)}{\Gamma(n+1)^2}.$$

By (2.7) and (5.16), for $\rho < \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \leq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0 \right)$$
$$\leq \lambda + \frac{1}{2} \log 2\pi e + \log R - \frac{\sqrt{2}\rho R}{\sqrt{\pi}} + \log 4.$$

Thus, if $R < (4e^{\lambda}\sqrt{2\pi})^{-1}$ then for all $\rho < \frac{\sqrt{\pi}}{R\sqrt{2}}$,

$$\lim_{n\to\infty} \mathbb{P}_n^0 \left(\max_{x_i \in N_n \cap Z_{0,n}} |x_i| \leq n^{\frac{3}{2}-\alpha} R \,\middle|\, N_n(Z_{0,n}) > 0 \right) = 0.$$

This completes the proofs of (iii) and (iv). □

Remark 5.3.1. To separate data more efficiently, we would ideally like to assume a relationship between λ_n and γ_n such that the cells of the tessellation contain more than one point with high probability. The assumption that $\lim_{n\to\infty} \mathbb{E}_n^{0,!}[N_n(Z_{0,n})] = \infty$ does not ensure that $\lim_{n\to\infty} \mathbb{P}_n^0(N_n(Z_{0,n}) > 1) = 1$, however. The second moment method does not help, since this lower bound goes to zero as n goes to infinity for all λ_n, and thus it remains an open question what scaling of λ_n and γ_n is needed to ensure $\lim_{n\to\infty} \mathbb{P}_n^0(N_n(Z_{0,n}) > 1) = 1$.

5.4 Summary

Our results can be summarized in terms of phenomena that successively take place when increasing ρ for a given α and incrementing α, when parameterizing the intensity of hyperplanes as ρn^α. As soon as α is positive, one finds a data arbitrarily close and encoded differently w.h.p. In addition, a displacement of order $\sqrt{n}$ in a random direction leads to an encoding which is different w.h.p. When moving to $\alpha > \frac{1}{2}$, a displacement of order one in a random direction leads to an encoding which is different w.h.p. Further phenomena start appearing when $\alpha = 1$ (Shannon regime). When increasing ρ, one first gets a small volume for the typical cell, and then for the zero cell w.h.p. At this scale, one can also control distortion, namely the fact that the most distant data point encoded like the typical data is at distance at most $\sqrt{n}R$ w.h.p. by a proper choice of ρ with ρ arbitrarily small as R grows. A new phenomenon appears at $\alpha = \frac{3}{2}$ where a sufficiently large ρ guarantees that the most distant data point encoded like the typical data is at distance at most R w.h.p. The following table illustrates how and when this collection of phenomena take place when increasing α and ρ .

Table 5.1: Labels for different separation and distortion criteria

Measure of good separation/low distortion	Label
$\mathbb{P}(X_n(\mathcal{F}_{B_n(r)}) = 0)$	A
$\mathbb{P}(Y_n \in Z_{0,n})$ (Gaussian displ.)	B
$\mathbb{P}(Y_{n,\delta} \in Z_{0,n})$ (Displ. at dist. δ)	C
$\mathbb{E}[V_n(Z)]$	D
$\mathbb{E}[V_n(Z_{0,n})]$	E
$\mathbb{P}(R_M(Z_{0,n}) > r)$	F

Table 5.2: Limit of separation and distortion metrics as $n \to \infty$ for different values of α and ρ when $\gamma_n \sim \rho n^\alpha$.

	$\alpha = 0$, $\rho > 0$	$\alpha \in (0, \frac{1}{2})$, $\rho > 0$	$\alpha = \frac{1}{2}$, $\rho > 0$	$\alpha \in (\frac{1}{2}, 1)$, $\rho > 0$
A	$e^{-\rho r}$	0	0	0
B	$e^{-\sqrt{\frac{2}{\pi}}\rho\sigma}$	0	0	0
C	1	1	$e^{-\sqrt{\frac{2}{\pi}}\rho\delta}$	0
D	∞	∞	∞	∞
E	∞	∞	∞	∞
F	1	1	1	1

	$\alpha = 1$			$\alpha \in (1, \frac{3}{2})$	$\alpha = \frac{3}{2}$		
		$\rho = \frac{1}{\sqrt{e}}$	$\rho = \frac{\pi}{\sqrt{e}}$	$\rho > 0$		ρ_ℓ	ρ_u
A	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0
D	∞	0	0	0	0	0	0
E	∞	∞	0	0	0	0	0
F	1	1	1	1	1	open	0

Remark 5.4.1. In Table 5.2, the only distortion measure which was included is $\mathbb{P}(R_M(Z_{0,n}) > r)$, but as mentioned, we could also consider $\mathbb{P}(R_M(Z_{0,n}) > \sqrt{n}r)$, which follows the information theoretic Shannon regime discussed later in Section 5.6.2. In this case the threshold above which this probability is small in high dimensions is for $\alpha = 1$ and $\rho > \rho_u$, and by Remark 5.2.1, this is the scaling at which the centroids of the cells have intensity growing like $e^{n\lambda}$ with dimension n for some $\lambda \in \mathbb{R}$.

5.5 Dimension Reduction

If it is known beforehand that the data lie in a lower dimensional subspace of $\mathbb{R}^n$, then the number of random hyperplanes needed to encode it may be much less than was evaluated above. If the subspace is known, we can tessellate the subspace directly. But if only the dimension of the subspace known, then we can model the subspace containing the data as a uniform random subspace in $\mathbb{R}^n$ independent of X_n. Let $\mathcal{L}$ be a random subspace in $\mathbb{R}^n$ of dimension $m(n)$, independent of the hyperplane tessellation X. If we assume that the data all lie in $\mathcal{L}$, then instead of considering the zero cell Z_0 of X in $\mathbb{R}^n$, we can consider the zero cell $Z_0^{(\mathcal{L})}$ of the tessellation induced by the intersection of X with $\mathcal{L}$. By radial symmetry, we can just consider a fixed subspace L. It is known that $X \cap L$ is a Poisson hyperplane process with intensity measure

$$\Theta_L(\cdot) = \gamma_m \int_{\mathbb{S}_L} \int_{\mathbb{R}} 1\{tu + (u^\perp \cap L) \in \cdot\} dt \sigma_{m-1}(du),$$

where $\gamma_m = \frac{\omega_m \omega_{n+1}}{\omega_n \omega_{m+1}} \gamma$. In [39], the authors showed that

$$\mathbb{E}[V_m(Z_0 \cap L)] = \Gamma(m+1)\kappa_m \left(\frac{\pi \omega_n}{\gamma_n \omega_{n+1}} \right)^m, \tag{5.18}$$

and established the following results on higher moments:

$$\Gamma(m+1)^k \kappa_m^k \left(\frac{\pi \omega_n}{\gamma_n \omega_{n+1}}\right)^{km} \leq \mathbb{E}[V_m(Z_0 \cap L)^k] \leq \Gamma(2m+1)\kappa_m^k \left(\frac{\pi \omega_n}{\gamma_n \omega_{n+1}}\right)^{2m}. \tag{5.19}$$

Proposition 5.3.5 can be extended to this case:

Proposition 5.5.1. *Let $\mathcal{L}_n$ be a random subspace of $\mathbb{R}^n$ with dimension $m_n < n$ such that $m_n \to \infty$ as $n \to \infty$. Let X_n be a stationary and isotropic Poisson hyperplane process in $\mathbb{R}^n$ with intensity γ_n. Then, if $\gamma_n \sim \rho\sqrt{m_n n}$ for some fixed $\rho > 0$,*

$$\lim_{n\to\infty} \mathbb{E}[V_{m_n}(Z_{0,n} \cap \mathcal{L}_n)] = \begin{cases} 0, & \rho > \frac{\pi}{\sqrt{e}} \\ 1, & \rho < \frac{\pi}{\sqrt{e}}. \end{cases}$$

Similarly, Theorem 5.3.8 can be extended to:

Proposition 5.5.2. *Let $\mathcal{L}_n$ be a random subspace of $\mathbb{R}^n$ with dimension $m_n < n$ such that $m_n \to \infty$ as $n \to \infty$. Let X_n be a stationary and isotropic Poisson hyperplane process in $\mathbb{R}^n$ with intensity γ_n, and let $R > 0$ Then, if $\gamma_n \sim \rho n^{\alpha-1} m_n$ as $n \to \infty$, then there exists ρ_u such that for all $\rho > \rho_u$,*

$$\lim_{n\to\infty} \mathbb{P}\left(R_M(Z_{0,n} \cap \mathcal{L}_n) \geq n^{\frac{3}{2}-\alpha} R\right) = 0,$$

and there exists ρ_ℓ such that for all $\rho < \rho_\ell$,

$$\lim_{n\to\infty} \mathbb{P}\left(R_M(Z_{0,n} \cap \mathcal{L}_n) \leq n^{\frac{3}{2}-\alpha} R\right) = 0.$$

5.6 Comments

5.6.1 One-Bit Compressed Sensing Comments

In this paper, the compression of the data can be considered as a sequence of one-bit measurements, where each bit gives the side of a random hyperplane the data lies on. This is the paradigm of one-bit compressed sensing, and the aim of this section is to further connect this theory with the results in this paper.

Traditional compressed sensing is concerned with recovering a signal $x \in \mathbb{R}^n$ from a measurement vector $y = Ax \in \mathbb{R}^m$, where A is some $m \times n$ measurement matrix ($m \leq n$). The goal is to find the smallest m such that the signal x can be recovered from y. If m is less than n, this problem is ill-posed. However, Tao and Candes [19] showed that under the assumption that x is s-sparse, i.e. $|\text{supp}(x)| \leq s$, x can be recovered from $y = Ax$, where A is Gaussian matrix, with $m = O\left(s \log \frac{n}{s}\right)$ measurements.

In general the measurement vector in this set-up requires infinite bit precision. One-bit compressed sensing was introduced by Baraniuk and Boufounos in [13] and aims to recover x from the most severely quantized measurements possible: $y = \text{sign}(Ax)$. This contains just one-bit per measurement. Note that taking these measurements loses all information regarding the norm of x, so we can only hope to recover $x/|x|$. The goal is then to find a $x^* \in S^{n-1}$ such that $|x/|x| - x^*| < \delta$ for some error δ. To reconstruct the signal from m measurements, Plan and Vershynin showed that one can solve the convex optimization

$$\min \|x\|_1 \qquad \text{subject to} \qquad \operatorname{sign}(Ax) \equiv y \text{ and } \|Ax\|_1 = m, \tag{5.20}$$

where A is a $m \times n$ matrix with i.i.d. standard Gaussian entries, see Theorem 1.1 in [69]. The original signal is recovered with small error if it can be guaranteed that the reconstructed signal is close in Euclidean distance to the original signal with high probability. Plan and Vershynin showed this error guarantee specifically for sparse or almost sparse signals using the following two results. First, they showed that if the original signal is effectively sparse (see Remark 1 in [69]), the signal returned from the optimization (5.20) will also be effectively sparse. Second they use the fact that there is a tessellation of the signal space $S^{n-1} \cap \Sigma_s$, where $\Sigma_s := \{s - \text{sparse signals}\}$, with $m = O(s \log^2(n/s))$ hyperplanes where all cells in the tessellation will have diameter at most δ, i.e., all sparse signals within a cell of the tessellation will be with δ-distance apart from eachother. Thus, the recovered signal will be within distance δ of the original signal with high probability. In fact, they showed a more general result in [70] that, for a subset $K \subseteq S^{n-1}$, all cells of a tessellation with $m \geq C\delta^{-6}\omega(K)^2$ hyperplanes will have diameter at most δ with probability as least $1 - 3e^{-c\delta^4 m}$, where $\omega(K)$ is the Gaussian mean width of the set K.

Some recent work has shown that the same geometric techniques can be used to recover a signal x, both direction and magnitude, if it is known that $|x| \leq R < \infty$. Instead of linear hyperplanes tessellating $K \subset \mathbb{S}^{n-1}$, consider a bounded set $K \subset \mathbb{R}^n$ and tessellate it with affine hyperplanes with normal vectors a_i and translations from the origin t_i. It was shown in [7] that a s-sparse signal x with $|x| \leq R$ can be recovered with measurements of the form

$$y_i = \text{sign}(\langle a_i, x\rangle - t_i), i = 1, ..., m, \tag{5.21}$$

where $t_1, ..., t_m \sim \mathcal{N}(0, R^2)$ are independent of $a_1, ..., a_m$. It is proved that the following program recovers the signal with small error:

$$\text{argmin}\|z\|_1 \qquad \text{subject to} \qquad |z| \leq R \text{ and } y_i(\langle a_i, z\rangle - t_i) \geq 0, \qquad \forall i = 1, ..., m. \tag{5.22}$$

More specifically, Theorem 2 in [7] states that with probability at least $1 - 3\exp(-c\delta^4 m)$, the following holds for all $x \in B_n(R) \cap \Sigma_s$: For $n \geq 2m$ and $m \geq C\delta^{-4} s \log(n/s)$, and for y obtained from the measurement model (5.21), the solution x^* to the program (5.22) satisfies $|x - x^*| \leq \delta R$.

Also, Knudson et al. [48] showed that if t is a Gaussian vector with variance depending on R, x can be recovered if $|x| \leq R$ by lifting to one dimension higher and using the program (5.20). They also showed you can estimate the magnitude (but not direction) of a signal x in an annulus $r \leq |x| \leq R$ up to error δ with $m \gtrsim R^4 r^{-2} \delta^{-2}$ measurements from evaluating the inverse Gaussian error function.

If we remove the norm constraint on the signal, one can use a stationary and isotropic hyperplane tessellation to obtain an infinite sequence of one-bit measurements encoding the signal. Instead of minimizing the number of hyperplanes, the intensity of hyperplanes is minimized, as done throughout this paper for the various separation/distortion metrics. The encoding scheme corresponding to a stationary and isotropic Poisson hyperplane tessellation is given as follows. Letting $\{u_i\}_{i\in\mathbb{Z}}$ be an i.i.d sequence of normal Gaussian random vectors in $\mathbb{R}^n$, and $\{t_i\}_{i\in\mathbb{Z}}$ be the support of a Poisson point process of intensity γ in $\mathbb{R}$, then the encoding is given by the one-bit measurements

$$y_i = \operatorname{sign}\left(\langle u_i/|u_i|, x\rangle - t_i\right), \qquad i \in \mathbb{Z}.$$

The collection of hyperplanes $\{H(u_i, t_i)\}_{i\in\mathbb{Z}}$ tessellates all of $\mathbb{R}^n$ and forms a stationary and isotropic Poisson hyperplane process with intensity γ, and all data within a single cell of the tessellation have the same encoding. The results in the paper provide an analysis of the quality of the compression, in terms of theoretical error bounds on the separation of a typical signal from other signals or the distortion of a typical signal. These are based on some metric of the cell that a typical signal lies in, i.e., the zero cell by stationarity.

The paradigm of one-bit compressed sensing requires the ability to recover the original data given only its one-bit encoding. Given an encoding, if one can identify a member of the cell corresponding to this sequence of bits, one can use this as an approximation of the original data.

The convex optimization recovery technique used in the literature for the constrained norm case will return a signal x^* that is one of the vertices of the cell, and knowing that all cells have small diameters ensures that recovered signal is close the original. The analogous strategy for the Poisson hyperplane compression requires showing that the vertex of the zero cell that is furthest from the origin is close in Euclidean distance, and thus the measure of distortion needed to ensure signal recovery through this convex optimization strategy is Theorem 5.3.8. To ensure that the farthest vertex of the cell containing the original signal is within error distance δ the intensity of hyperplanes γ_n must be on the order of $n^{3/2}$.

An alternative method for reconstruction that returns a point of the cell more likely to be close to the typical signal would provide a more efficient compression. For example, if the reconstruction returns a uniformly distributed signal in the cell determined by the measurements using, for instance, the algorithm for finding an approximate uniform random point in a convex set in [24], this could be guaranteed to be close to the original signal with high probability using results from [66].

As seen later, a deterministic grid actually performs better than the isotropic Poisson hyperplane tessellation in the full dimensional case in the sense that a smaller constant ρ is needed to ensure that the furthest vertex, or a uniform random vector in the cell, is close with high probability. However, if the data is sparse, or somehow lower-dimensional, this may make the isotropic case more desirable. In the case of a deterministic grid, only in the best case scenario will the intersection of the tessellation with a random m-dimensional subspace be a m-dimensional grid. However, in the isotropic case, the intersection will always have the distribution of a m-dimensional isotropic hyperplane tessellation. A more complete analysis of the case of sparse and lower dimensional data is left for future work.

5.6.2 Information Theoretic Comments

The aim of this section is to connect the results of the present paper to classical information theory.

5.6.2.1 Channel Coding

Consider first channel coding. The additive noise channel features the transmission of codewords in $\mathbb{R}^n$ (n is referred to the block-length of the code) through a noisy channel. The white Gaussian noise special case is of the same nature as that considered in Proposition 5.3.3: each coordinate of a transmitted codeword is additively blurred by an independent $\mathcal{N}(0, \sigma^2)$ random variable.

In the viewpoint introduced by Poltyrev [71], the codebook is a stationary point process in $\mathbb{R}^n$ (e.g., a Poisson point process in the random coding case) and the decoding scheme consists in saying that the codeword c was transmitted if the received message is in the Voronoi cell of c. The latter is the maximal likelihood decoder. In the regime where the point process has intensity $e^{n\rho}$ for some $\rho \in \mathbb{R}$, there is a threshold for ρ below which the correct codeword is decoded with a probability tending to one as n tends to infinity, and above which the probability of error tends to 1 as n tends to infinity. In Shannon's channel coding theory, the codewords are constrained to satisfy some power constraint requiring that the Euclidean norm of a codeword be less than or equal to $\sqrt{nP}$, for some P which is the power per symbol. As shown in [3] (Lemma 2 and Theorem 7), the Poltyrev viewpoint can be connected to Shannon's channel coding theorem in the high signal to noise ratio case, namely when P tends to infinity. In particular the Shannon capacity then grows like $\frac{1}{2}\log(2\pi eP)$ when $P \to \infty$, and the Poltyrev capacity is what one gets asymptotically when subtracting $\frac{1}{2}\log(2\pi eP)$ from the Shannon capacity.

