Error Exponents in the AWGN Channel

A Project Report

submitted by

MALI SUNDARESAN S

in partial fulfilment of the requirements
for the award of the degrees of

BACHELOR OF TECHNOLOGY
and
MASTER OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY MADRAS

JUNE 2014

THESIS CERTIFICATE

This is to certify that the thesis titled **Error Exponents in the AWGN Channel**, submitted by **Mali Sundaresan S**, to the Indian Institute of Technology, Madras, for the award of the degree of **Bachelor of Technology** and **Master of Technology**, is a bonafide record of the research work done by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Radhakrishna Ganti
Project Guide
Assistant Professor
Dept. of Electrical Engineering
IIT-Madras, 600 036

Place: Chennai

Date: 13th June 2014

ACKNOWLEDGEMENTS

When I started this project a year back, I had heard of neither point process nor error exponents before. Today I have my own thesis on these topics. It is all down to my guide, Prof. Radhakrishna Ganti. He gave me the right reference materials, and enough time to get a hang of things. The most important thing was his optimism and encouragement even at times when things weren't going in the right direction.

My thanks to the Department of Electrical Engineering for providing excellent lab and computer facilities. My sincere thanks to the professors of the department who taught me various courses both at undergraduate and post graduate level . I would also like to thank my faculty advisor Dr. Srikrishna Bashyam for his guidance over the past five years.

The Network and Stochastic Systems Lab is easily the best workplace I've worked in and it's because of the awesome people. I'll definitely come back to work here if the opportunity presents.

Special thanks to my family and friends ("Dhendams") for being the support and inspiration in my life.

Lastly, thanks to all the scientists who are or have ever been out there, from Aristotle to the guys who worked in my field - Shannon or Kabatiansky. The methodology of Science, of passing rational conclusions across generations of human beings is mankind's greatest legacy and I am proud and thankful to have been a part of it.

ABSTRACT

In Shannon's 1958 paper, he finds the upper and lower bounds on the probability of error of optimal codes in the AWGN Channel. When rate is less than Capacity, these bounds are exponentials of the form $C\exp(-nE)$ with E (called error exponent) positive. However, there is a gap between upper and lower bounds on the error exponent E for lower rate cases meaning the exact exponent is unknown for these code rates. A similar gap exists in the coding without restrictions in n dimensional space scenario considered by Polytrev. Both these problems are related to each other as well as to the long standing question of optimal packing of spheres in n dimensions.

We use the theory of stochastic geometry and point processes to attack Codes without Restrictions. It can be shown that using Poisson and Matern Point Process to model random codes together leads to the best known bounds [6]. The repulsive nature of Determinantal Point Processes compared to Poisson make them a better model and we try to construct appropriate determinantal point process to get better bounds. However it becomes clear that they aren't useful because of too many resrictions on the kernel function. This leads us to look for general constraints for first and second moment densities of stationary point processes. We find a valid stationary point process that matches the best known bounds of today. This is the only continuous intensity function to give the best known exponent. Besides, this result gives hope to construct similar functions that might lead to better lower bounds.

TABLE OF CONTENTS

Part I

BACKGROUND

CHAPTER 1

CODING WITH RESTRICTIONS

We start our research by looking back at Shannon's classic paper [1]. Here, Shannon considers optimal codes in the Additive White Gaussian Noise (AWGN) channel under Maximum Likelihood Estimation (MLE) decoding system. Upper bounds and lower bounds on the probability of error for optimal codes P_e are computed and it is found that the bounds are close together for rates near channel capacity as well as zero, but diverge in between.

1.1 The Problem

1.1.1 Statement

1. Code is a set of M points in n dimensional space under the restriction each code word be on the surface of sphere of radius $\sqrt{nP}$, with P the signal power.
2. Noise is AWGN with noise power N. This means the code points are randomly displaced by an n dimensional Gaussian probability distribution of variance N in each dimension.
3. P_e is Probability of error of a code.

$$P_e \triangleq \sum_{i=1}^{M} \frac{1}{M} \text{[Probability(Error given i th code word sent)]}$$