5.6.2.2 Loss-less One-bit Compression Source Coding

Consider now source coding, which is more directly related to the setting considered in the present paper. Consider a source with i.i.d. $N(0, \sigma^2)$ symbols. If there are n such symbols, with n (also called block-length) large, they lie in a ball of radius $\sqrt{n\sigma^2}$, which has volume about $e^{n\frac{1}{2}\log(2\pi e\sigma^2)}$. If one wants to represent in a loss-less way all typical sequences of this type by $2^{\beta n}$ binary compression sequences, namely all binary sequences of length βn, the volume per sequence should tend to 0. That is

$$e^{n\frac{1}{2}\log(2\pi e\sigma^2)}e^{-\beta n\log(2)}$$

should go to 0 when n tends to infinity. This shows that the best (smallest) compression rate β for such a signal is $\beta_c = \frac{1}{2}\log(2\pi e\sigma^2)/\log(2)$. This is sharp and generalizes to all sources with a well defined entropy rate. This is formalized in the source coding theorem.

In our case, we have no structure in the signal, which corresponds to letting σ^2 tend to ∞. The unconstrained setting developed in the present paper can hence be seen as an *analogue of the Poltyrev regime* for source coding. In addition, we focus on a *specific coding scheme* which is that of Poisson hyperplanes one-bit compression.

Before going down this path, let us discuss some questions related to coding in this one-bit compressive setting. (1) What is the codebook? A first natural answer consists in associating one codeword sampled at random to each cell, with the uniform sampling taking place in a conditionally independent way given the hyperplane tessellation. Another possibility is the center of the smallest ball containing the zero cell (the out-ball). A third one is the center of the largest ball contained in the zero cell (the in-ball). (2) What is the decoding algorithm? By this, we mean the way to retrieve the codeword, as defined above, from the sequence of bits characterizing the cell as described in Section 5.6.1.

For unconstrained one-bit data compression, the analogue of the Shannon threshold β_c is the density $\gamma_n = \rho n^\alpha$ of hyperplanes that separates the situations where the mean volume of the typical cell tends to 0 and infinity, respectively. As shown above, this critical density lies in the Shannon regime, namely for $\alpha = 1$. More precisely, if $\gamma_n = \rho n$, with $\rho < \rho_c = \frac{1}{\sqrt{e}}$, then this mean volume tends to infinity, whereas if $\rho > \rho_c$, then it tends to 0. In other words, for one-bit compressive sensing based on Poisson isotropic hyperplanes, the Palm-Shannon-Poltyrev source coding rate is $\alpha_c = 1$ and $\rho_c = \frac{1}{\sqrt{e}}$. The proposed name comes from the fact that one looks at the typical cell, with typicality defined in the Palm sense (e.g., with respect to the point process of centers of the out-balls). The threshold that separates the situations where the mean volume of zero cell tends to 0 and infinity, respectively, could be called the Feller-Shannon-Poltyrev threshold and is obtained for a density of hyperplanes with $\alpha_c = 1$ and $\rho_c = \frac{\pi}{\sqrt{e}}$. The proposed name comes from "Feller's paradox" which states that the interval of a stationary point process on $\mathbb{R}$ containing the origin is larger than the typical interval. The Feller-Shannon-Poltyrev rate is of the same order as the Palm-Shannon-Poltyrev one, but π times larger.

5.6.2.3 Lossy One-bit Compression Source Coding

In the classical lossy source coding case, one looks for a codebook such that the distortion between a signal and its encoding be less than or equal to D. The most common distortion constraint is that the signal be at Euclidean distance order less than or equal to $\sqrt{nD}$ from the sequence it is encoded by. The rate-distortion function then specifies what is the best coding rate ensuring this constraint.

The framework discussed in the present paper can be seen as some Poltyrev version of lossy source coding with codebooks corresponding to one-bit data compression. As for the loss-less case, the first dichotomy is whether one takes the Palm viewpoint of the typical codeword or the Feller viewpoint of the typical data point. The cell of the former is Z, whereas that containing the latter is Z_0. Let us first discuss the equivalent of the classical distortion defined above in the Palm case. If the codewords are the centers of the out-balls, then a natural definition of Palm distortion is in terms of the radius of the out-ball of the typical cell. For instance, in this case, the rate-distortion function would give the smallest intensity of hyperplanes $\gamma_n = \rho n^\alpha$ such that this radius is less than or equal to $\sqrt{n}R$, as a function of R. This Palm-Shannon-Poltyrev out-ball rate-distortion function is not known to the best of our knowledge. However, the Feller version of this problem is precisely solved by Theorems 5.3.10 and 5.3.8. For instance, in the case of Theorem 5.3.8, the parameters in question are $\alpha = 1$ and $\rho_u(R) = x_u \frac{\sqrt{\pi}}{R\sqrt{2}}$, with x_u the constant defined in the proof of the theorem. Hence the function $R \to n\rho_u(R)$ can be seen as the *rate-distortion function* for this version of the problem. Note that for this definition of distortion, lossy coding with a radius R large enough requires a smaller hyperplane intensity than that guaranteeing the Palm volume to go to zero (which can be seen as an analogue of loss-less coding): the exponent is the same, namely $\alpha = 1$, but the multiplicative constant $\rho(u)$ goes to 0 as R tends to infinity. As expected, relaxing the distortion constraint allows one to use smaller codes.

The paper also determines various other rate-separation functions of the Feller type. A first instance is the Feller-Shannon-Poltyrev in-ball function, which gives the smallest hyperplane intensity such that the closest data point not encoded in the same way as the origin lies at a distance at least δ. This last condition is equivalent to having the radius of the largest ball centered at the origin and contained in the zero cell being larger than or equal to δ. By the same arguments as in Proposition 5.3.1, the associated threshold is $\alpha_c = 0$.

If $\gamma_n = \rho$, the probability that this distance is at least δ is $\exp(-2\rho\delta)$. A second example is the Feller-Shannon-Poltyrev linear contact function, which gives the smallest hyperplane intensity such that the closest data point in some random direction and not encoded as the origin is at distance more than $\sqrt{nD}$. By the arguments of Proposition 5.3.2, the threshold is again $\alpha_c = 0$ and if $\gamma_n = \rho$, the probability that this distance is at least $\sqrt{nD}$ is $\exp(-\frac{\sqrt{2}}{\sqrt{\pi}}\rho D)$.

5.6.3 Why Isotropic Poisson Hyperplanes

We discuss here some mathematical reasons justifying the framework proposed here for a one-bit compression based on Poisson isotropic hyperplanes. Other natural options in the Poisson hyperplane framework are Poisson Manhattan hyperplanes, where all hyperplanes are orthogonal to the orthonormal basis of $\mathbb{R}^n$. An even simpler hyperplane system is the square one (referred to as the deterministic grid below). The following tables summarize the results available on basic quantities related to these tessellations, when the distance to the nearest hyperplane is the same in expectation. The results are proved at the end of the section.

For all criteria in Table 5.3, the Poisson isotropic setting outperforms the two other options. For the expected volume of the zero cell (first column), the isotropic Poisson tessellation is the best, i.e., has the smallest expected volume. This fact is the main justification of the use of this Poisson isotropic structure in the context of one-bit compression: this allows the code with the smallest volume of data encoded as the typical data, among all three options.

Table 5.3: Comparison of quantities for different tessellations with intensity γ in $\mathbb{R}^n$.

Type of tessellation	$\mathbb{E}[V(Z_0)]$	$\mathbb{E}[V(Z)]$	$\mathbb{P}(x \notin Z_0)$
Deterministic Grid	$\left(\frac{2n}{\gamma}\right)^n$	$\left(\frac{2n}{\gamma}\right)^n$	$1_{\{\lVert x\rVert_\infty \geq \frac{n}{\gamma}\}}$
Poisson Manhattan	$\left(\frac{2n}{\gamma}\right)^n$	$\frac{1}{\kappa_n}\left(\frac{n\kappa_n}{\gamma\kappa_{n-1}}\right)^n$	$1 - \exp(-\frac{\gamma}{n}\lVert x\rVert_1)$
Poisson Isotropic	$n!\kappa_n\left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^n$	$\frac{1}{\kappa_n}\left(\frac{n\kappa_n}{\gamma\kappa_{n-1}}\right)^n$	$1 - \exp\left(-\frac{2\gamma\kappa_{n-1}}{n\kappa_n}\lvert x\rvert\right)$

The Poisson isotropic setting is also better than the other two in terms of the probability of separation of the typical data from data point x. We see from the last column that isotropic Poisson hyperplanes outperforms the other two options orderwise: the thresholds for the latter have order $\alpha = 1$, whereas that of the former has order $\alpha = 1/2$ only.

In contrast, consider now a uniform random vector Y chosen in the zero cell and take as a distortion criterion the "norm" of Y, defined as $\mathbb{E}[|Y|^2]^{\frac{1}{2}}$. The deterministic grid has the smallest norm and the Poisson grid has the second smallest norm. From Proposition 4.1 in [66], the isotropic Poisson tessellation gives an upper bound of this norm, where the upper bound is larger than the other two cases. For the quantity R_M, or equivalently, the furthest vertex of the zero cell from the origin, the results are the same, with the deterministic grid performing better than the Poisson grid, and the isotropic Poisson tessellation having an upper bound greater than the other two cases, since $x_u \approx 3$. For both quantities to be small, the scaling with dimension n needed for γ is $n^{3/2}$ for all three tessellations.

Table 5.4: Comparison of quantities for different tessellations with intensity γ in $\mathbb{R}^n$.

Type of tessellation	$\mathbb{E}[\lvert Y\rvert^2]^{\frac{1}{2}}$	R_M
Deterministic Grid	$\frac{n^{3/2}}{\sqrt{3}\gamma}$	$\frac{n^{3/2}}{\gamma}$
Poisson Manhattan	$\frac{n^{3/2}}{\gamma}$	$\frac{\sqrt{7}n^{\frac{3}{2}}}{\sqrt{2}\gamma}$
Poisson Isotropic	$\lesssim \frac{\sqrt{\pi}n^{3/2}}{\sqrt{2}\gamma}$	$\lesssim x_u \frac{\sqrt{\pi}n^{3/2}}{\sqrt{2}\gamma}$

We now give the proofs.

To compute the norm of the uniform random vector in the zero cell of the deterministic grid, consider the fixed cube of width $\frac{2n}{\gamma}$. Let $Y_n \sim \text{Uniform}\left([-n/\gamma, n/\gamma]^n\right)$. Then, by the strong law of large numbers,

$$\frac{|Y_n|^2}{n} = \frac{\sum_{k=1}^n Y_{n,k}^2}{n} \to \mathbb{E}[Y_{n,1}^2],$$

as $n \to \infty$. Then, since $Y_{n,1} \sim \text{Uniform}([-n/\gamma, n/\gamma])$,

$$\mathbb{E}[Y_{n,1}^2] = \frac{1}{3}\left(\frac{n^2}{\gamma^2} - \frac{n^2}{\gamma^2} + \frac{n^2}{\gamma^2}\right) = \frac{n^2}{3\gamma^2}.$$

Thus, $|Y_n| \sim \frac{n^{3/2}}{\sqrt{3}\gamma}$, as $n \to \infty$. The other quantities are immediate.

The Poisson Manhattan tessellation is defined as follows. Let X be a Poisson hyperplane tessellation in $\mathbb{R}^n$ with intensity γ and directional distribution ϕ that has mass $\frac{1}{2n}$ on each positive and negative axis, i.e. the normal vectors of the hyperplanes are the usual basis directions $\pm e_1, ..., \pm e_n$. Since equal weight is placed on each direction, the normal vectors of the hyperplanes form independent Poisson point processes of intensity $\frac{\gamma}{n}$ on each axis.

For each $i = 1, \dots n$, let $N_i = \{T_k^i\}$ be the Poisson point process of intersection points on the $\pm e_i$ axis with the usual convention that $T_0^i \leq 0 < T_1^i$. Then, the zero cell Z_0 of X is defined as

$$Z_0 = \prod_{i=1}^{n} [T_0^i, T_1^i].$$

Note that the interval $[T_0^i, T_1^i]$ will not have an exponential distribution, since we are requiring that 0 is in the interval, biasing for larger intervals. We obtain the distribution of the length of the interval by using the Palm distributions of $\{N_i\}_{i=1}^n$. By Slivnyak's theorem, $\mathbb{P}_N = \mathbb{P}^0_{N-\delta_0}$, so the distribution of length of the interval is the same as

$$\mathbb{P}(T_1^i - T_0^i \in A) = \mathbb{P}^0(T_1^i + |T_{-1}^i| \in A).$$

Under $\mathbb{P}^0$, i.e. conditioned on $T_0 = 0$, T_1 and $|T_{-1}|$ are independent exponential random variables with parameter $\frac{\gamma}{n}$. Then, we first see that

$$\mathbb{E}(V_n(Z_0)) = \prod_{i=1}^{n} \mathbb{E}(T_1^i - T_0^i) = \mathbb{E}^0(T_1^1 + |T_{-1}^1|)^n = \left(\frac{2\gamma}{n}\right)^n.$$

Also, for Y such that conditioned on X, $Y \sim \text{Uniform}(Z_0)$, the law of large numbers implies that as $n \to \infty$,

$$\frac{|Y|^2}{n} = \frac{\sum_{i=1}^n Y_i^2}{n} \to \mathbb{E}[Y_1^2] \text{ a.s.}$$

Using the fact that $(Y_i | T_0^i, T_1^i) \sim \text{Uniform}([T_0^i, T_1^i])$, we have

$$\begin{aligned}\mathbb{E}[Y_i^2] &= \mathbb{E}[\mathbb{E}[Y_i^2 | T_0^i, T_1^i]] = \mathbb{E}\left[\frac{T_0^2 + T_0 T_1 + T_1^2}{3}\right] = \frac{1}{3}\left(\mathbb{E}T_0^2 - \mathbb{E}^0[|T_{-1}|]\mathbb{E}^0[T_1] + \mathbb{E}T_1^2\right) \\ &= \frac{1}{3}\left(\frac{2n^2}{\gamma^2} - \frac{n^2}{\gamma^2} + \frac{2n^2}{\gamma^2}\right) = \frac{n^2}{\gamma^2}.\end{aligned}$$

Thus, $|Y_n|^2 \sim \frac{n^{3/2}}{\gamma}$ as $n \to \infty$.

For the Poisson Manhattan, the quantity R_M is given by

$$R_M^2 = |(\max\{T_1^1, |T_0^1|\}, ..., \max\{T_1^n, |T_0^n|\}|^2 = \sum_{i=1}^{n} (\max\{T_1^n, |T_0^n|\}^2).$$

By the law of large numbers, as $n \to \infty$,

$$\frac{R_M^2}{n} \to \mathbb{E}[\max\{T_1^n, |T_0^n|\}^2], \text{ a.s.}$$

The distribution of $\max\{T_1, T_0\}$ is

$$\mathbb{P}(\max\{T_1, T_0\} \leq x) = \mathbb{P}^0(\max\{T_1, |T_{-1}|\} \leq x) = (1 - e^{-\frac{\gamma}{n}x})^2.$$

Then, using integration by parts,

$$\begin{aligned}
\mathbb{E}[\max\{T_1^n, |T_0^n|\}^2] &= \int_0^\infty 2x\mathbb{P}(\max\{T_1, |T_0|\} \geq x)dx \\
&= \int_0^\infty 2x(1 - (1 - e^{-\frac{\gamma}{n}x})^2)dx = \int_0^\infty 2x(1 - (1 - 2e^{-\frac{\gamma}{n}x} + e^{-\frac{2\gamma}{n}}))dx \\
&= \int_0^\infty 2x(2e^{-\frac{\gamma}{n}x} - e^{-\frac{2\gamma}{n}x})dx = \frac{4n^2}{\gamma^2} - \frac{n^2}{2\gamma^2} = \frac{7n^2}{2\gamma^2}.
\end{aligned}$$

Thus, R_M is concentrated near $\frac{\sqrt{7}n^{3/2}}{\sqrt{2}\gamma}$ for large n.

Chapter 6

Uniform Random Vectors in Zero Cells of Stationary Poisson Mosaics[1]

6.1 Introduction

Random mosaics, also called random tessellations, have long been studied in stochastic geometry and give rise to interesting classes of random polytopes. Recently, there has been more interest in high dimensional tessellations, partially due to applications in signal processing [69] and information theory [3]. For these applications, it is important to understand the asymptotic geometric properties of the polytopes induced by random tessellations, in order to decode and reconstruct signals with small error.

Some well-known classes of random mosaics are built from Poisson point processes, either in $\mathbb{R}^n$ or in the space of hyperplanes in $\mathbb{R}^n$. Statistics of the cells of these random mosaics have been well-studied, particularly in dimensions $n = 2$ and $n = 3$. See [75] and [18] for more background and further references. Some recent work has focused on high dimensional Poisson mosaics, particularly on the volume and shape of the zero cell and typical cell as dimension n tends to infinity ([2], [43], and [40]). The zero cell is the cell of

[1]This chapter is based on the following manuscript: *E. O'Reilly. Thin-shell concentration for zero cells of stationary Poisson mosaics. arXiv:1809.04134, September 2018.*

the tessellation containing the origin, and the distribution of the typical cell is obtained by averaging over all cells in a large bounded subset and then increasing this subset to the entire space. The volume of these cells has been studied in [2] and [43], as well as some analysis of their shape in high dimensions. For example, in [2], it is proved that the volume of the intersection of the typical cell of a Poisson-Voronoi tessellation with intensity λ and a co-centered ball of volume u tends to $\lambda^{-1}(1 - e^{-\lambda u})$ as the space dimension tends to infinity.

In this chapter, we aim to better understand the nature of the zero cell of stationary Poisson mosaics in high dimensions by considering the random vector that, conditioned on the random mosaic, is uniformly distributed in the centered zero cell. By "centered", we mean that an appropriately chosen center of the cell is located at the origin, for instance, the center of the largest ball contained within the cell. We will study to what extent the phenomenon of thin-shell concentration, as described in Section 2.4.1, occurs for this random vector for specific models. If this random vector is concentrated around its mean in high dimensions, then most of the volume of the zero cell of the random mosaic will be contained within a narrow annulus.

One motivation for the study of the norm of this random vector is data compression. Random mosaics can be used to compress data in $\mathbb{R}^n$ such that all data contained in the same cell of the tessellation will have the same encoding. This is the case, for instance, in one-bit compressed sensing using hyperplane tessellations, see [69] and [70]. Reconstructing the original data with small error requires that all data within the same cell of the tessellation are close together. The volume of the cell is not a useful metric in this case, since a very thin cell could have small volume and also contain signals that lie very far apart. The norm of the random vector studied in this paper is a more useful metric to ensure the mass of the cell does not lie far away from the center.