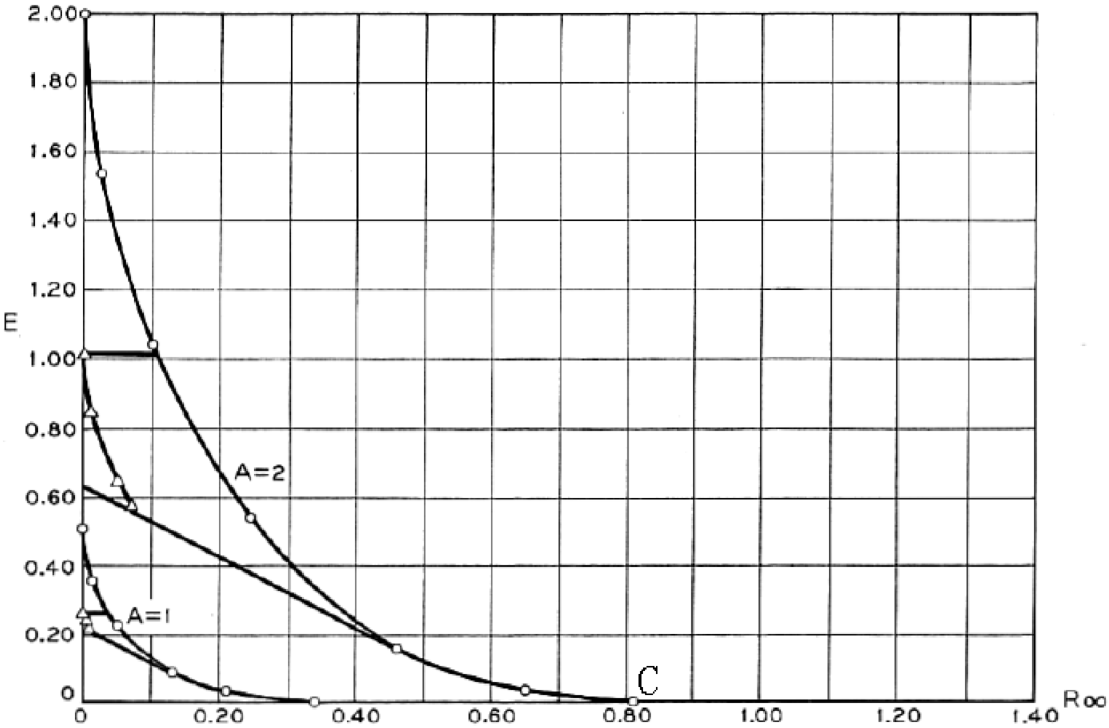
4. In this situation what is the probability of error for optimal codes? In other words what is the theoretical exponent of probability of error of these optimal codes?

1.1.2 Some Definitions

1. A is the signal to noise ratio parameter defined as $\sqrt{\frac{P}{N}}$. It will turn out useful because all expressions depend only on the P and N ratio.
2. C is channel capacity $\frac{1}{2}\log(A^2+1)$
3. R is rate of the code defined as $\frac{1}{n}\log(M)$ and in all our discussion $0 \le R \le C$
4. $E(R)$ is error exponent (also called reliability function) $\limsup_{n\to\infty} -\frac{1}{n}\log(P_{e,opt}(R,n))$, where $P_{e,opt}$ is the lowest possible P_e

1.2 Shannon's bounds

The figures [1] show the error exponent bounds from Shannon for two values of SNR A.



It can be seen that the lower and upper bounds on the exponent coincide for

$$\frac{1}{2}\log(\frac{1}{2}+\frac{A}{4}+\frac{1}{2}\sqrt{\frac{A^2}{4}+1}) \leq R \leq C = \frac{1}{2}\log(A^2+1)$$

At all lower rates the exact exponent is unknown.

1.3 New Upper Bound

An improved bound [2] to the minimum distance upper bound of Shannon was found in 2000. The figure [2] below shows the different bounds known today.

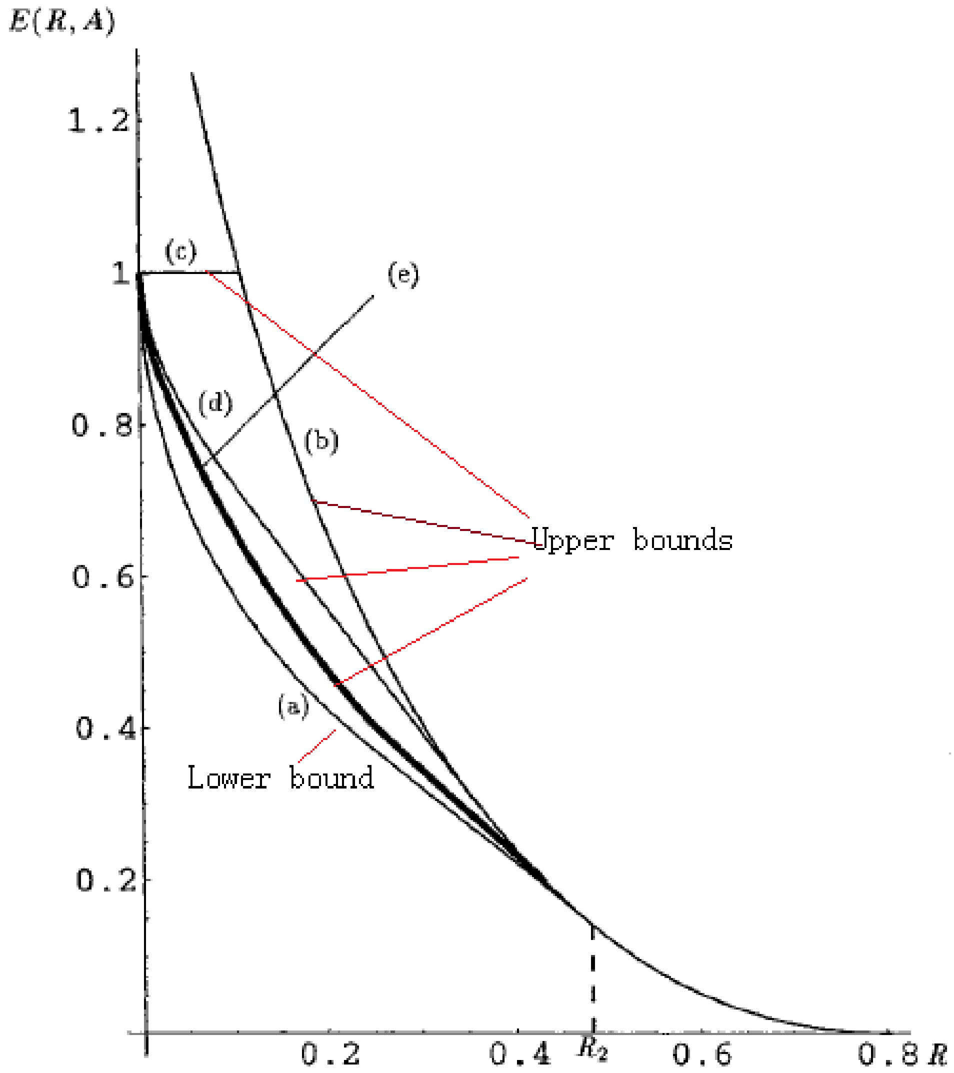


Fig. . Bounds on the reliability function ($A = 4$). (a) "Random coding" exponent (10.I–III). (b) Sphere-packing bound (9). (c) Minimum-distance bound (11). (d) Minimum-distance bound (12). (e) The new bound (14). R_2 denotes the critical rate. Each of the curves (d)–(e) includes a segment of the common tangent to the curve and the sphere-packing exponent.

1.4 Ideas Used

Let's look at the ideas used to derive the various bounds

1.4.1 Random Coding Lower Bound ((a) in the figure)

Consider this experiment: Take the M points of the code one by one and put them independently and randomly with probability measure propotional to the surface area, on the sphere of radius $\sqrt{nP}$. If you keep repeating this experiment, all possible codes are got. And then the average probability of error is calculated across all possible codes. The contribution to the average probability of error due to the first code point is $-\frac{1}{M}\left\{1-[1-\Omega(\theta)/\Omega(\pi)]^{M-1}\right\} dQ(\theta)$, where $Q(\theta)$ is defined as the probability of uniform gaussian noise taking the signal point out of the cone of half angle θ with axis passing through the signal point. Integrating over all θ and adding contribution from other M-1 points gives the lower bound on the error exponent (reliability function).

1.4.2 Upper Bounds on the Reliability function ((b) - (e))

Probability of error is given by $1/M \sum_i P(error\, given\, i^{th} codepoint\, was\, sent)$, assuming each point has the same probability $1/M$ of being picked. Each of the terms in the summation can be calculated once we know the Voronoi tesselation, by integrating the probability of gaussian noise outiede the voronoi cell of the i^{th}code point. But it is not possible to calculate a general expression for this as tesselations are complex and very different for different codes.

Sphere Packing bound (b)

Circles are easier to handle than polygonal Voronoi cells. So draw a circle centered around the code point with the same surface area as the polygon. The probability of gaussian noise taking the centre signal point outside the circle is lesser than that of noise taking it outside the polygon (because integral of gaussian is lesser inside the polygon than the circle). The probability of noise taking signal outside the circle (which is a cone in 3d and higher) is our $Q(\theta)$ we saw before. So by finding approximations for Q, Shannon got this upper bound on the error exponent.