The distribution of the vector described above is shown to depend on the typical cell, as shown in Lemma 6.2.1, due to the fact that the distribution of the zero cell has a Radon-Nikodym derivative with respect to the distribution of the typical cell. Here, we restrict to studying two types of stationary random mosaics, a stationary Poisson-Voronoi mosaic and a stationary and isotropic Poisson hyperplane mosaic, since in both cases there exists an explicit representation for the distribution of the typical cell that allows for computations. Both of these random mosaics are isotropic, that is, their distribution is invariant under rotations about the origin. This implies that the random vector chosen uniformly from the centered zero cell will be radially symmetric. In the Poisson-Voronoi case, we show that this random vector is also log-concave, and thus satisfies the thin-shell estimate (2.9). Strong concentration inequalities are also obtained directly.

The main complementary results can be stated as follows. For each n, let X_n be a stationary random mosaic in $\mathbb{R}^n$ where the intensity of cell centroids is $e^{n\lambda_n}$ and assume $\lim_{n\to\infty} \lambda_n = \lambda \in \mathbb{R}$. Let Y_n denote a random vector in $\mathbb{R}^n$ such that, conditionally on X_n, Y_n is uniformly distributed in the centered zero cell of X_n. For the Poisson-Voronoi mosaic, we show that

$|Y_n|/\sqrt{n}$ concentrates to $e^{-\lambda}(2\pi e)^{-\frac{1}{2}}$ as the dimension n increases. Exponential rates of convergence are also computed, as shown in Theorem 6.3.3.

In the case of the zero cell of a Poisson hyperplane tessellation, we show there exists an interval (R_ℓ, R_u) such that $|Y_n|/\sqrt{n}$ will be contained in this interval with high probability in high dimensions. Rates of convergence are also computed in this case as shown in Theorem 6.4.2.

6.2 Preliminaries

An application of Theorem 2.3.2 gives the density of a vector uniformly sampled in the zero cell of a random mosaic.

Lemma 6.2.1. *Let X be a stationary random mosaic in $\mathbb{R}^n$ with zero cell Z_0 and typical cell Z with respect to the center function $c : \mathcal{C}' \to \mathbb{R}^n$ as previously defined. Let Y be a random vector in $\mathbb{R}^n$ such that conditioned on X,*

$$Y \sim \text{Uniform}(Z_0 - c(Z_0)).$$

Then, for all measurable g,

$$\mathbb{E}[g(Y)] = \int_{\mathbb{R}^n} g(x) \frac{\mathbb{P}(x \in Z)}{\mathbb{E}[V(Z)]} dx,$$

i.e., $Y \in \mathbb{R}^n$ has density $f_Y(x) = \frac{\mathbb{P}(x \in Z)}{\mathbb{E}[V(Z)]}$.

Proof. First, by the definition of Y,

$$\mathbb{E}[g(Y)] = \mathbb{E}\left[\mathbb{E}[g(Y)|Z_0]\right] = \mathbb{E}\left[\frac{1}{V(Z_0)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in Z_0 - c(Z_0)\}} dx\right].$$

Note that the function $f(\cdot) = \frac{1}{V(\cdot)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in \cdot - c(\cdot)\}} dx$ is invariant under translations, since for any $t \in \mathbb{R}$ and $K \in \mathcal{K}'$,

$$\begin{aligned} f(K+t) &= \frac{1}{V(K+t)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in K+t-c(K+t)\}} dx \\ &= \frac{1}{V(K)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in K+t-c(K)-t\}} dx \\ &= \frac{1}{V(K)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in K-c(K)\}} dx = f(K). \end{aligned}$$

Thus, by Theorem 2.3.2, and since $c(Z) = 0$,

$$\begin{aligned} \mathbb{E}[g(Y)] &= \mathbb{E}\left[\frac{1}{V(Z_0)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in Z_0 - c(Z_0)\}} dx\right] \\ &= \frac{1}{\mathbb{E}[V(Z)]} \mathbb{E}\left[V(Z) \frac{1}{V(Z)} \int_{\mathbb{R}^n} g(x) 1_{\{x \in Z - c(Z)\}} dx\right] \\ &= \mathbb{E}\left[\int_{\mathbb{R}^n} g(x) \frac{1_{\{x \in Z\}}}{\mathbb{E}[V(Z)]} dx\right]. \end{aligned}$$

Finally, applying Fubini's Theorem gives the result. □

6.3 Poisson-Voronoi Mosaic

A special type of random mosaic comes from the Voronoi cells of a Poisson point process in $\mathbb{R}^n$. Let N be a stationary Poisson point process with intensity λ and, for $x \in N$, define the Voronoi cell of N with center x by

$$C(x, N) := \{z \in R^n : |z - x| \leq |z - y| \text{ for all } y \in N\}.$$

The collection $X := \{C(x, N) : x \in N\}$ is a stationary random mosaic and is called the Poisson-Voronoi mosaic induced by N. The intensity λ of the underlying Poisson point process is the cell intensity of the induced mosaic.

First we show that for a Poisson-Voronoi mosaic X, the density of the random vector Y that is uniformly distributed in Z_0 conditioned on X, is log-concave, and we then compute the moments of its norm.

Lemma 6.3.1. *Let Z_0 be the zero cell of the stationary and isotropic Poisson-Voronoi tessellation associated to $N \sim PPP(\lambda)$ in $\mathbb{R}^n$. Define the random vector Y, such that conditioned on Z_0,*

$$Y \sim \text{Uniform}(Z_0 - c(Z_0)).$$

Then, Y has a log-concave density and for all $k \in \mathbb{N}$,

$$\mathbb{E}[|Y|^k] = \frac{\Gamma(1+\frac{k}{n})}{(\lambda\kappa_n)^{\frac{k}{n}}} = O\left(\lambda^{-\frac{k}{n}} n^{\frac{k}{2}}\right).$$

Proof. By Lemma 6.2.1, the density of Y is $\frac{\mathbb{P}(x\in Z)}{\mathbb{E}[\mathrm{V}(Z)]}$. By the fact that $\mathbb{E}[\mathrm{V}(Z)] = \frac{1}{\lambda}$ and by Slivnyak's theorem [75],

$$\frac{\mathbb{P}(x \in Z)}{\mathbb{E}[\mathrm{V}(Z)]} = \lambda\mathbb{P}^{0,!}(x \in Z_0) = \lambda\mathbb{P}(N(B(x,|x|)) = 0) = \lambda e^{-\lambda\kappa_n|x|^n},$$

and this is clearly log-concave. Thus, the density of Y is log-concave.

For the moments, switching to polar coordinates and using another change of variables ($y = \lambda\kappa_n r^n$) gives

$$\begin{aligned}
\mathbb{E}[|Y|^k] &= \lambda \int_{\mathbb{R}^n} |x|^k e^{-\lambda\kappa_n|x|^n} dx = \lambda n\kappa_n \int_0^\infty r^{n+k-1} e^{-\lambda\kappa_n r^n} dr \\
&= \lambda n\kappa_n \int_0^\infty \left(\frac{y}{\lambda\kappa_n}\right)^{1+\frac{k}{n}-\frac{1}{n}} e^{-y} \frac{1}{n\lambda\kappa_n} \left(\frac{y}{\lambda\kappa_n}\right)^{\frac{1}{n}-1} dy \\
&= (\lambda\kappa_n)^{-\frac{k}{n}} \int_0^\infty y^{\frac{k}{n}} e^{-y} dy = \frac{\Gamma(1+\frac{k}{n})}{(\lambda\kappa_n)^{\frac{k}{n}}}.
\end{aligned}$$

Then, by (2.7), as $n \to \infty$,

$$\mathbb{E}[|Y|^k] \sim \frac{n^{\frac{k}{2}}}{\lambda^{\frac{k}{n}}(2\pi e)^{\frac{k}{2}}}.$$

$\square$

The fact that Y has a radial and log-concave density already implies that $|Y|$ concentrates to

$$\mathbb{E}[|Y|^2]^{\frac{1}{2}} = \frac{\Gamma(1+\frac{2}{n})^{\frac{1}{2}}}{(\lambda\kappa_n)^{\frac{1}{n}}} \sim \frac{\sqrt{n}}{\lambda^{\frac{1}{n}}\sqrt{2\pi e}}, \tag{6.1}$$

in high dimensions by the thin-shell estimate (2.9). However, we can prove strong concentration inequalities by direct computation.

Theorem 6.3.2. *Let X be a stationary Poisson-Voronoi mosaic in $\mathbb{R}^n$ with intensity λ and Y a random vector such that, conditioned on X, $Y \sim \mathrm{Uniform}(Z_0 - c(Z_0))$. Let $\sigma^2 = \mathbb{E}|Y|^2$. Then, there exists $c > 0$ such that for all $t > 0$,*

$$\mathbb{P}\left(|Y| \geq (1+t)\sigma\right) \leq e^{-e^{cn\ln(1+t)}},$$

and for all $t \in (0,1)$ and $n \geq 2$,

$$\mathbb{P}\left(|Y| \leq (1-t)\sigma\right) \leq e^{n\ln(1-t)}.$$

Proof. By Lemma 6.3.1,

$$\begin{aligned}\mathbb{P}(|Y| \leq R) &= \lambda \int_{B(R)} e^{-\lambda\kappa_n|x|^n} dx = \lambda n \kappa_n \int_0^R r^{n-1} e^{-\lambda\kappa_n r^n} dr \\ &= \int_0^{\lambda\kappa_n R^n} e^{-y} dy = 1 - e^{-\lambda\kappa_n R^n},\end{aligned}$$

and by (6.1),

$$\mathbb{P}(|Y| \leq (1-t)\sigma) = 1 - e^{-\Gamma(1+\frac{2}{n})^{\frac{n}{2}}(1-t)^n}.$$

By the inequality $1 - e^{-x} \leq x$ for all $x \geq 0$ and the assumption $n \geq 2$,

$$\mathbb{P}(|Y| \leq (1-t)\sigma) \leq \Gamma\left(1+\frac{2}{n}\right)^{\frac{n}{2}}(1-t)^n \leq \Gamma(2)^{\frac{n}{2}} e^{n\ln(1-t)} = e^{n\ln(1-t)}.$$

Similarly, we have

$$\mathbb{P}(|Y| \geq (1+t)\sigma) = \int_{\lambda\kappa_n R^n}^{\infty} e^{-y} dy = e^{-\Gamma(1+\frac{2}{n})^{\frac{n}{2}}(1+t)^n}.$$

Then, by the property $\Gamma(s+1) = s\Gamma(s)$ and the inequality $1+s \leq e^s$ for $s > 0$,

$$\Gamma\left(1+\frac{2}{n}\right)^{\frac{n}{2}} = \left(\frac{\Gamma(2+\frac{2}{n})}{(1+\frac{2}{n})}\right)^{\frac{n}{2}} \geq \frac{1}{e}.$$

Hence, for $c = e^{-1}$,

$$\mathbb{P}(|Y| \geq (1+t)\sigma) \leq e^{-e^{cn\ln(1+t)}}.$$

□

Considering a sequence of these vectors in increasing dimensions, we obtain the following threshold result when the cell intensity of the random mosaics grows exponentially with dimension.

Theorem 6.3.3. *Let $Y_n \sim \mathrm{Uniform}(Z_{0,n} - c(Z_{0,n}))$, where $Z_{0,n}$ is the zero cell of a stationary Poisson-Voronoi tessellation in $\mathbb{R}^n$ with intensity $e^{n\lambda_n}$. Assume $\lim_{n\to\infty} \lambda_n = \lambda \in \mathbb{R}$. Then,*

$$\lim_{n\to\infty} \mathbb{P}(|Y_n| \leq \sqrt{n}R) = \begin{cases} 0, & R < e^{-\lambda}(2\pi e)^{-\frac{1}{2}} \\ 1, & R > e^{-\lambda}(2\pi e)^{-\frac{1}{2}} \end{cases}.$$

For $R < e^{-\lambda}(2\pi e)^{-\frac{1}{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(|Y_n| \le \sqrt{n}R) = \lambda + \frac{1}{2}\ln(2\pi e) + \ln R,$$

and for $R > e^{-\lambda}(2\pi e)^{-\frac{1}{2}}$,

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(-\frac{1}{n}\ln \mathbb{P}(|Y_n| \ge \sqrt{n}R)\right) = \lambda + \frac{1}{2}\ln(2\pi e) + \ln R.$$

Proof. Switching to polar coordinates and then using another change of variables gives

$$\mathbb{P}(|Y_n| \le \sqrt{n}R) = e^{n\lambda_n} \int_{B(\sqrt{n}R)} e^{-e^{n\lambda_n}\kappa_n |x|^n} dx = e^{n\lambda_n} n\kappa_n \int_0^{\sqrt{n}R} r^{n-1} e^{-e^{n\lambda_n}\kappa_n r^n} dr$$
$$= \int_0^{e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n} e^{-y} dy = 1 - e^{-e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n}.$$

For $R < e^{-\lambda}(2\pi e)^{-\frac{1}{2}}$, by (2.7),

$$e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n \sim \frac{1}{\sqrt{\pi n}}\left(e^{\lambda_n}(2\pi e)^{\frac{1}{2}}R\right)^n \to 0, \text{ as } n \to \infty.$$

implying that $\mathbb{P}(|Y_n| \le \sqrt{n}R) \to 0$ as $n \to \infty$. This also implies that there is an $\alpha > 0$ such that for all n large enough,

$$\alpha(e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n) \le 1 - e^{-e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n} \le e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n.$$

Thus,

$$\lim_{n\to\infty} -\frac{1}{n} \ln \mathbb{P}(|Y_n| \le \sqrt{n}R) = \lambda + \frac{1}{2}\ln(2\pi e) + \ln R.$$

Also, similarly to above, for $R > e^{-\lambda}(2\pi e)^{-\frac{1}{2}}$,

$$\mathbb{P}(|Y_n| \ge \sqrt{n}R) = \int_{e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n}^{\infty} e^{-y} dy = e^{-e^{n\lambda_n}\kappa_n(\sqrt{n}R)^n} \to 0,$$

as $n \to \infty$, and

$$\lim_{n\to\infty} \frac{1}{n} \ln \left(-\frac{1}{n} \ln \mathbb{P}(|X| \geq \sqrt{n}R) \right) = \lambda + \frac{1}{2} \ln(2\pi e) + \ln R.$$

$\square$

6.4 Poisson Hyperplane Mosaic

The second type of random mosaic we consider is the mosaic induced by a stationary and isotropic Poisson hyperplane process X in $\mathbb{R}^n$. Recall that a hyperplane process in $\mathbb{R}^n$ is a point process in the space of $n-1$ dimensional affine subspaces in $\mathbb{R}^n$, denoted by $\mathcal{H}^n$.

In the case where the stationary mosaic is induced by a stationary and isotropic Poisson hyperplane process, we first have the following proposition. For $C \in \mathcal{K}'$, let $c(C)$ be the center of the inball of C.

Proposition 6.4.1. *Let X be a stationary and isotropic Poisson hyperplane mosaic in $\mathbb{R}^n$ with cell intensity λ. Let Y be the random vector such that, conditional on X,*

$$Y \sim \text{Uniform}(Z_0 - c(Z_0)).$$

Then, for all $k \geq 0$,

$$\mathbb{E}[|Y|^k] \leq \frac{\Gamma(n+k+1)}{\Gamma(n+1)2^k} \left(\frac{\kappa_n}{\lambda}\right)^{k/n}.$$

Next, we consider a sequence of these random vectors in increasing dimensions n, and obtain the following result.

Theorem 6.4.2. *For each n, let X_n be a stationary and isotropic Poisson hyperplane process with cell intensity $e^{n\lambda_n}$. Assume $\lim_{n\to\infty} \lambda_n = \lambda \in \mathbb{R}$. Let $Z_{0,n}$ be the zero cell of X_n, and define the random vectors Y_n such that, conditional on X_n,*

$$Y_n \sim \text{Uniform}(Z_{0,n} - c(Z_{0,n})).$$

Then, for all $R > e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$,

$$\lim_{n\to\infty} \mathbb{P}(|Y_n| \geq \sqrt{n}R) = 0,$$

and there is a R_ℓ such that $0 < R_\ell < e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$ and for all $R < R_\ell$,

$$\lim_{n\to\infty} \mathbb{P}(|Y_n| \leq \sqrt{n}R) = 0.$$

Also, for $R > e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(|Y_n| \geq \sqrt{n}R) \leq \lambda + \frac{1}{2}\ln\left(\frac{2e}{\pi}\right) + \ln R - \frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}},$$

and for $R < e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$,

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(|Y_n| \leq \sqrt{n}R) \leq \lambda + \frac{1}{2}\ln(2\pi e) + \ln R - \frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}} - \ln 2.$$

Remark 6.4.1. The lower bound R_ℓ is the radius satisfying

$$\lambda + \frac{1}{2}\ln(2\pi e) + \ln R_\ell - \frac{e^\lambda R_\ell\sqrt{2}}{\sqrt{\pi e}} - \ln 2 = 0.$$

Then, for all $R > 0$,

$$\lambda + \frac{1}{2}\ln(2\pi e) + \ln R - \frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}} - \ln 2 < \lambda + \frac{1}{2}\ln(2\pi e) + \ln R,$$

and thus, $\lambda + \frac{1}{2}\ln(2\pi e) + \ln R_\ell > 0$, implying that

$$R_\ell > e^{-\lambda}(2\pi e)^{-\frac{1}{2}}.$$

This implies that that vector Y_n chosen with respect to the Poisson-Voronoi zero cell will have a smaller norm in high dimensions than the vector Y_n chosen with respect to the zero cell of the Poisson hyperplane mosaic.

Remark 6.4.2. The assumption on λ_n can be generalized to $\lambda_n \sim e^{n\lambda}n^{n\alpha}$ for some $\alpha \in \mathbb{R}$. By (2.4), this implies that the scaling for the intensity of hyperplanes is $\gamma_n = O(n^{\alpha+1})$. Then a similar result holds with the probabilities

$$\mathbb{P}(|Y_n| \leq R_n) \text{ and } \mathbb{P}(|Y_n| \geq R_n),$$

where $R_n = Rn^{\frac{1}{2}-\alpha}$. This requirement gives two special cases: $\lambda_n \sim e^{n\lambda}$, $\gamma_n = O(n)$, $R_n = O(\sqrt{n})$ and $\lambda_n \sim e^{n\lambda}n^{\frac{n}{2}}$, $\gamma_n = O(n^{3/2})$, $R_n = O(1)$.

Before proving the theorem, recall the following special functions. The beta function is defined by

$$B(x,y) := \int_0^1 t^{x-1}(1-t)^{y-1}dt.$$

The incomplete beta function is defined as $B(x;a,b) := \int_0^x t^{a-1}(1-t)^{b-1}dt$, and the regularized incomplete beta function is

$$I_x(a,b) := \frac{B(x;a,b)}{B(a,b)}.$$

Recall that the Gamma function is defined by $\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}dt$, and we define the upper and lower incomplete gamma functions by

$$\Gamma(x,R) := \int_R^\infty t^{x-1}e^{-t}dt, \text{ and } \Gamma_\ell(x,R) := \int_0^R t^{x-1}e^{-t}dt,$$

respectively. The following series of lemmas are needed before proving the results.

Lemma 6.4.3. *Let X be a stationary and isotropic Poisson hyperplane process with intensity γ. Then, letting $[0,x]$ denote the line segment between 0 and the point x,*

$$\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right) = \gamma|x|\left[\frac{2\kappa_{n-1}}{n\kappa_n}\left(1-\frac{r^2}{|x|^2}\right)^{\frac{n-1}{2}} - \frac{r}{|x|}I_{1-\frac{r^2}{|x|^2}}\left(\frac{n-1}{2},\frac{1}{2}\right)\right]1_{\{|x|\geq r\}}.$$

Proof. Note that if $r > |x|$, then a hyperplane cannot hit $[0,x]$ and not hit the open ball $B(r)$ at the same time and thus $\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right)$ is zero. If $r \leq |x|$, then by Theorem 2.3.3,

$$\begin{aligned}
\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right) &= 2\gamma\int_{S^{n-1}}\int_0^\infty 1_{\{H(u,t)\cap[0,x]\neq 0\}}1_{\{H(u,t)\cap B(r)=\emptyset\}}dt\sigma_{n-1}(du) \\
&= 2\gamma\int_{S^{n-1}}\int_0^\infty 1_{\{r\leq t<\langle x,u\rangle_+\}}1_{\{r\leq\langle x,u\rangle\}}dt\sigma_{n-1}(du) \\
&= 2\gamma\int_{\{v\in S^{n-1}:\langle v,x\rangle\geq r\}}(\langle x,u\rangle_+ - r)\sigma_{n-1}(du) \\
&= 2\gamma|x|\int_{\{v\in S^{n-1}:\langle v,x\rangle\geq r\}}\left\langle\frac{x}{|x|},u\right\rangle\sigma_{n-1}(du) \\
&\qquad - 2r\gamma\sigma_{n-1}(\{v\in S^{n-1}:\langle v,x\rangle_+\geq r\}),
\end{aligned}$$

where $a_+ = \max\{a,0\}$. To compute the first integral, first note that the integral does not depend on the direction of x, only on the norm $|x|$. We can then assume $x = |x|e_n$, where $e_n = (0, ..., 0, 1)$, and

$$\begin{aligned}
\int_{\{v\in S^{n-1}:\langle v,|x|e_n\rangle\geq r\}} &\langle e_n, u\rangle_+ \sigma_{n-1}(du) = \int_{\{v\in S^{n-1}:v_n\geq \frac{r}{|x|}\}} u_n\sigma_{n-1}(du) \\
&= \frac{\omega_{n-2}}{\omega_{n-1}} \int_{\frac{r}{|x|}}^{1} \int_{\mathbb{S}^{n-2}} t(1-t^2)^{\frac{n-3}{2}} \sigma_{n-2}(du)dt \\
&= \frac{(n-1)\kappa_{n-1}}{n\kappa_n} \int_{\frac{r}{|x|}}^{1} t(1-t^2)^{\frac{n-3}{2}} dt = \frac{(n-1)\kappa_{n-1}}{n\kappa_n} \int_0^{1-\frac{r^2}{|x|^2}} \frac{1}{2} s^{\frac{n-3}{2}} ds \\
&= \frac{(n-1)\kappa_{n-1}}{2n\kappa_n} \frac{2}{n-1} \left(1 - \frac{r^2}{|x|^2}\right)^{\frac{n-1}{2}} = \frac{\kappa_{n-1}}{n\kappa_n} \left(1 - \frac{r^2}{|x|^2}\right)^{\frac{n-1}{2}}.
\end{aligned}$$

The fractional area of a spherical cap is given by

$$\sigma_{n-1}(\{v \in S^{n-1} : \langle v, x\rangle_+ \geq r\}) = \frac{1}{2} I_{1-\frac{r^2}{|x|^2}} \left(\frac{n-1}{2}, \frac{1}{2}\right).$$

Then,

$$\begin{aligned}
\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right) &= \left[2\gamma|x|\frac{\kappa_{n-1}}{n\kappa_n}\left(1 - \frac{r^2}{|x|^2}\right)^{\frac{n-1}{2}} - r\gamma I_{1-\frac{r^2}{|x|^2}}\left(\frac{n-1}{2}, \frac{1}{2}\right)\right] 1_{\{|x|\geq r\}} \\
&= \gamma|x| \left[\frac{2\kappa_{n-1}}{n\kappa_n}\left(1 - \frac{r^2}{|x|^2}\right)^{\frac{n-1}{2}} - \frac{r}{|x|} I_{1-\frac{r^2}{|x|^2}}\left(\frac{n-1}{2}, \frac{1}{2}\right)\right] 1_{\{|x|\geq r\}}.
\end{aligned}$$

$\square$

Lemma 6.4.4. *Let Z_0 be the zero cell of a stationary and isotropic Poisson hyperplane tessellation with intensity γ in $\mathbb{R}^n$. Conditioned on Z_0, let $Y \sim$* Uniform$(Z_0 - c(Z_0))$*. Then, for $R > 0$,*

$$\mathbb{P}(|Y| \geq R) \leq \frac{\Gamma_u\left(n+1, 2\gamma R\frac{\kappa_{n-1}}{n\kappa_n}\right)}{\Gamma(n+1)},$$

and

$$\mathbb{P}(|Y| \leq R) \leq \frac{n\kappa_n^2}{4^n}\left(\frac{n\kappa_n}{2\kappa_{n-1}}\right)\left[\Gamma_\ell\left(n+1, 2\gamma R\frac{\kappa_{n-1}}{n\kappa_n}\right) + \Gamma(n)\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^{n+1}\right].$$

Proof. By Lemma 6.2.1, the density of Y is

$$f_Y(x) = \frac{\mathbb{P}(x \in Z)}{\mathbb{E}\left[V(Z)\right]}.$$

Using the representation of Z in Theorem 2.3.5,

$$\mathbb{P}(x \in Z) = \frac{\gamma^{n+1}}{\lambda(n+1)} \int_0^\infty \int_{(S^{n-1})^{n+1}} e^{-2\gamma r} \mathbb{P}\left(x \in \bigcap_{H \in X \cap \mathcal{F}^{B(r)}} H_0^+ \cap \bigcap_{j=0}^{n} H^-(u_j, r)\right)$$
$$\cdot \triangle_n(u_0, ..., u_n) 1_P(u_0, ..., u_n) \prod_{i=0}^{n} \phi(du_i) dr.$$

First, we see that

$$\mathbb{P}\left(x \in \bigcap_{H \in X \cap \mathcal{F}^{B(r)}} H_0^+ \cap \bigcap_{j=0}^{n} H^-(u_j, r)\right) = \prod_{j=0}^{n} 1_{\{x \in H^-(u_j,r)\}} \mathbb{P}\left(X\left(\mathcal{F}_{[0,x]}^{B(r)}\right) = 0\right)$$
$$= \prod_{j=0}^{n} 1_{\{x \in H^-(u_j,r)\}} e^{-\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right)}.$$

Then,

$$\mathbb{P}(|Y| \geq R) = \int_{B(R)^c} \frac{\mathbb{P}(x \in Z)}{\mathbb{E}[V(Z)]} dx$$
$$= \int_{B(R)^c} \frac{\gamma^{n+1}}{(n+1)} \int_0^\infty \int_{(S^{n-1})^{n+1}} e^{-2\gamma r} \mathbb{P}\left(x \in \bigcap_{H \in X \cap \mathcal{F}^{B(r)}} H_0^+ \cap \bigcap_{j=0}^{n} H^-(u_j, r)\right)$$
$$\cdot \triangle_n(u_0, ..., u_n) 1_P(u_0, ..., u_n) \prod_{i=0}^{n} \sigma_{n-1}(du_i) dr$$
$$= \frac{\gamma^{n+1}}{(n+1)} \int_{B(R)^c} \int_0^\infty \int_{(S^{n-1})^{n+1}} e^{-2\gamma r} e^{-\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right)} \triangle_n 1_P \prod_{i=0}^{n} 1_{\{x \in H^-(u_i,r)\}} \sigma_{n-1}(du_i) dr dx$$
$$= \frac{\gamma^{n+1}}{(n+1)} \int_{B(R)^c} \int_0^\infty e^{-2\gamma r} e^{-\Theta\left(\mathcal{F}_{[0,x]}^{B(r)}\right)} \int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{i=0}^{n} 1_{\{\langle x, u_i \rangle \leq r\}} \sigma_{n-1}(du_i) dr dx.$$

Making the change of variables $t = \frac{r}{|x|}$, observing that the innermost integral does not depend on the direction of x, and using Fubini's Theorem,

$$\begin{aligned}
&\mathbb{P}(|Y| \geq R) \\
&= \frac{\gamma^{n+1}}{(n+1)} \int_{B(R)^c} |x| \int_0^\infty e^{-2\gamma|x|t - \Theta\left(\mathcal{F}_{[0,x]}^{B(|x|t)}\right)} \int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{i=0}^n 1_{\{\langle e_n, u_i\rangle \leq t\}} \sigma_{n-1}(du_i) dt dx \\
&= \frac{\gamma^{n+1}}{(n+1)} \int_0^\infty \int_{B(R)^c} |x| e^{-2\gamma|x|t - \Theta\left(\mathcal{F}_{[0,x]}^{B(|x|t)}\right)} dx \int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{i=0}^n 1_{\{\langle e_n, u_i\rangle \leq t\}} \sigma_{n-1}(du_i) dt.
\end{aligned}$$

By Lemma 6.4.3 and a change to polar coordinates,

$$\begin{aligned}
&\int_{B(R)^c} |x| e^{-2\gamma|x|t} e^{-\Theta\left(\mathcal{F}_{[0,x]}^{B(|x|t)}\right)} dx \\
&\qquad = \int_{B(R)^c} |x| e^{-2\gamma|x|t - \gamma|x|\left[\frac{2\kappa_{n-1}}{n\kappa_n}\left(1-t^2\right)^{\frac{n-1}{2}} - tI_{(1-t^2)}\left(\frac{n-1}{2}, \frac{1}{2}\right)\right]1_{\{t\leq 1\}}} dx \\
&\qquad = n\kappa_n \int_R^\infty r^n e^{-\gamma r\left[2t + \left(\frac{2\kappa_{n-1}}{n\kappa_n}\left(1-t^2\right)^{\frac{n-1}{2}} - tI_{1-t^2}\left(\frac{n-1}{2}, \frac{1}{2}\right)\right)1_{\{t\leq 1\}}\right]} dr.
\end{aligned}$$

Now, using the identity $I_{1-x}(a,b) = 1 - I_x(b,a)$,

$$\int_{B(R)^c} |x| e^{-2\gamma|x|t} e^{-\Theta\left(\mathcal{F}_{[0,x]}^{B(|x|t)}\right)} dx = n\kappa_n \int_R^\infty r^n e^{-\gamma r f_n(t)} dx,$$

where

$$f_n(t) := \begin{cases} t + \frac{2\kappa_{n-1}}{n\kappa_n}\left(1-t^2\right)^{\frac{n-1}{2}} + tI_{t^2}\left(\frac{1}{2}, \frac{n-1}{2}\right), & 0 \leq t \leq 1 \\ 2t, & t \geq 1. \end{cases}$$

Note that f_n is differentiable and that

$$f_n(0) = \frac{2\kappa_{n-1}}{n\kappa_n}. \tag{6.2}$$

Then, by the change of variables $y = \gamma f_n(t) r$,

$$\begin{aligned}
n\kappa_n \int_R^\infty r^n e^{-\gamma f_n(t) r} dr &= \frac{n\kappa_n}{(\gamma f_n(t))^{n+1}} \int_{\gamma f_n(t)R}^\infty y^n e^{-y} dy \\
&= \frac{n\kappa_n}{(\gamma f_n(t))^{n+1}} \Gamma_u(n+1, \gamma f_n(t)R).
\end{aligned}$$

This gives us that

$$\mathbb{P}(|Y| \geq R) =$$
$$\frac{n\kappa_n}{(n+1)} \int_0^\infty \frac{\Gamma_u(n+1, \gamma f_n(t) R)}{f_n(t)^{n+1}} \int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{i=0}^n 1_{\{\langle e_n, u_i \rangle \leq t\}} \sigma_{n-1}(du_i) dt.$$

Since the upper incomplete gamma function is decreasing in its second argument, for all $t \geq 0$,

$$\Gamma_u(n+1, \gamma f_n(t) R) \leq \Gamma_u\left(n+1, \gamma \frac{2\kappa_{n-1}}{n\kappa_n} R\right),$$

where we have used (6.2). So, we have the upper bound

$$\mathbb{P}(|Y| \geq R) \leq \frac{\Gamma_u(n+1, \gamma \frac{2\kappa_{n-1}}{n\kappa_n} R)}{\Gamma(n+1)}$$
$$\cdot \left[\frac{n\kappa_n}{(n+1)} \int_0^\infty \frac{\Gamma(n+1)}{f_n(t)^{n+1}} \left(\int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{j=0}^n 1_{\{\langle e_n, u_j \rangle \leq t\}} \sigma_{n-1}(du_j)\right) dt\right].$$

The term in the parentheses is the value of the integral $\int_{\mathbb{R}^n} \frac{\mathbb{P}(x \in Z)}{\mathbb{E}[V(Z)]} dx$ and is thus equal to 1. Hence,

$$\mathbb{P}(|Y| \geq R) \leq \frac{\Gamma_u\left(n+1, \gamma R \frac{2\kappa_{n-1}}{n\kappa_n}\right)}{\Gamma(n+1)}.$$

To obtain the upper bound for $\mathbb{P}(|Y| \leq R)$, we can follow a similar procedure up to the equality

$$\mathbb{P}(|Y| \leq R)$$
$$= \frac{n\kappa_n}{(n+1)} \int_0^\infty \frac{\Gamma_\ell(n+1, \gamma f_n(t) R)}{f_n(t)^{n+1}} \int_{(S^{n-1})^{n+1}} \triangle_n 1_P \prod_{i=0}^n 1_{\{\langle e_n, u_i \rangle \leq t\}} \sigma_{n-1}(du_i) dt.$$

The lower incomplete gamma function is not decreasing in t like the upper incomplete gamma function, so we cannot proceed exactly as above.

First we use the upper bound

$$\int_{(S^{n-1})^{n+1}} \triangle_n(u_0,\ldots,u_n)1_P \prod_{i=0}^{n} 1_{\{\langle e_n,u_i\rangle \le t\}}\sigma_{n-1}(du_i)$$
$$\le \int_{(S^{n-1})^{n+1}} \triangle_n(u_0,\ldots,u_n)1_P \prod_{i=0}^{n} \sigma_{n-1}(du_i).$$

Then, by the fact that

$$\frac{n2^n}{(n+1)(\omega_n)^2(\kappa_{n-1})^n}\triangle(u_0,\ldots,u_n)1_P(u_0,\ldots,u_n)\prod_{i=0}^{n} d\sigma_{n-1}(u_i)$$

is a joint density (see equation (11) in [16]),

$$\int_{(S^{n-1})^{n+1}} \prod_{j=0}^{n} \triangle_n(u_0,\ldots,u_n)1_P \prod_{i=0}^{n} \sigma_{n-1}(du_i) = \frac{\kappa_n(n+1)}{2^n}\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^n,$$

and thus

$$\mathbb{P}(|Y|\le R) \le \frac{n\kappa_n^2}{2^n}\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^n \int_0^\infty \frac{\Gamma_\ell(n+1,\gamma f_n(t)R)}{f_n(t)^{n+1}}dt.$$

Then, note that for $t \ge 1$, $f_n(t) = 2t$, so

$$\int_1^\infty \frac{\Gamma_\ell(n+1,\gamma f_n(t)R)}{f_n(t)^{n+1}}dt \le \frac{1}{2^{n+1}}\int_1^\infty \frac{\Gamma(n+1)}{t^{n+1}}dt = \frac{\Gamma(n)}{2^{n+1}}.$$

Thus,

$$\mathbb{P}(|Y|\le R) \le \frac{n\kappa_n^2}{2^n}\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^n \left[\int_0^1 \frac{\Gamma_\ell(n+1,\gamma f_n(t)R)}{f_n(t)^{n+1}}dt + \frac{\Gamma(n)}{2^{n+1}}\right].$$

Now, we note that the function

$$h_n(t) := \frac{\Gamma_\ell(n+1,\gamma f_n(t)R)}{f_n(t)^{n+1}}$$

is decreasing and thus reaches its maximum at $t = 0$. It suffices to show $h_n'(t) \leq 0$. Indeed, we first note that the derivative of $f_n'(t)$,

$$f_n'(t) = \begin{cases} 1 + I_{t^2}\left(\frac{1}{2}, \frac{n-1}{2}\right), & 0 \leq t \leq 1 \\ 2, & t \geq 1, \end{cases}$$

is positive. Then, by the Fundamental Theorem of Calculus,

$$\frac{d}{dt}\Gamma_\ell(n+1, \gamma f_n(t)R) = e^{-\gamma f_n(t)R}\left(\gamma f_n(t)R\right)^n (\gamma R)\, f_n'(t),$$

and by the quotient rule,

$$\begin{aligned} h_n'(t) &= -\frac{1}{f_n(t)^{n+2}}\left((n+1)f_n'(t)\Gamma_\ell(n+1, \gamma f_n(t)R) - e^{-\gamma f_n(t)R}\left(\gamma f_n(t)R\right)^{n+1} f_n'(t)\right) \\ &= -\frac{f_n'(t)}{f_n(t)^{n+2}}\left(\Gamma_\ell(n+1, \gamma R)(n+1) - e^{-\gamma f(t)R}\left(\gamma f_n(t)R\right)^{n+1}\right). \end{aligned}$$

Since f_n and f_n' are positive, it suffices to show the following inequality for h' to be negative:

$$e^{-\gamma f(t)R}\left(\gamma f_n(t)R\right)^{n+1} \leq \Gamma_\ell(n+1, \gamma f_n(t)R)(n+1). \tag{6.3}$$

Indeed, since $e^{-t} \geq e^{-x}$ for all $t \in [0, x]$,

$$\Gamma_\ell(n+1, x) = \int_0^x e^{-t}t^n dt \geq e^{-x}\int_0^x t^n dt = e^{-x}\frac{x^{n+1}}{n+1}.$$

Letting $x = 2\gamma f(t)R$ gives (6.3), and hence $h'(t) \leq 0$.