Minimum Distance bound ((c) and (d))

Same problem as before: to calculate the probability of noise taking the i^{th}code point outside its Voronoi cell. At lower rates, the geometry doesn't seem to matter as much as the closest neighbour distance. So for each point consider only the probability of noise taking the codeword closer to its nearest neighbout than itself. This is nothing but integral of the gaussian over a half plane, which is much easier to calculate for each point than the integral outside voronoi cell.

New Upper Bound (e)

Listyn did better than considering only the nearest neighbour. He used results done on the distance distributions of points in spaces[3] to get better bounds. In effect, he used how many neighbours a point has on an average(instead of taking only one neighbour as in minimum distance), and used that to get tighter bounds.

CHAPTER 2

CODING WITHOUT RESTRICTIONS

2.1 The Problem

This modified version of the Shannon problem was first considered by Polytrev [4].

2.1.1 Code as Infinite Constellation(IC)

Consider any countable collection of points in $\mathbb{R}^n$ as our code S. Let $Cube_n(a)$ denote the n dimesional cube containing points with coordinate magnitude not exceeding $a/2$. A quantity of importance here is the normalized logarithmic density of a code defined as

$$\rho(S) \; = \frac{1}{n}(ln(\limsup_{a\to\infty}\frac{\mid S \cap Cube_n(a) \mid}{a^n}))$$

which is in effect a measure of number of codewords in a given volume.

2.1.2 Error

AWG(Additive White Gaussian) Noise displaces our code points randomly. Maximum Likelihood estimation is used by the decoder for determining the code point sent. The conditional probability of error when any point s is sent is:

$$\zeta(s) = Pr\left\{d(s, s+z) \geq d(s^{'}, s+z)\, for\, some\, s^{'} \in S, s^{'} \neq s\right\}$$

This (just like in the Shannon regime) is nothing but the gaussian noise integrated outside the voronoi cell of s.

The average probability of error of code S is given by

$$\zeta(S) = \limsup_{a\to\infty} \frac{1}{\mid S \cap Cube_n(a) \mid} \sum_{s\in S\cap Cube_n(a)} \zeta(s)$$

where $\zeta(s)$ are the conditional probability of noise taking the point s outside its voronoi cell. $\zeta(s)$ for the individual code points depend on the variance per coordinate σ^2 of the AWGN. Analogous to finding sets of achievable (rate,error probability) pairs for the Shannon regime, our objective here is to find the set of achievable (ρ, ζ) over the choice of S.

2.1.3 Polytrev Capacity

The generalized capacity (or Poltrev Capacity) of AWGN channel without restrictions C_∞ is defined as the largest number such that for $\rho < C_\infty$ there exists, for a sufficiently large n, an IC with arbitrarily small decoding error probability. Polytrev proved that

$$C_\infty = \frac{1}{2} ln(\frac{1}{2\pi e \sigma^2})$$

2.1.4 Error Exponent

The reliability function (or error exponent) is defined by the equation below:

$$\eta(\rho, \sigma^2) = -lim_n(\frac{1}{n} log\zeta_{opt}(n, \rho, \sigma^2))$$

where ζ_{opt} is the infimum of $\zeta(S)$ over all codes having parameters (n, ρ, σ^2). Just like we found that the Shannon reliability function depended only on the ratio of signal power to noise power, here we see that the error depends only on α defined by

$$\alpha = \frac{e^{-2\rho}}{2\pi e \sigma^2}$$

α is called the generalized signal to noise ratio. (It is easy to visualize the fact that the error probability depends only on α. Any change to ρ and σ that preserves the α ratio, is a geometrically identical situation). Thus the function $\eta(\rho, \sigma^2)$ is written simply as $\eta(\alpha)$. From here on $\eta(\alpha)$ will be called the reliability function of the AWGN channel without restrictions. Note that in Shannon regime there was a dependence on Code Rate besides the SNR. In this case there is dependence only on GSNR.

2.2 Bounds

The lower bound of the reliability function is called the Polytrev exponent $\pi(\alpha)$. It is obtained using the random coding argument similar to previous regime.

$$\pi(\alpha) = \begin{cases} \frac{\alpha^2}{2} - \frac{1}{2} - ln(\alpha) & if\ 1 \leq \alpha < \sqrt{2} \\ \frac{1}{2} - ln2 + ln(\alpha) & if\ \sqrt{2} \leq \alpha < 2 \\ \frac{\alpha^2}{8} & if\ \alpha \geq 2 \end{cases}$$