Thus, by (6.2),

$$\mathbb{P}(|Y| \leq R) \leq \frac{n\kappa_n^2}{4^n}\left(\frac{n\kappa_n}{2\kappa_{n-1}}\right)\left[\Gamma_\ell\left(n+1, 2\gamma R\frac{\kappa_{n-1}}{n\kappa_n}\right) + \Gamma(n)\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^{n+1}\right].$$

$\square$

We can now prove the main results.

6.4.1 Proof of Proposition 6.4.1

By Lemma 6.4.4,

$$\begin{aligned}
\mathbb{E}[|Y|^k] &= k\int_0^\infty y^{k-1}\mathbb{P}(|Y|\geq y)dy \\
&\leq k\int_0^\infty y^{k-1}\frac{\Gamma_u(n+1, 2\gamma\frac{\kappa_{n-1}}{n\kappa_n}y)}{\Gamma(n+1)}dy \\
&= \frac{k}{\Gamma(n+1)}\int_0^\infty y^{k-1}\left(\int_{2\gamma\frac{\kappa_{n-1}}{n\kappa_n}y}^\infty t^n e^{-t}dt\right)dy.
\end{aligned}$$

Then, by Fubini's Theorem,

$$\begin{aligned}
\mathbb{E}[|Y|^k] &\leq \frac{k}{\Gamma(n+1)}\int_0^\infty t^n e^{-t}\left(\int_0^{\frac{t}{2\gamma}\frac{n\kappa_n}{\kappa_{n-1}}} y^{k-1}dy\right)dt \\
&= \frac{1}{\Gamma(n+1)}\int_0^\infty t^n e^{-t}\left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}t\right)^k dt = \frac{\Gamma(n+k+1)}{\Gamma(n+1)}\left(\frac{n\kappa_n}{2\gamma\kappa_{n-1}}\right)^k.
\end{aligned}$$

Then, by (2.4),

$$\mathbb{E}[|Y|^k] \leq \frac{\Gamma(n+k+1)}{\Gamma(n+1)2^k}\left(\frac{\kappa_n}{\lambda}\right)^{k/n}.$$

6.4.2 Proof of Theorem 6.4.2

By the assumption on λ_n and (2.4), the intensity γ_n of X_n satisfies

$$\gamma_n \sim \frac{e^\lambda}{\sqrt{e}}n \text{ as } n\to\infty.$$

Then, by (2.7),

$$\lim_{n\to\infty}\gamma_n\sqrt{n}R\frac{2\kappa_{n-1}}{n^2\kappa_n} = \lim_{n\to\infty}\frac{e^\lambda}{\sqrt{e}}\sqrt{n}R\frac{2\kappa_{n-1}}{n\kappa_n} = \lim_{n\to\infty}\frac{2e^\lambda}{\sqrt{e}}\frac{\sqrt{n}R}{\sqrt{2\pi n}} = \frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}}.$$

Let $c_n = e^\lambda\frac{2\sqrt{n}R}{\kappa_n^{1/n}}$, and $c = \frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}}$. Then, by a modified application of Laplace's method (see B.0.2), for $\frac{e^\lambda R\sqrt{2}}{\sqrt{\pi e}} > 1$,

$$\Gamma_u(n+1, 2\gamma_n\sqrt{n}R\frac{\kappa_{n-1}}{n\kappa_n}) = \Gamma_u(n+1, c_n n) \sim n^n\frac{ce^{-n(c_n-\log c_n)}}{(c-1)}, \tag{6.4}$$

and for $\frac{e^{\lambda}R\sqrt{2}}{\sqrt{\pi e}} < 1$,

$$\Gamma_\ell(n+1, 2\gamma_n\sqrt{n}R\frac{\kappa_{n-1}}{n\kappa_n}) = \Gamma_\ell(n+1, c_n n) \sim n^n \frac{ce^{-n(c_n-\log c_n)}}{(1-c)}. \tag{6.5}$$

Then, by Lemma 6.4.4, (6.4), and Stirling's formula, as $n \to \infty$,

$$\mathbb{P}(|Y_n| \geq \sqrt{n}R) \leq \frac{\Gamma_u\left(n+1, c_n n\right)}{\Gamma(n+1)} \sim \frac{ce^{-n(c_n-\log c_n-1)}}{\sqrt{2\pi n}(c-1)}.$$

Then, for $R > e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$,

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(|Y_n| \geq \sqrt{n}R) &\leq \limsup_{n\to\infty} 1 - c_n + \ln c_n = 1 - c + \ln c \\ &= -\frac{e^{\lambda}R\sqrt{2}}{\sqrt{\pi e}} + \ln \frac{e^{\lambda}R\sqrt{2e}}{\sqrt{\pi}} = \lambda + \frac{1}{2}\ln\left(\frac{2e}{\pi}\right) + \ln R - \frac{e^{\lambda}R\sqrt{2}}{\sqrt{\pi e}}.\end{aligned}$$

Similarly, for $R < e^{-\lambda}\frac{\sqrt{\pi e}}{\sqrt{2}}$,

$$\mathbb{P}(|Y_n| \leq \sqrt{n}R) \leq \frac{n\kappa_n^2}{4^n}\left(\frac{n\kappa_n}{2\kappa_{n-1}}\right)\Gamma_\ell\left(n+1, c_n n\right)\left(1 + \frac{\Gamma(n)\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^{n+1}}{\Gamma_\ell\left(n+1, c_n n\right)}\right).$$

Then, for all λ and R,

$$\lim_{n\to\infty} \frac{\Gamma(n)\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^{n+1}}{\Gamma_\ell\left(n+1, c_n n\right)} = \lim_{n\to\infty} \frac{1}{n}\left(\frac{\kappa_{n-1}}{n\kappa_n}\right)^{n+1} \frac{\Gamma(n+1)}{\Gamma_\ell\left(n+1, c_n n\right)} = 0.$$

Hence, by (6.5) and Stirling's formula, as $n \to \infty$,

$$\mathbb{P}(|Y_n| \leq \sqrt{n}R) \lesssim \frac{n\kappa_n^2}{4^n}\left(\frac{n\kappa_n}{2\kappa_{n-1}}\right)\Gamma_\ell\left(n+1, c_n n\right) \sim \left(\frac{\pi}{2}\right)^n \frac{ce^{-n(c_n-\ln c_n-1)}}{\sqrt{2\pi n}(1-c)},$$

and thus

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(|Y_n| \leq \sqrt{n}R) \leq \lambda + \frac{1}{2}\ln\left(\frac{2e}{\pi}\right) + \ln R - \frac{e^{\lambda}R\sqrt{2}}{\sqrt{\pi e}} - \ln 2 + \ln \pi.$$

Appendices

Appendix A

Proofs from Chapter 4

A.1 Proof of (4.3)

For each n, let $X_n \sim DPP(K_n)$ in $\mathbb{R}^n$ be stationary with intensity $K_n(0) = e^{n\rho}$. From (4.1), there exists $\tilde{R} := \frac{1}{\sqrt{2\pi e}e^{\rho}}$ such that

$$\lim_{n\to\infty} \mathbb{E}[X_n(B_n(\sqrt{n}R))] = \begin{cases} 0, & R < \tilde{R} \\ \infty, & R > \tilde{R}. \end{cases}$$

Now,

$$\mathbb{E}[X_n(B_n(\sqrt{n}R))] - \mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))] = \frac{1}{e^{n\rho}} \int_{B_n(\sqrt{n}R)} K_n(x)^2 \mathrm{d}x,$$

and, by Parseval's theorem and the condition on K_n for the existence of Y_n,

$$\frac{1}{e^{n\rho}} \int_{B_n(\sqrt{n}R)} K_n(x)^2 \mathrm{d}x \leq \frac{1}{e^{n\rho}} \int_{\mathbb{R}^n} \hat{K}_n(\xi)^2 \mathrm{d}\xi \leq \frac{1}{e^{n\rho}} \int_{\mathbb{R}^n} \hat{K}_n(\xi) \mathrm{d}\xi = 1.$$

Also, since $\frac{1}{e^{n\rho}} \int_{B_n(\sqrt{n}R)} K_n(x)^2 \mathrm{d}x \geq 0$, the following bounds hold:

$$\mathbb{E}[X_n(B_n(\sqrt{n}R))] - 1 \leq \mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))] \leq \mathbb{E}[X_n(B_n(\sqrt{n}R))].$$

Thus, the threshold remains the same for the reduced Palm expectation:

$$\lim_{n\to\infty} \mathbb{E}[X_n^{0,!}(B_n(\sqrt{n}R))] = \begin{cases} 0, & R < \tilde{R} \\ \infty, & R > \tilde{R}. \end{cases}$$

By the first moment inequality and Proposition 5.1 in [10], one has the following bounds:

$$1 - \mathbb{E}[X_n^{0,!}(B(\sqrt{n}R))] \leq \mathbb{P}(X_n^{0,!}(B_n(\sqrt{n}R)) = 0) \leq \exp\left(-\mathbb{E}[X_n^{0,!}(B(\sqrt{n}R))]\right).$$

Thus, $\lim_{n\to\infty} \mathbb{P}(X_n^{0,!}(B_n(\sqrt{n}R)) > 0) = \begin{cases} 0, & R < \tilde{R} \\ 1 & R > \tilde{R}. \end{cases}$

A.2 Proof of Lemma 4.3.1

The assumption $\frac{|Y_n|}{\sqrt{n}} \to R^*$ in probability means that for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{|Y_n|}{\sqrt{n}} - R^*\right| > \varepsilon\right) \to 0, \text{ as } n \to \infty.$$

Now, assume $R < R^*$. Then, there exists $\varepsilon > 0$ such that $R = R^* - \mathrm{Var}epsilon.$ Thus,

$$\mathbb{P}(|Y_n| \leq \sqrt{n}R) = \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} \leq R^* - \varepsilon\right) \leq \mathbb{P}\left(\left|\frac{|Y_n|}{\sqrt{n}} - R^*\right| > \varepsilon\right) \to 0 \text{ as } n \to \infty.$$

Second, assume $R > R^*$. Then, there exists $\varepsilon > 0$ such that $R = R^* + \mathrm{Var}epsilon$, and

$$\mathbb{P}(|Y_n| \leq \sqrt{n}R) = 1 - \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} > R^* + \varepsilon\right) \geq 1 - \mathbb{P}\left(\left|\frac{|Y_n|}{\sqrt{n}} - R^*\right| > \varepsilon\right) \to 1.$$

Then, by the assumption on Y_n and Theorem 3.3.2,

$$\mathbb{P}(\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset) = \mathbb{P}\left(|Y_n| \leq \sqrt{n}R\right) \to \begin{cases} 0, & R < R^* \\ 1, & R > R^*. \end{cases}$$

A.3 Proof of Proposition 4.3.2

Since for all n, Y_n is isotropic, Y_n as defined in Lemma 4.3.1 has a radially symmetric density. Thus, Y_n has the same distribution as the product R_nU_n, where R_n is equal in distribution to $|Y_n|$, U_n is uniformly distributed on $\mathbb{S}^{n-1}$, and R_n and U_n are independent. Letting $\sigma_n^2 = \mathbb{E}|Y_n|^2$ for each n, $\frac{\sqrt{n}}{\sigma_n}Y_n$ then satisfies the conditions of Theorem 2.4.1 for each n. Then, by Theorem 2.4.1, for any $\delta > 0$, there exist absolute constants $C, c > 0$ such that

$$\mathbb{P}\left(\left|\frac{|Y_n|}{\sigma_n} - 1\right| \geq \delta\right) \leq Ce^{-c\sqrt{n}\min(\delta,\delta^3)}.$$

Now, let $\delta \in (0,1)$. Then,

$$\mathbb{P}(\eta_n(B_n(\sigma_n(1-\delta))) > 0|\eta_n \neq \emptyset) = \mathbb{P}\left(\frac{|Y_n|}{\sigma_n} \leq 1-\delta\right) \leq Ce^{-c\sqrt{n}\delta^3},$$

since $\min(\delta^3, \delta) = \delta^3$ for $\delta \in (0,1)$. Similarly, for any $\delta > 0$,

$$\mathbb{P}(\eta_n(\mathbb{R}^n \setminus B_n(\sigma_n(1+\delta))) > 0|\eta_n \neq \emptyset) = \mathbb{P}\left(\frac{|Y_n|}{\sigma_n} \geq 1+\delta\right) \leq Ce^{-c\sqrt{n}\min(\delta^3,\delta)}.$$

Now, assume $\frac{\sigma_n}{\sqrt{n}} \to R^* \in (0,\infty)$ as $n \to \infty$. For $R < R^*$, there exists $\varepsilon \in (0,1)$ such that $R = R^*(1-\varepsilon)$. Then, for all n large enough, $\frac{\sqrt{n}R^*}{\sigma_n} < \frac{1-\varepsilon/2}{1-\varepsilon}$ and

$$\begin{aligned}
\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0|\eta_n \neq \emptyset] &= \mathbb{P}\left(|Y_n| \leq \sqrt{n}R\right) = \mathbb{P}\left(\frac{|Y_n|}{\sigma_n} \leq \frac{\sqrt{n}R}{\sigma_n}\right) \\
&= \mathbb{P}\left(\frac{|Y_n|}{\sigma_n} \leq \frac{\sqrt{n}R^*(1-\varepsilon)}{\sigma_n}\right) \leq \mathbb{P}\left(\frac{|Y_n|}{\sigma_n} \leq 1 - \frac{\varepsilon}{2}\right) \\
&\leq \mathbb{P}\left(\left|\frac{|Y_n|}{\sigma_n} - 1\right| \geq \frac{\varepsilon}{2}\right) \leq Ce^{-c\sqrt{n}(\varepsilon/2)^3}.
\end{aligned}$$

Thus for all $R < R^*$, there exists a constant $C(\varepsilon(R)) = c\varepsilon^3/2^3$ such that

$$\liminf_{n\to\infty} -\frac{1}{\sqrt{n}} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0|\eta_n \neq \emptyset] \geq C(\varepsilon(R)).$$

A similar argument gives that for all $R > R^*$, there exists $C(\varepsilon(R))$ such that

$$\liminf_{n\to\infty} -\frac{1}{\sqrt{n}} \ln \mathbb{P}[\eta_n(\mathbb{R}^n \setminus B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] \geq C(\varepsilon(R)).$$

This implies the threshold (4.5).

A.4 Proof of Proposition 4.3.3

If $\frac{|Y_n|}{\sqrt{n}}$ satisfies a large deviations principle with convex rate function I, then by definition,

$$\begin{aligned} -\inf_{r<R} I(r) &\leq \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} \leq R\right) \\ &\leq \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} \leq R\right) \leq -\inf_{r\leq R} I(r). \end{aligned}$$

Thus,

$$\begin{aligned} -\inf_{r<R} I(r) &\leq \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] \\ &\leq \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] \leq -\inf_{r\leq R} I(r). \end{aligned}$$

By the assumption that the rate function I is strictly convex, there exists a unique R^* such that $I(R^*) = 0$. Note that $\inf_{\{r\leq R\}} I(r)$ is then zero for $R > R^*$. Thus,

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = \begin{cases} 0, & R < R^* \\ 1, & R > R^*. \end{cases}$$

Let $R < R^*$. If the rate function I is continuous at R, then the above inequalities become equalities and

$$\lim_{n\to\infty} -\frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0 | \eta_n \neq \emptyset] = I(R).$$

A.5 Proof of Lemma 4.4.1

The proof shows that the sequence of random variables satisfies the conditions of the Gärtner-Ellis theorem (see [27]). First,

$$\mathbb{E}[e^{s|Y_n|^2}] = \frac{e^{2n\rho}}{\binom{m-1+n/2}{m-1}^2 \|K_n\|_2^2} \overbrace{\int_{\mathbb{R}^n} e^{-\left(\frac{2}{\alpha^2 m}-s\right)|x|^2} \left(L_{m-1}^{n/2}\left(\frac{1}{m}\left|\frac{x}{\alpha}\right|^2\right)\right)^2 \mathrm{d}x}^{\mathrm{I(s)}} .$$

Writing out the polynomial, the integral I above becomes

$$I(s) = \sum_{k,j=0}^{m-1} \binom{m-1+n/2}{m-1-k}\binom{m-1+n/2}{m-1-j} \frac{(-1)^{k+j}}{k!j!(m\alpha^2)^{k+j}} \int_{\mathbb{R}^n} e^{-\left(\frac{2}{\alpha^2 m}-s\right)|x|^2} |x|^{2k+2j} \mathrm{d}x.$$

A quick calculation shows that for $a > 0$,

$$\int_{\mathbb{R}^n} e^{-a|x|^2}|x|^b \mathrm{d}x = \frac{\pi^{n/2}}{a^{\frac{n+b}{2}}} \frac{\Gamma\left(\frac{n}{2}+\frac{b}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}. \tag{A.1}$$

Then, if $s < \frac{2}{\alpha^2 m}$,

$$I(s) = \frac{\pi^{n/2}}{\left(\frac{2}{\alpha^2 m}-s\right)^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \sum_{k,j=0}^{m-1} \binom{m-1+n/2}{m-1-k}\binom{m-1+n/2}{m-1-j} \frac{(-1)^{k+j}\Gamma\left(\frac{n}{2}+k+j\right)}{k!j!\left(2-sm\alpha^2\right)^{k+j}},$$

and $I(s) = \infty$ otherwise. For each $k, j \in \mathbb{N}$,

$$\begin{aligned} &\binom{m-1+n/2}{m-1-k}\binom{m-1+n/2}{m-1-j}\Gamma\left(\frac{n}{2}+k+j\right) \\ &\qquad \sim \frac{1}{(m-1-k)!(m-1-j)!}\left(\frac{n}{2}\right)^{2m-2}\Gamma\left(\frac{n}{2}\right), \end{aligned} \tag{A.2}$$

as $n \to \infty$. So, $I(s)$ has the following asymptotic expansion for $s < \frac{2}{\alpha^2 m}$ as $n \to \infty$:

$$I(s) \sim \frac{\pi^{n/2}}{\left(\frac{2}{\alpha^2 m}-s\right)^{\frac{n}{2}}} \left(\frac{n}{2}\right)^{2m-2} \sum_{k=0}^{m-1}\sum_{j=0}^{m-1} \frac{(-1)^{k+j}}{k!j!(m-1-k)!(m-1-j)!} \frac{1}{\left(2-sm\alpha^2\right)^{k+j}}.$$

By (4.9) and (A.2),

$$\frac{1}{e^{2n\rho}}\|K_n\|_2^2 \sim \frac{\alpha^n}{\binom{m-1+n/2}{m-1}^2}\left(\frac{m\pi}{2}\right)^{\frac{n}{2}}\left(\frac{n}{2}\right)^{2m-2}\sum_{k,j=0}^{m-1}\frac{(-1)^{k+j}}{k!j!(m-1-k)!(m-1-j)!}\frac{1}{2^{k+j}}, \tag{A.3}$$

and hence,

$$\mathbb{E}[e^{s|Y_n|^2}] \sim \left(1-\frac{s\alpha^2 m}{2}\right)^{-\frac{n}{2}}\left(\frac{\sum_{k,j=0}^{m-1}\frac{(-1)^{k+j}}{k!j!(m-1-k)!(m-1-j)!}\frac{1}{(2-sm\alpha^2)^{k+j}}}{\sum_{k,j=0}^{m-1}\frac{(-1)^{k+j}}{k!j!(m-1-k)!(m-1-j)!}\frac{1}{2^{k+j}}}\right),$$

as $n\to\infty$. Thus,

$$\Lambda(s) = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[e^{s|Y_n|^2}] = -\frac{1}{2}\log\left(1-\frac{s\alpha^2 m}{2}\right) \text{ if } s<\frac{2}{\alpha^2 m},$$

and is infinite otherwise. It is clear that $0\in(D(\Lambda))^\circ$, where $D(\Lambda)=\{s\in\mathbb{R}:\Lambda(s)<\infty\}$. Thus, the Gärtner-Ellis conditions are satisfied. The rate function for the LDP is computed with the optimization

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}[x\lambda-\Lambda(\lambda)] = \sup_{\lambda\in\mathbb{R}}\left[x\lambda+\frac{1}{2}\log\left(1-\frac{\lambda\alpha^2 m}{2}\right)\right].$$

Then, since

$$0=\frac{d}{d\lambda}\left[x\lambda+\frac{1}{2}\log\left(1-\frac{\lambda\alpha^2 m}{2}\right)\right] = x-\frac{\alpha^2 m}{4-2\alpha^2 m\lambda} \text{ if and only if } \lambda=\frac{2}{\alpha^2 m}-\frac{1}{2x},$$

the rate function is

$$\Lambda^*(x) = x\left(\frac{2}{\alpha^2 m}-\frac{1}{2x}\right)+\frac{1}{2}\log\left(1-\frac{\left(\frac{2}{\alpha^2 m}-\frac{1}{2x}\right)\alpha^2 m}{2}\right) = \frac{2x}{\alpha^2 m}-\frac{1}{2}+\frac{1}{2}\log\left(\frac{\alpha^2 m}{4x}\right).$$

Then by the contraction principle (see [27]), the sequence $\frac{|Y_n|}{\sqrt{n}}$ satisfies an LDP with rate function

$$\Lambda^*(x) = \frac{2x^2}{\alpha^2 m}-\frac{1}{2}+\frac{1}{2}\log\left(\frac{\alpha^2 m}{4x^2}\right).$$

Note that $\Lambda^*(x)=0$ if and only if $x=\sqrt{m}\frac{\alpha}{2}$, implying $\frac{|Y_n|}{\sqrt{n}}\to\sqrt{m}\frac{\alpha}{2}$ in probability.

A.6 Proof of Proposition 4.4.2

Proof. For each n, let Y_n be a random vector in $\mathbb{R}^n$ with density $\frac{K_n^2}{\|K_n\|_2^2}$. By Lemma 4.4.1, for $R < \sqrt{m}\frac{\alpha}{2}$,

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} \le R\right) = \frac{2R^2}{\alpha^2 m} - \frac{1}{2} + \frac{1}{2}\log\left(\frac{\alpha^2 m}{4R^2}\right).$$

Then by (A.3), as $n \to \infty$,

$$\mathbb{P}[\eta_n \neq \emptyset] = \frac{1}{e^{n\rho}}\|K_n\|_2^2 \sim \left(\frac{e^{2\rho}\alpha^2 m\pi}{2}\right)^{\frac{n}{2}} \sum_{k,j=0}^{m-1} \frac{(-1)^{k+j}}{k!j!(m-1-k)!(m-1-j)!}\frac{1}{2^{k+j}}, \tag{A.4}$$

Thus,

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0] \\
&= \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}[\eta_n \neq \emptyset] + \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}\left(\frac{|Y_n|}{\sqrt{n}} \le R\right) \\
&= \begin{cases} -\rho - \log\alpha - \frac{1}{2}\log\left(\frac{m\pi}{2}\right) + \left(\frac{2R^2}{\alpha^2 m} - \frac{1}{2} + \frac{1}{2}\log\left(\frac{\alpha^2 m}{4R^2}\right)\right), & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ -\rho - \log\alpha - \frac{1}{2}\log\left(\frac{m\pi}{2}\right), & R > \sqrt{m}\frac{\alpha}{2} \end{cases} \\
&= \begin{cases} -\rho - \frac{1}{2}\log 2\pi e + \frac{2R^2}{\alpha^2 m} - \log R, & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ -\rho - \log\alpha - \frac{1}{2}\log\frac{m\pi}{2}, & R > \sqrt{m}\frac{\alpha}{2}. \end{cases}
\end{aligned}$$

□

A.7 Proof of Lemma 4.4.3

Since for all n, $\hat{K}_n \in C^2(\mathbb{R}^n)$, Parseval's theorem implies

$$\mathbb{E}[|Y_n|^2] = \frac{1}{\|K_n\|_2^2}\int_{\mathbb{R}^n} |x|^2 K_n(x)^2 \mathrm{d}x = \frac{1}{\|\hat{K}_n\|_2^2}\int_{\mathbb{R}^n} -\frac{\triangle \hat{K}_n(\xi)}{(2\pi)^2}\hat{K}_n(\xi)\mathrm{d}\xi. \tag{A.5}$$

To compute the Laplacian of $\hat{K}$, we first see that for each i,

$$
\begin{aligned}
\frac{\partial^2}{\partial x_i^2} e^{-|\alpha x|^\nu} &= \frac{\partial}{\partial x_i}(-\nu\alpha^\nu x_i |x|^{\nu-2} e^{-|\alpha x|^\nu}) \\
&= -\nu\alpha^\nu |x|^{\nu-2} e^{-|\alpha x|^\nu} - \nu\alpha^\nu x_i \left(\tfrac{\partial}{\partial x_i}|x|^{\nu-2}\right) e^{-|\alpha x|^\nu} + (\nu\alpha^\nu x_i |x|^{\nu-2})^2 e^{-|\alpha x|^\nu} \\
&= e^{-|\alpha x|^\nu}\left(-\nu\alpha^\nu|x|^{\nu-2} - \nu(\nu-2)\alpha^\nu x_i^2|x|^{\nu-4} + \nu^2\alpha^{2\nu}x_i^2|x|^{2\nu-4}\right) \\
&= e^{-|\alpha x|^\nu}\left(x_i^2(\nu^2\alpha^{2\nu}|x|^{2\nu-4} - \nu(\nu-2)\alpha^\nu|x|^{\nu-4}) - \nu\alpha^\nu|x|^{\nu-2}\right).
\end{aligned}
$$

Then,

$$
\begin{aligned}
\triangle e^{-|\alpha x|^\nu} &= \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} e^{-|\alpha x|^\nu} \\
&= \sum_{i=1}^n e^{-|\alpha x|^\nu}\left(x_i^2(\nu^2\alpha^{2\nu}|x|^{2\nu-4} - \nu(\nu-2)\alpha^\nu|x|^{\nu-4}) - \nu\alpha^\nu|x|^{\nu-2}\right) \\
&= e^{-|\alpha x|^\nu}\left(|x|^2(\nu^2\alpha^{2\nu}|x|^{2\nu-4} - \nu(\nu-2)\alpha^\nu|x|^{\nu-4}) - n\nu\alpha^\nu|x|^{\nu-2}\right) \\
&= e^{-|\alpha x|^\nu}\left(\nu^2\alpha^{2\nu}|x|^{2\nu-2} - (\nu(\nu-2)\alpha^\nu + n\nu\alpha^\nu)|x|^{\nu-2}\right).
\end{aligned}
$$

Thus by (A.5) and (4.12),

$$
\begin{aligned}
&\mathbb{E}[|Y_n|^2] \\
&= \frac{\Gamma(\frac{n}{2}+1)\alpha_n^n 2^{\frac{n}{\nu}}}{4\pi^2\pi^{n/2}\Gamma(\frac{n}{\nu}+1)} \int_{\mathbb{R}^n} e^{-2|\alpha_n x|^\nu}\left((\nu(\nu-2)\alpha_n^\nu + n\nu\alpha_n^\nu)|x|^{\nu-2} - \nu^2\alpha_n^{2\nu}|x|^{2\nu-2}\right)\mathrm{d}x \\
&= \frac{\Gamma(\frac{n}{2}+1)\alpha_n^{n+\nu}2^{\frac{n}{\nu}}\nu}{4\pi^2\pi^{\frac{n}{2}}\Gamma(\frac{n}{\nu}+1)}\left[(\nu-2+n)\int_{\mathbb{R}^n}|x|^{\nu-2}e^{-2|\alpha_n x|^\nu}\mathrm{d}x - \nu\alpha_n^\nu\int_{\mathbb{R}^n}e^{-2|\alpha_n x|^\nu}|x|^{2\nu-2}dx\right].
\end{aligned}
$$

Then, using (A.1),

$$
\begin{aligned}
\mathbb{E}[|Y_n|^2] &= n\frac{\alpha_n^{n+\nu}2^{\frac{n}{\nu}}\nu}{4\pi^2\Gamma(\frac{n}{\nu}+1)}\left[-\frac{\nu\alpha_n^\nu\Gamma(\frac{n+2\nu-2}{\nu})}{\nu 2^{(n+2\nu-2)/\nu}\alpha_n^{n+2\nu-2}} + \frac{(\nu-2+n)\Gamma(\frac{n+\nu-2}{\nu})}{\nu 2^{(n+\nu-2)/\nu}\alpha_n^{n+\nu-2}}\right] \\
&= n\frac{2^{2/\nu}\alpha_n^2}{4\pi^2\Gamma(\frac{n}{\nu}+1)}\left[\frac{(\nu-2+n)}{2}\Gamma\left(\frac{n-2}{\nu}+1\right) - \frac{\nu}{4}\Gamma\left(\frac{n-2}{\nu}+2\right)\right] \\
&= n\frac{2^{2/\nu}\alpha_n^2\Gamma\left(\frac{n-2}{\nu}+1\right)}{4\pi^2\Gamma(\frac{n}{\nu}+1)}\left[\frac{n}{4}+\frac{\nu}{4}-\frac{1}{2}\right].
\end{aligned}
$$

By the asymptotic formula for the Gamma function, as $n \to \infty$,

$$\begin{aligned}
&\mathbb{E}[|Y_n|^2] \\
&\sim n\frac{\alpha_n^2 2^{\frac{2}{\nu}}}{4\pi^2}\left(\sqrt{\frac{\nu}{2\pi n}}\left(\frac{\nu e}{n}\right)^{\frac{n}{\nu}}\right)\left(\sqrt{\frac{2\pi(n-2)}{\nu}}\left(\frac{n-2}{\nu e}\right)^{\frac{(n-2)}{\nu}}\right)\left[\frac{n}{4}+\frac{\nu}{4}-\frac{1}{2}\right] \\
&= n\frac{\alpha_n^2 2^{2/\nu}}{4\pi^2}\frac{\sqrt{n-2}}{\sqrt{n}}\left(1-\frac{2}{n}\right)^{\frac{n}{\nu}}\left(\frac{n-2}{\nu e}\right)^{-\frac{2}{\nu}}\left[\frac{n}{4}+\frac{\nu}{4}-\frac{1}{2}\right] \sim n^{2-2/\nu}\alpha_n^2\frac{(2\nu)^{2/\nu}}{16\pi^2}.
\end{aligned}$$

By assumption, $\alpha_n \sim \alpha n^{\frac{1}{\nu}-\frac{1}{2}}$ for some constant $\alpha \in (0,\infty)$. Thus,

$$\lim_{n\to\infty}\frac{\mathbb{E}[|Y_n|^2]}{n} = \alpha^2\frac{(2\nu)^{2/\nu}}{16\pi^2}.$$

For the second moment of $|Y_n|^2$, Parseval's theorem is applied again and gives that

$$\mathbb{E}[(|Y_n|^2)^2] = \frac{1}{\|K_n\|_2^2}\int_{\mathbb{R}^n}(|x|^2K_n(x))^2\mathrm{d}x = \frac{1}{\|K_n\|_2^2}\int_{\mathbb{R}^n}\frac{(\triangle\hat{K}_n(\xi))^2}{(2\pi)^4}\mathrm{d}\xi. \quad \text{(A.6)}$$

Then, by the above computation of the Laplacian of $\hat{K}$, (A.1), and (4.12),

$$
\begin{aligned}
&\mathbb{E}[(|Y_n|^2)^2] \\
&= \frac{\Gamma(\frac{n}{2}+1)\alpha_n^n 2^{n/\nu}\nu^2\alpha_n^{2\nu}}{(2\pi)^4\pi^{n/2}\Gamma(\frac{n}{\nu}+1)} \int_{\mathbb{R}^n} e^{(-2|\alpha_n x|^\nu)} \left(\nu\alpha_n^\nu |x|^{2\nu-2} - (\nu-2+n)|x|^{\nu-2}\right)^2 \mathrm{d}x \\
&= \frac{\Gamma(\frac{n}{2}+1)\alpha_n^n 2^{n/\nu}\nu^2\alpha_n^{2\nu}}{(2\pi)^4\pi^{n/2}\Gamma(\frac{n}{\nu}+1)} \Bigg[(\nu\alpha_n^\nu)^2 \int_{\mathbb{R}^n} e^{-2|\alpha_n x|^\nu}|x|^{4\nu-4}\mathrm{d}x \\
&\qquad - 2\nu\alpha_n^\nu(\nu-2+n)\int_{\mathbb{R}^n} e^{-2|\alpha_n x|^\nu}|x|^{3\nu-4}dx + (\nu-2+n)^2\int_{\mathbb{R}^n} e^{-2|\alpha_n x|^\nu}|x|^{2\nu-4}\mathrm{d}x\Bigg] \\
&= n\frac{\alpha_n^n 2^{n/\nu}\nu^2\alpha_n^{2\nu}}{(2\pi)^4\Gamma(\frac{n}{\nu}+1)}\Bigg[\frac{(\nu\alpha_n^\nu)^2\Gamma(\frac{n+4\nu-4}{\nu})}{\nu 2^{(n+4\nu-4)/\nu}\alpha_n^{n+4\nu-4}} \\
&\qquad - \frac{2\nu\alpha_n^\nu(\nu-2+n)\Gamma(\frac{n+3\nu-4}{\nu})}{\nu 2^{(n+3\nu-4)/\nu}\alpha_n^{n+3\nu-4}} + \frac{(\nu-2+n)^2\Gamma(\frac{n+2\nu-4}{\nu})}{\nu 2^{(n+2\nu-4)/\nu}\alpha_n^{n+2\nu-4}}\Bigg] \\
&= \frac{n2^{4/\nu}\nu^2\alpha_n^4}{(2\pi)^4\Gamma(\frac{n}{\nu}+1)}\left[\frac{\nu\Gamma\left(\frac{n-4}{\nu}+4\right)}{2^4} - \frac{2(\nu-2+n)\Gamma\left(\frac{n-4}{\nu}+3\right)}{2^3} + \frac{(\nu-2+n)^2\Gamma\left(\frac{n-4}{\nu}+2\right)}{\nu 2^2}\right] \\
&= n\frac{2^{4/\nu}\alpha_n^4\Gamma\left(\frac{n-4}{\nu}+1\right)}{(2\pi)^4\Gamma\left(\frac{n}{\nu}+1\right)}\Bigg[\frac{\nu^3}{2^4}\left(\frac{n-4}{\nu}+3\right)\left(\frac{n-4}{\nu}+2\right)\left(\frac{n-4}{\nu}+1\right) \\
&\qquad - \frac{\nu^2(n+\nu-2)}{2^2}\left(\frac{n-4}{\nu}+2\right)\left(\frac{n-4}{\nu}+1\right) + \frac{\nu(n+\nu-2)^2}{2^2}\left(\frac{n-4}{\nu}+1\right)\Bigg] \\
&= n\frac{2^{4/\nu}\alpha_n^4}{(2\pi)^4}\frac{\Gamma\left(\frac{n-4}{\nu}+1\right)}{\Gamma\left(\frac{n}{\nu}+1\right)}\left(\frac{n^3}{2^4} - \frac{n^3}{2^2} + \frac{n^3}{2^2} + o(n^3)\right) \\
&= n^4\frac{2^{4/\nu}\alpha_n^4}{(2\pi)^4}\frac{\Gamma\left(\frac{n-4}{\nu}+1\right)}{\Gamma\left(\frac{n}{\nu}+1\right)}\left(\frac{1}{16}+o(1)\right) \\
&\qquad \sim n^4\frac{2^{4/\nu}\alpha_n^4}{16(2\pi)^4}\sqrt{\frac{\nu}{2\pi n}}\left(\frac{\nu e}{n}\right)^{\frac{n}{\nu}}\sqrt{\frac{2\pi(n-4)}{\nu}}\left(\frac{n-4}{\nu e}\right)^{\frac{n-4}{\nu}} \\
&= n^4\sqrt{\frac{n-4}{n}}\left(1-\frac{4}{n}\right)^{\frac{n}{\nu}}\left(\frac{n-4}{\nu e}\right)^{-\frac{4}{\nu}}\frac{\alpha_n^4 2^{4/\nu}}{16(2\pi)^4} \sim n^4(n-4)^{-\frac{4}{\nu}}\frac{\alpha_n^4(2\nu)^{4/\nu}}{16(2\pi)^4}.
\end{aligned}
$$

Again, since $\alpha_n \sim \alpha n^{\frac{1}{\nu}-\frac{1}{2}}$, $\mathbb{E}[(|Y_n|^2)^2] = O(n^2)$, and

$$\lim_{n\to\infty}\frac{\mathbb{E}[(|Y_n|^2)^2]}{n^2} = \alpha^4\frac{(2\nu)^{4/\nu}}{16(2\pi)^4}.$$

Note that this limit is exactly the square of the limit of the expectation of $|Y_n|^2/n$, implying

$$\mathrm{Var}\left(\frac{|Y_n|^2}{n^2}\right) = \frac{\mathbb{E}[(|Y_n|^2)^2]}{n^2} - \left(\frac{\mathbb{E}[|Y_n|^2]}{n}\right)^2 \to 0 \text{ as } n \to \infty.$$

Thus, by Chebychev's inequality, $\frac{|Y_n|}{\sqrt{n}} \to \alpha\frac{(2\nu)^{1/\nu}}{4\pi}$ in probability.

A.8 Proof of Proposition 4.4.5

First, for $k \geq 0$, we see that

$$\begin{aligned}
&\int_{\mathbb{R}^n} |x|^k K(x)^2 \mathrm{d}x \\
&= \int_{\mathbb{R}^n} |x|^k \left(e^{n\rho}2^{(\sigma+n)/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)\frac{J_{(\sigma+n)/2}(2|x/\alpha|\sqrt{(\sigma+n)/2})}{(2|x/\alpha|\sqrt{(\sigma+n)/2})^{(\sigma+n)/2}}\right)^2 \mathrm{d}x \\
&= e^{2n\rho}2^{(\sigma+n)}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2 \int_{\mathbb{R}^n} |x|^k \frac{J_{(\sigma+n)/2}(2|x/\alpha|\sqrt{(\sigma+n)/2})^2}{(2|x/\alpha|\sqrt{(\sigma+n)/2})^{(\sigma+n)}}\mathrm{d}x \\
&= e^{2n\rho}2^{(\sigma+n)}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2 \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\int_0^\infty r^{n-1}r^k \frac{J_{(\sigma+n)/2}(2(r/\alpha)\sqrt{(\sigma+n)/2})^2}{(2(r/\alpha)\sqrt{(\sigma+n)/2})^{(\sigma+n)}}\mathrm{d}x,
\end{aligned}$$

and by the change of variables $y = \left(\frac{2}{\alpha}\sqrt{\frac{\sigma+n}{2}}\right) r$,

$$\begin{aligned}
&= e^{2n\rho}2^{\sigma+n}\frac{2\pi^{n/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2}{\Gamma(\frac{n}{2})}\int_0^\infty \left(\frac{2}{\alpha}\sqrt{\frac{\sigma+n}{2}}\right)^{-k-n+1}\frac{J_{(\sigma+n)/2}(y)^2}{y^{\sigma+1-k}}\left(\frac{2}{\alpha}\sqrt{\frac{\sigma+n}{2}}\right)^{-1}\mathrm{d}y \\
&= e^{2n\rho}2^{\sigma+n}\frac{2\pi^{n/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2\alpha^{k+n}}{\Gamma(\frac{n}{2})(2(\sigma+n))^{\frac{k+n}{2}}}\int_0^\infty \frac{J_{(\sigma+n)/2}(y)^2}{y^{\sigma+1-k}}\mathrm{d}y.
\end{aligned}$$

For $\sigma + 1 - k > 0$, from [1, 10.22.57],

$$\int_0^\infty \frac{J_{(\sigma+n)/2}(y)^2}{y^{\sigma+1-k}}\mathrm{d}y = \frac{\Gamma\left(\frac{n}{2}+\frac{k}{2}\right)\Gamma(\sigma+1-k)}{2^{\sigma-k+1}\Gamma\left(\frac{\sigma-k}{2}+1\right)^2\Gamma\left(\sigma-\frac{k}{2}+\frac{n}{2}+1\right)},$$

and thus,

$$\begin{aligned}
&\int_{\mathbb{R}^n} |x|^k K(x)^2 \mathrm{d}x \\
&= e^{2n\rho} 2^{\sigma+n} \frac{2\pi^{n/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2 \alpha^{k+n}}{\Gamma(\frac{n}{2})(2(\sigma+n))^{\frac{k+n}{2}}} \frac{\Gamma\left(\frac{n}{2}+\frac{k}{2}\right)\Gamma(\sigma+1-k)}{2^{\sigma-k+1}\Gamma\left(\frac{\sigma-k}{2}+1\right)^2\Gamma\left(\sigma-\frac{k}{2}+\frac{n}{2}+1\right)} \\
&= e^{2n\rho} \frac{(2\pi)^{n/2}\alpha^{k+n}2^{k/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2}{(\sigma+n)^{\frac{k+n}{2}}\Gamma(\frac{n}{2})} \frac{\Gamma\left(\frac{n}{2}+\frac{k}{2}\right)\Gamma(\sigma+1-k)}{\Gamma\left(\frac{\sigma-k}{2}+1\right)^2\Gamma\left(\sigma-\frac{k}{2}+\frac{n}{2}+1\right)}.
\end{aligned}$$

Then, for $\sigma > 0$,

$$\begin{aligned}
&\mathbb{E}[|Y_n|] = \tfrac{1}{\|K_n\|_2^2} \int_{\mathbb{R}^n} |x| K_n(x)^2 \mathrm{d}x \\
&= \frac{(2\pi)^{\frac{n}{2}}\alpha^{1+n}2^{1/2}\Gamma\left(\frac{\sigma+n+2}{2}\right)^2\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)\Gamma(\sigma)}{(\sigma+n)^{\frac{1+n}{2}}\Gamma(\frac{n}{2})\Gamma\left(\frac{\sigma+1}{2}\right)^2\Gamma\left(\sigma-\frac{1}{2}+\frac{n}{2}+1\right)} \frac{(\sigma+n)^{\frac{n}{2}}\Gamma\left(\frac{\sigma}{2}+1\right)^2\Gamma\left(\sigma+\frac{n}{2}+1\right)}{(2\pi)^{\frac{n}{2}}\alpha^n\Gamma(\sigma+1)\Gamma\left(\frac{\sigma}{2}+\frac{n}{2}+1\right)^2} \\
&= \frac{\alpha 2^{1/2}}{(\sigma+n)^{1/2}\Gamma(\frac{n}{2})} \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)\Gamma(\sigma)\Gamma\left(\frac{\sigma}{2}+1\right)^2\Gamma\left(\sigma+\frac{n}{2}+1\right)}{\Gamma\left(\frac{\sigma}{2}+\frac{1}{2}\right)^2\Gamma\left(\sigma+\frac{n}{2}+\frac{1}{2}\right)\Gamma(\sigma+1)} \\
&\sim \frac{\alpha 2^{1/2}}{(\sigma+n)^{1/2}\Gamma(\frac{n}{2})} \frac{\Gamma\left(\frac{n}{2}\right)\left(\frac{n}{2}\right)^{\frac{1}{2}}\Gamma(\sigma)\Gamma\left(\frac{\sigma}{2}+1\right)^2\Gamma\left(\frac{n}{2}\right)\left(\frac{n}{2}\right)^{\sigma+1}}{\Gamma\left(\frac{\sigma+1}{2}\right)^2\Gamma\left(\frac{n}{2}\right)\left(\frac{n}{2}\right)^{\sigma+\frac{1}{2}}\Gamma(\sigma+1)} \\
&= \frac{\alpha 2^{1/2}}{(\sigma+n)^{1/2}} \frac{\left(\frac{n}{2}\right)\Gamma(\sigma)\Gamma\left(\frac{\sigma}{2}+1\right)^2}{\Gamma\left(\frac{\sigma+1}{2}\right)^2\Gamma(\sigma+1)} \sim n^{1/2}\frac{\alpha}{2^{1/2}} \frac{\Gamma(\sigma)\Gamma\left(\frac{\sigma}{2}+1\right)^2}{\Gamma\left(\frac{\sigma+1}{2}\right)^2\Gamma(\sigma+1)} = O(n^{\frac{1}{2}}).
\end{aligned}$$

Now, let $\beta > \frac{1}{2}$. By Markov's inequality,

$$\lim_{n\to\infty} \mathbb{P}[\eta_n(B_n(Rn^\beta)^c) > 0 | \eta_n \neq \emptyset] = \lim_{n\to\infty} \mathbb{P}\left(|Y_n| \geq Rn^\beta\right) \leq \lim_{n\to\infty} \frac{\mathbb{E}|Y_n|}{Rn^\beta} = 0.$$

For $\sigma = 0$, one can take any $k \in (0,1)$, and a similar analysis shows that $\mathbb{E}[|Y_n|^k] = O(n^{k/2})$, and the result still holds by applying Markov's inequality to $\mathbb{P}\left(|Y_n|^k \geq R^k n^{k\beta}\right)$.

A.9 Proof of Proposition 4.4.6

First, from [36, 6.576.3], we have for all $\nu > 0$ and $k > 2\nu - 1$,

$$\int_0^\infty r^k \mathbb{K}_\nu \left(\frac{r}{\alpha}\right)^2 dr = \frac{2^{-2+k}\alpha^{k+1}}{\Gamma(1+k)} \Gamma\left(\frac{1+k}{2} + \nu\right) \Gamma\left(\frac{k+1}{2}\right)^2 \Gamma\left(\frac{1+k}{2} - \nu\right), \tag{A.7}$$

where $\mathbb{K}_\nu$ is the modified Bessel function of the second kind. Then, for the Whittle-Matérn Kernel (4.16),

$$\begin{aligned}\int_{\mathbb{R}^n} K_n(x)^2 dx &= \int_{\mathbb{R}^n} e^{2n\rho} \frac{2^{2-2\nu}}{\Gamma(\nu)^2} \frac{|x|^{2\nu}}{\alpha^{2\nu}} \mathbb{K}_\nu \left(\frac{|x|}{\alpha}\right)^2 dx \\ &= \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} e^{2n\rho} \frac{2^{2-2\nu}}{\Gamma(\nu)^2\alpha^{2\nu}} \int_0^\infty r^{n-1} r^{2\nu} \mathbb{K}_\nu(r)^2 dr \\ &= e^{2n\rho} \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \frac{2^{2-2\nu}}{\Gamma(\nu)^2\alpha^{2\nu}} \int_0^\infty r^{n-1+2\nu} \mathbb{K}_\nu(r)^2 dr.\end{aligned}$$

Then by (A.7),

$$\begin{aligned}&\int_0^\infty r^{n-1+2\nu} \mathbb{K}_\nu(r)^2 dr \\ &= \frac{2^{-3+n+2\nu}\alpha^{n+2\nu}}{\Gamma(n+2\nu)} \Gamma\left(\frac{n+2\nu}{2} + \nu\right) \Gamma\left(\frac{n+2\nu}{2}\right)^2 \Gamma\left(\frac{n+2\nu}{2} - \nu\right) \\ &= \frac{2^{-3+n+2\nu}\alpha^{n+2\nu}}{\Gamma(n+2\nu)} \Gamma\left(\frac{n}{2} + 2\nu\right) \Gamma\left(\frac{n}{2} + \nu\right)^2 \Gamma\left(\frac{n}{2}\right).\end{aligned}$$

Similarly,

$$\int_{\mathbb{R}^n} |x|^2 K_n(x)^2 dx = e^{2n\rho} \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \frac{2^{2-2\nu}}{\Gamma(\nu)^2\alpha^{2\nu}} \int_0^\infty r^{n+1+2\nu} \mathbb{K}_\nu(r)^2 dr.$$

and also by (A.7),

$$\begin{aligned}&\int_0^\infty r^{n+1+2\nu} \mathbb{K}_n(r)^2 dr \\ &\quad = \frac{2^{-1+n+2\nu}\alpha^{n+2+2\nu}}{\Gamma(n+2+2\nu)} \Gamma\left(\frac{n}{2} + 2\nu + 1\right) \Gamma\left(\frac{n}{2} + \nu + 1\right)^2 \Gamma\left(\frac{n}{2} + 1\right)\end{aligned}$$

Then,

$$
\begin{aligned}
\mathbb{E}[|Y_n|^2] &= \frac{\int_{\mathbb{R}^n} |x|^2 K_n(x)^2 dx}{\int_{\mathbb{R}^n} K_n(x)^2 dx} \\
&= \frac{(2\alpha)^2 \Gamma(n+2\nu)\Gamma\left(\frac{n}{2}+2\nu+1\right)\Gamma\left(\frac{n}{2}+\nu+1\right)^2\Gamma\left(\frac{n}{2}+1\right)}{\Gamma(n+2+2\nu)\Gamma\left(\frac{n}{2}+2\nu\right)\Gamma\left(\frac{n}{2}+\nu\right)^2\Gamma\left(\frac{n}{2}\right)} \\
&= \frac{(2\alpha)^2\left(\frac{n}{2}+2\nu\right)\left(\frac{n}{2}+\nu\right)^2\left(\frac{n}{2}\right)}{(n+1+2\nu)(n+2\nu)} \sim \left(\frac{\alpha}{2}\right)^2 n,
\end{aligned}
$$

as $n \to \infty$, and this implies

$$
\frac{\mathbb{E}[|Y_n|^2]^{\frac{1}{2}}}{\sqrt{n}} \to \frac{\alpha}{2}, \text{ as } n \to \infty.
$$

Thus, since the Whittle Matérn kernel is log-concave, the conclusion holds by Theorem 4.3.2.

A.10 Proof of Proposition 4.4.7

First, recall the the beta function satisfies

$$
B(x,y) := \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t = \int_0^\infty t^{x-1}(1+t)^{-(x+y)}\mathrm{d}t = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.
$$

Then, for any $k \geq 0$,

$$
\begin{aligned}
\int_{\mathbb{R}^n} |x|^k K_n(x)^2 \mathrm{d}x &= \int_{\mathbb{R}^n} |x|^k \frac{e^{2n\rho}}{(1+|\frac{x}{\alpha_n}|^2)^{2\nu+n}}\mathrm{d}x \\
&= e^{2n\rho}\frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\int_0^\infty r^{n-1+k}\left(1+\frac{r^2}{\alpha_n^2}\right)^{-2\nu-n}\mathrm{d}r \\
&= e^{2n\rho}\frac{\pi^{n/2}}{\Gamma(\frac{n}{2})}\alpha_n^{n+k}\int_0^\infty t^{\frac{n}{2}-1+\frac{k}{2}}(1+t)^{-(2\nu+n)}\mathrm{d}t \\
&= e^{2n\rho}\frac{\pi^{n/2}}{\Gamma(\frac{n}{2})}\alpha_n^{n+k} B\left(\frac{n}{2}+\frac{k}{2}, 2\nu+\frac{n}{2}-\frac{k}{2}\right).
\end{aligned}
$$

Thus, the expectation of $|Y_n|^2$ is

$$
\begin{aligned}
\mathbb{E}[|Y_n|^2] &= \frac{1}{\|K_n\|_2^2}\int_{\mathbb{R}^n}|x|^2K_n(x)^2\mathrm{d}x = \alpha_n^2\frac{B(\frac{n}{2}+1, 2\nu+\frac{n}{2}-1)}{B(\frac{n}{2}, 2\nu+\frac{n}{2})} \\
&= \alpha_n^2\frac{\Gamma(\frac{n}{2}+1)\Gamma(2\nu+\frac{n}{2}-1)\Gamma(n+2\nu)}{\Gamma(n+2\nu)\Gamma(\frac{n}{2})\Gamma(2\nu+\frac{n}{2})} = \alpha_n^2\frac{n}{2(\frac{n}{2}+2\nu-1)} \\
&= \alpha_n^2\frac{n}{n+4\nu-2},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[|Y_n|^4] &= \alpha_n^4\frac{B(\frac{n}{2}+2, 2\nu+\frac{n}{2}-2)}{B(\frac{n}{2}, 2\nu+\frac{n}{2})} = \alpha_n^4\frac{\Gamma(\frac{n}{2}+2)\Gamma(2\nu+\frac{n}{2}-2)\Gamma(n+2\nu)}{\Gamma(n+2\nu)\Gamma(\frac{n}{2})\Gamma(2\nu+\frac{n}{2})} \\
&= \alpha_n^4\frac{(\frac{n}{2}+1)\frac{n}{2}}{(2\nu+\frac{n}{2}-2)(2\nu+\frac{n}{2}-1)} = \alpha_n^4\frac{n(n+2)}{(n+4\nu-4)(n+4\nu-2)}.
\end{aligned}
$$

Thus, by the assumption that $\alpha_n \sim \alpha n^{\frac{1}{2}}$ as $n \to \infty$ for some $\alpha > 0$,

$$
\lim_{n\to\infty}\frac{\mathbb{E}[|Y_n|^2]}{n} = \alpha^2 \text{ and } \lim_{n\to\infty}\frac{\mathrm{Var}(|Y_n|^2)}{n^2} = 0.
$$

Thus, by Chebychev's inequality, $\frac{|Y_n|}{\sqrt{n}} \to \alpha$ in probability.

A.11 Proof of Proposition 4.5.1

By Proposition 4.4.2,

$$
\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n}\ln\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0] \\
&= \begin{cases} -\rho - \frac{1}{2}\log 2\pi e + \frac{2R^2}{\alpha^2 m} - \log R, & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ -\rho - \log\alpha - \frac{1}{2}\log\frac{m\pi}{2}, & R > \sqrt{m}\frac{\alpha}{2}. \end{cases}
\end{aligned}
$$

Recall that $\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[X_n(B_n(\sqrt{n}R))] = \rho + \frac{1}{2}\log 2\pi e + \log R$. Thus,

$$
\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n} \ln \frac{\mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0]}{\mathbb{E}[X_n(B_n(\sqrt{n}R))]} \\
&= \lim_{n\to\infty} -\frac{1}{n} \ln \mathbb{P}[\eta_n(B_n(\sqrt{n}R)) > 0] + \frac{1}{n} \ln \mathbb{E}[X_n(B_n(\sqrt{n}R))] \\
&= \begin{cases} -\rho - \frac{1}{2}\log 2\pi e + \frac{2R^2}{\alpha^2 m} - \log R + \rho + \frac{1}{2}\log 2\pi e + \log R, & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ -\rho - \log\alpha - \frac{1}{2}\log\frac{m\pi}{2} + \rho + \frac{1}{2}\log 2\pi e + \log R, & R > \sqrt{m}\frac{\alpha}{2} \end{cases} \\
&= \begin{cases} \frac{2R^2}{\alpha^2 m}, & 0 < R < \sqrt{m}\frac{\alpha}{2} \\ \frac{1}{2} + \log 2 - \log\alpha - \frac{1}{2}\log m + \log R, & R > \sqrt{m}\frac{\alpha}{2}. \end{cases}
\end{aligned}
$$

Appendix B

Laplace Method

Lemma B.0.1. *Let $f(t)$ be a function such that $f(t)$ achieves its minimum at $t = a$ on the interval $[a, b)$ and $f'(t)$ is continuous. Also assume* $\lim_{n\to\infty} a_n = a \in \mathbb{R}$*. If $f'(a) > 0$, then as $n \to \infty$*

$$\int_{a_n}^{b} e^{-nf(t)}dt \sim \frac{e^{-nf(a_n)}}{nf'(a)}.$$

If $f(t)$ achieves its minimum at $t = b$ over the interval $(a, b]$, $\lim_{n\to\infty} b_n = b$*, and $f'(b) < 0$, then*

$$\int_{a}^{b_n} e^{-nf(t)}dt \sim -\frac{e^{-nf(b_n)}}{nf'(b)}.$$

Proof. Let $\varepsilon > 0$. By the continuity of f', there exists $\min_n(b - a_n) > \delta > 0$ such that $|t - a| < 2\delta$ implies $f'(t) \leq f'(a) + \varepsilon$. By Taylor's theorem, for each $t \in [a_n, b]$, there is some $\xi_t \in (a_n, t)$ such that

$$f(t) = f(a_n) + f'(\xi_t)(t - a_n).$$

Then, for t such that $|t - a_n| < \delta$ and n large enough such that $|a_n - a| < \delta$, we have that $|\xi_t - a_n| < \delta$, and thus by the triangle inequality, $|\xi_t - a| < 2\delta$, which implies that for all n large enough,

$$f(t) \leq f(a_n) + (f'(a) + \varepsilon)(t - a_n).$$

Since the integrand is positive,

$$\int_{a_n}^{b} e^{-nf(t)}dt \geq \int_{a_n}^{a_n+\delta} e^{-nf(t)}dt \geq \int_{a_n}^{a_n+\delta} e^{-n(f(a_n)+(f'(a)+\varepsilon)(t-a_n))}dt$$
$$= e^{-nf(a_n)} \int_{0}^{\delta n(f'(a)+\varepsilon)} e^{-y}dy = \frac{e^{-nf(a_n)}}{n(f'(a)+\varepsilon)}\left(1-e^{-\delta n(f'(a)+\varepsilon)}\right).$$

Then,

$$\liminf_{n\to\infty} \frac{\int_{a_n}^{b} e^{-nf(t)}dt}{\frac{e^{-nf(a_n)}}{n(f'(a)+\varepsilon)}} \geq \liminf_{n\to\infty}\left(1-e^{-\delta n(f'(a)+\varepsilon)}\right) = 1.$$

For $\delta > 0$, let N be such that for all $n > N$, $a_n + \delta > a + \frac{\delta}{2}$. Then, for the upper bound, define

$$C := \inf_{t\in[a+\frac{\delta}{2},b]} f(t) > f(a),$$

where the last inequality follows from the hypothesis that f achieves its minimum at a on the interval $[a,b)$. By a similar Taylor series argument, we have for all $|t-a_n| < \delta$, and n large enough,

$$f(t) \geq f(a_n) + (f'(a)-\varepsilon)(t-a_n).$$

Define $\eta := C - f(a) > 0$. Then, for all $t \in [a+\frac{\delta}{2}, b]$, $f(t) > f(a)+\eta$. Then, for all n large enough,

$$\int_{a_n}^{b} e^{-nf(t)}dt = \int_{a_n}^{a_n+\delta} e^{-nf(t)}dt + \int_{a_n+\delta}^{b} e^{-nf(t)}dt$$
$$\leq \int_{a_n}^{a_n+\delta} e^{-n(f(a_n)+(f'(a)-\varepsilon)(t-a_n)}dt + \int_{a+\frac{\delta}{2}}^{b} e^{-nC}dt$$
$$< (b-a)e^{-nC} + \frac{e^{-nf(a_n)}}{n(f'(a)-\varepsilon)}\int_0^{\delta n(f'(a)-\epsilon)} e^{-y}dy$$
$$= (b-a)e^{-nC} + \frac{e^{-nf(a_n)}}{n(f'(a)-\varepsilon)}\left(1-e^{-\delta n(f'(a)-\varepsilon)}\right).$$

Then,

$$\limsup_{n\to\infty} \frac{\int_{a_n}^{b} e^{-nf(t)}dt}{\frac{e^{-nf(a)}}{n(f'(a)-\epsilon)}} \leq \limsup_{n\to\infty}\{(b-a)n(f'(a)-\varepsilon)e^{-n\eta}+1-e^{-\delta n(f'(a)+\varepsilon)}\}=1,$$

since $\eta > 0$. These limits hold for all ε, and thus,

$$\lim_{n\to\infty} \frac{\int_{a_n}^{b} e^{-nf(t)}dt}{\frac{e^{-nf(a_n)}}{nf'(a)}} = 1.$$

□

Lemma B.0.2. *Assume that* $\lim_{n\to\infty} c_n = c \in (0,\infty)$. *Then, as* $n \to \infty$, *if* $c > 1$,

$$\Gamma_u(n+1, c_n n) \sim n^n \frac{ce^{n(\ln c_n - c_n)}}{(c-1)},$$

and if $c < 1$,

$$\Gamma_\ell(n+1, c_n n) \sim n^n \frac{ce^{n(\ln c_n - c_n)}}{(1-c)}.$$

Proof. By a change of variables,

$$\Gamma_u(n+1, c_n n) = \int_{c_n n}^{\infty} e^{-t}t^n dt = n^{n+1}\int_{c_n}^{\infty} e^{-n(y-\log y)}dy.$$

For $c > 1$, letting $f(y) = y - \log(y)$, f has a minimum at c on the interval $[c,\infty)$. Also observe that $f'(y) = 1 - \frac{1}{y}$, and for $c > 1$, $f'(c) = 1 - \frac{1}{c} > 0$. Thus, by Lemma B.0.1,

$$\Gamma_u(n+1, c_n n) \sim n^{n+1}\frac{e^{-n(c_n - \ln c_n)}}{n\left(1-\frac{1}{c}\right)} = n^n \frac{ce^{n(\ln c_n - c_n)}}{(c-1)}.$$

Similarly,

$$\Gamma_\ell(n+1, c_n n) = \int_{0}^{c_n n} e^{-t}t^n dt = n^{n+1}\int_{0}^{c_n} e^{-n(y-\log y)}dy,$$

and for $c < 1$, $f(y) = y - \log(y)$ hits its minimum at c on the interval $(0, c]$ and $f'(c) = 1 - \frac{1}{c} < 0$. Then, by Lemma B.0.1,

$$\Gamma_\ell(n+1, c_n n) \sim n^n \frac{ce^{n(\ln c_n - c_n)}}{(1-c)}.$$

$\square$

Bibliography

[1] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.14 of 2017-09-18. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[2] K. Alishahi and M. Sharifitabar. Volume degeneracy of the typical cell and the chord length distribution for Poisson-Voronoi tessellations in high dimensions. *Advances in Applied Probability*, 40:919–938, 2008.

[3] V. Anantharam and F. Baccelli. Capacity and error exponents of stationary point processes under random additive displacements. *Advances in Applied Probability*, 47(1):1–26, 2015.

[4] V. Anantharam and F. Baccelli. The Boolean model in the Shannon regime: Three thresholds and related asymptotics. 53(4):1001–1018, 2016.

[5] F. Baccelli and E. O'Reilly. Reach of repulsion for determinantal point processes in high dimensions. *Journal of Applied Probability*, 55(3), September 2018.

[6] F. Baccelli and E. O'Reilly. The stochastic geometry of unconstrained one-bit data compression. *arXiv:1810.06095*, 2018.

[7] R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. Exponential decay of reconstruction error from binary measurements of sparse signals. *IEEE Transactions on Information Theory*, 63(6):3368–3385, March 2017.

[8] D. Bilyk and M. T. Lacey. Random tessellations, restricted isometric embeddings, and one bit sensing. *arXiv:1512.06697*, December 2015.

[9] C. Biscio and F. Lavancier. Quantifying repulsiveness of determinantal point processes. *Bernoulli*, 22(4):2001–2028, 2016.

[10] B. Blaszczyszyn and D. Yogeshwaran. On comparison of clustering properties of point processes. *Advances in Applied Probability*, 46(1):1–20, 2014.

[11] S. Bobkov. *Geometric Aspects of Functional Analysis*, volume 1807 of *Lecture Notes in Mathematics*, chapter Spectral gap and concentration for some spherically symmetric probability measures, pages 37–43. Springer, 2003.

[12] A. Borodin and G. Olshanski. Distributions on partitions, point processes and the hypergeometric kernel. *Communications in Mathematical Physics*, 211:335–358, 2000.

[13] P. T. Boufounos and R. G. Baraniuk. 1-bit compressive sensing. Princeton, NJ, March 2008. Conference on Information Science and Systems (CISS).

[14] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.-H. Vritsiou. *Geometry of Isotropic Convex Bodies*, volume 196 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2014.

[15] P. Calka. The distribution of the smallest disks containing the Poisson-Voronoi typical cell and the Crofton cell in the plane. *Advances in Applied Probability*, 34:702–717, 2002.

[16] P. Calka. *Poissonian Tessellations of the Euclidean space. An extension of a result of R.E. Miles. In: De nouveaux résultants sur la géométrie des mosaïques de Poisson-Voronoi et des mosaïques poissoniennes d'hyperplans*. PhD thesis, Universite Lyon 1, September 2002.

[17] P. Calka. Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional Poisson–Voronoi tessellation and a Poisson line process. *Advances in Applied Probability*, 35:551–562, 2003.

[18] P. Calka. *New Perspectives in Stochastic Geometry*, chapter 5: Tessellations, pages 145–164. Oxford University Press, 2010.

[19] E. Candes and T. Tao. Near-optimal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 12:5406–5425, 2006.

[20] I. Cascos. *New Perspectives in Stochastic Geometry*, chapter 12: Data depth: multivariate statistics and geometry. Oxford University Press,

2010.

[21] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley Series in Probability and Statistics. Wiley, third edition, 2013.

[22] J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. Palm distributions for log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 44(1), October 2016.

[23] J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. A tutorial on Palm distributions for spatial point processes. *International Statistical Review*, 85:404–420, 2017.

[24] B. Cousins and S. Vempala. Gaussian cooling and $o^*(n^3)$ algorithms for volume and Gaussian volume. *SIAM Journal on Computing*, 47(3):1237–1273, 2018.

[25] A. Cuevas and R. Fraiman. *New Perspectives in Stochastic Geometry*, chapter 11: Set estimation. Oxford University Press, 2010.

[26] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer-Verlag, New York, second edition, 2003.

[27] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, 2nd edition, 1998.

[28] N. Deng, W. Zhou, and M. Haenggi. The Ginibre point process as a model for wireless networks with repulsion. *IEEE Transactions on Wireless Communications*, 14:107–121, 2015.

[29] G. Döge, K. Mecke, J. Møller, D. Stoyan, and R.P. Waagepetersen. Grand canonical simulations of hard-disk systems by simulated tempering. *International Journal of Modern Physics C*, 15:129–147, 2004.

[30] R. Eldan and B. Klartag. Approximately Gaussian marginals and the hyperplane conjecture. volume 545, pages 55–68, Providence, RI, 2011. Americal Mathematical Society, Contemporary Mathematics.

[31] M. Fradelizi, O. Guédon, and A. Pajor. Thin-shell concentration for convex measures. *Studia Mathematica*, 223(2), 2014.

[32] H.-O. Georgii and H. J. Yoo. Conditional intensity and Gibbsianness of determinantal point processes. *Journal of Statistical Physics*, 118(1-2):55–84, 2005.

[33] J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6:440–449, 1965.

[34] S. Glasauer and R. Schneider. Asymptotic approximation of smooth convex bodies by polytopes. *Forum Mathematicum*, 8:363–377, 1996.

[35] A. Goldman. The Palm measure and the Voronoi tessellation for the Ginibre process. *Annals of Applied Probability*, 20:90–128, 2000.

[36] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, series, and products.* Elsevier, seventh edition, 2007.

[37] L. Grafakos. *Classical Fourier Analysis.* Graduate Texts in Mathematics. Springer, 2008.

[38] O. Guédon and E. Milman. Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures. *Geometric and Functional Analysis*, 2011.

[39] J. Hörrmann and D. Hug. On the volume of the zero cell of a class of isotropic Poisson hyperplane tessellations. *Advances in Applied Probability*, 46(3):622–642, September 2014.

[40] J. Hörrmann, D. Hug, M. Reitzner, and C. Thäle. Poisson polyhedra in high dimensions. *Advances in Mathematics*, 281:1–39, 2015.

[41] J. B. Hough, M. Krishnapur, Y. Peres, and B. Viràg. Determinantal processes and independence. *Probability Surveys*, 3:206–229, 2006.

[42] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes.* American Mathematical Society, Providence, 2009.

[43] D. Hug, M. Reitzner, and R. Schneider. The limit shape of the zero cell in a stationary Poisson hyperplane tessellation. *The Annals of Probability*, 32(1B):1140–1167, 2004.

[44] K. Johansson. The arctic circle boundary and the Airy process. *The Annals of Probability*, 33:1–30, 2005.

[45] Z. Kabluchko, A. Marynych, D. Temesvari, and C. Thäle. Cones generated by random points on half-spheres and convex hulls of Poisson point processes. *arXiv:1801.08008v1*, January 2018.

[46] Z. Kabluchko, C. Thäle, and D. Zaporozhets. Beta polytopes Poisson polyhedra: f-vectors and angles. *arXiv:1805.01338v1*, May 2018.

[47] B. Klartag. A central limit theorem for convex sets. *Inventiones Mathematicae*, 168:91–131, 2007.

[48] K. Knudson, R. Saab, and R. Ward. One-bit compressive sensing with norm estimation. *IEEE Transaction on Information Theory*, 62(5), May 2016.

[49] A. Kulesza and B. Taskar. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning*, 5(2-3):123–286, 2012.

[50] T. Kuna, J.L. Lebowitz, and E.R. Speer. Realizability of point processes. *Journal of Statistical Physics*, 129(3):417–439, November 2007.

[51] F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference (extended version). Technical report, Available on arXiv:1205.4818, 2014.

[52] F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(4):853–877, 2015.

[53] Y. Li, F. Baccelli, H. S. Dhillon, and J. G. Andrews. Statistical modeling and probabilistic analysis of cellular networks with determinantal point processes. *IEEE Transactions on Communications*, 63(9):3405–3422, July 2015.

[54] T. Lindvall. On Strassen's theorem on stochastic domination. *Electronic Communications in Probability*, 4:51–59, 1999.

[55] R. Lyons. Determinantal probability measures. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 98:167–212, 2003.

[56] R. Lyons. Determinantal probability: Basic probability and conjectures. *Proceedings of International Congress of Mathematicians*, IV:137–161, 2014.

[57] O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7(1):83–122, March 1975.

[58] S. Mase, J. Møller, D. Stoyan, R.P. Waagepetersen, and G. Döge. Packing densities and simulated tempering for hard core Gibbs point processes. *Annals of the Institute of Statistical Mathematics*, 53:661–680, 2001.

[59] P. McCullagh and J. Møller. The permanental process. *Advances in Applied Probability*, 38:873–888, 2006.

[60] J. Møller, M. Nielsen, E. Porcu, and E. Rubak. Determinantal point process models on the sphere. *Bernoulli*, 24:1171–1201, 2018.

[61] J. Møller and E. O'Reilly. Couplings for determinantal point processes and their reduced Palm distribution with a view to quantifying repulsiveness. *arXiv:1806.07347*, June 2018.

[62] J. Møller and E. Rubak. Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spatial Statistics*, 18:4–23, 2016.

[63] J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[64] C. Müller. *Analysis of Spherical Symmetries in Euclidean Spaces*, volume 129 of *Applied Mathematical Sciences*. Springer, New York, 1998.

[65] B. S. Nagy, C. Foias, H. Bercovici, and L. Kérchy. *Harmonic Analysis of Operators on Hilbert Space*. Springer-Verlag, New York, 2010.

[66] E. O'Reilly. Thin-shell concentration for zero cells of stationary Poisson mosaics. *arXiv:1809.04134*, September 2018.

[67] E. O'Reilly, F. Baccelli, G. De Veciana, and H. Vikalo. End-to-end optimization of high-throughput DNA sequencing. *Journal of Computational Biology*, 23(10), 2016.

[68] V. Paulsen. *Completely Bounded Maps and Operator Algebras.* Cambridge University Press, Cambridge, 2002.

[69] Y. Plan and R. Vershynin. One-bit compressed sensing by linear programming. *Communications on Pure and Applied Mathematics*, 66:1275–1297, 2013.

[70] Y. Plan and R. Vershynin. Dimension reduction by random hyperplane tessellations. *Discrete and Computational Geometry*, 51:438–461, 2014.

[71] G. Poltyrev. On coding without restrictions for the AWGN channel. *IEEE Transactions on Information Theory*, 40(2):409–417, March 1994.

[72] B. D. Ripley. The second-order analysis of stationary point process. *Journal of Applied Probability*, 13(2):255–266, June 1976.

[73] W. Rudin. *Functional Analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., second edition, 1991.

[74] A. Scardicchio, C. E. Zachary, and S. Torquato. Statistical properties of determinantal point processes in high-dimensional Euclidean spaces. *Physical Review*, 79, 2009.

[75] R. Schneider and W. Weil. *Stochastic and Integral Geometry.* Springer, 2008.

[76] I. J. Schoenberg. Positive definite functions on spheres. *Duke Mathematical Journal*, 9:96–108, 1942.

[77] T. Shirai and Y. Takahashi. Random point fields associated with certain Fredholm determinants I: Fermion, Poisson, and Boson point processes. *Journal of Functional Analysis*, 205:414–463, 2003.

[78] J. Snoek, R. S. Zemel, and R. P. Adams. A determinantal point process latent variable model for inhibition in neural spiking data. In I. S. Francis, B. J. F. Manly, and F .C. Lam, editors, *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, pages 1932–1940. Electronic proceedings from the conference, "Neural Information Processing Systems 2013", 2013.

[79] A. Soshnikov. Determinantal random point fields. *Russian Mathematical Surveys*, 55:923–975, 2000.

[80] S. Torquato, A. Scardicchio, and C. E. Zachary. Point processes in arbitrary dimension from Fermionic gases, random matrix theory, and number theory. *Journal of Statistical Mechanics*, 2008.

[81] D. Yogeshwaran and R. J. Adler. On the topology of random complexes built over stationary point processes. *Annals of Applied Probability*, 25(6):3338–3380, 2015.

[82] Y. Yu. On normal variance-mean mixtures. *Statistics and Probability Letters*, 121:45–50, February 2017.