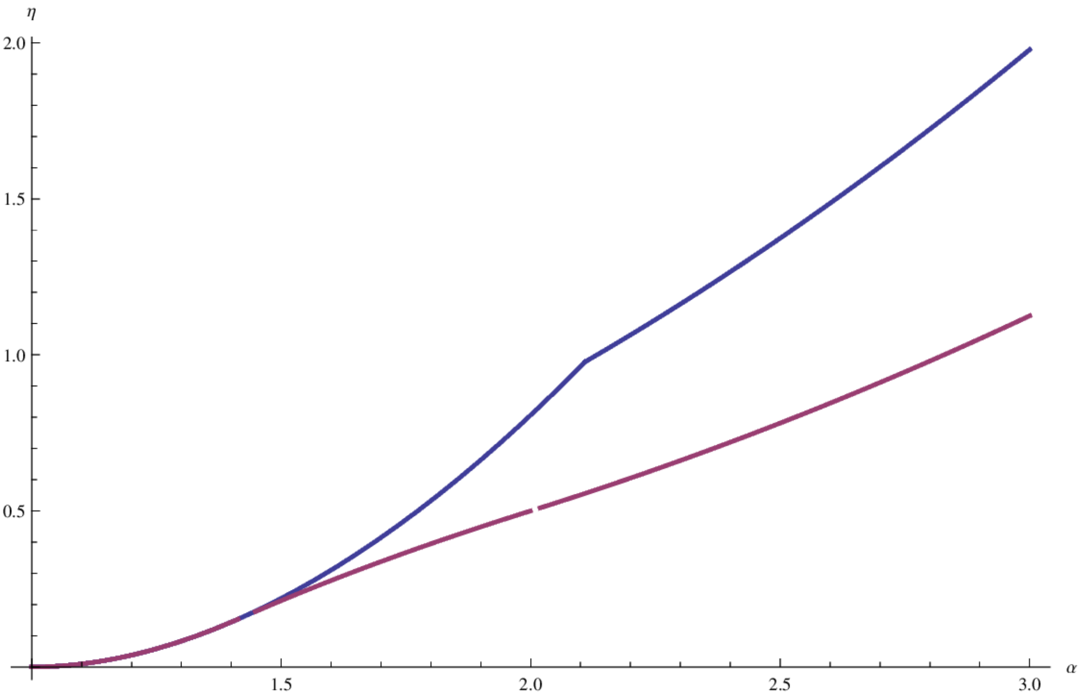
Polytrev also gave an upper bound using sphere packing for all $\alpha \geq 1$

$$\overline{\eta}(\alpha) \leq \frac{\alpha^2}{2} - \frac{1}{2} - ln(\alpha).$$

An improvement to this can be obtained for high α using Cor. 2 of [3]

$$\overline{\eta}(\alpha) \leq \frac{(0.663)^2\alpha^2}{2}$$

The following plot shows the best known bounds on error exponent versus α



There is a gap between the bounds for lower rates (high α) just like the Shannon regime.

2.3 Connection with restricted coding

Improvements in the bounds of Polytrev regime helps in the problem of gap between lower and upper bounds on the reliability function of the traditional AWGN channel. After all, the restricting codes to surface of a sphere is simply considering a space of one dimension less. We give a formal mathematical connection below.

2.3.1 Theorem

Let $E(R, A)$ denote the error exponents of the Shannon regime with code rate R and SNR A. With the generalized SNR of the Polytrev regime α let the error exponent be $\eta(\alpha)$. Then we have the following from [5]

$$\lim_{A\to\infty} E(\frac{1}{2}ln(\frac{1+A^2}{\alpha^2})\,,\ A) = \eta(\alpha)$$

2.3.2 Getting better bounds for Shannon regime

This the case of interest to us. We'll try to get better upper bounds for η using various new techniques. An improvement in the gap between the upper and lower bounds for $\eta(\alpha)$ will lead to improvement in the gap between the upper and lower bounds for the reliability function of the AWGN channel, atleast for large signal-to-noise ratios. In the rest of our study while we'll be looking only at the problem of error exponents in space, we need to understand any result obtained can be extended to Shannon regime.

CHAPTER 3

PALM APPROACH TO LOWER BOUNDS

Refer Appendix A for Point Process theory.

- Our code in Palm approach [5] will be a Point Process ϕ. The rate of code $R = \frac{1}{2}ln(\frac{1}{2\pi e\alpha^2\sigma^2})$ is related to intensity λ as

$$\lambda = \exp(nR) = (\alpha\sigma\sqrt{2\pi e})^{-n}$$

  Here α is the GSNR , σ is noise deviation and n the dimension of space.

- The noise are modelled as displacement vectors (ditto like Polytrev case), one for each point of the point process. The displacement vectors are independent of the points, each displacement vector being Gaussian with iid coordinates having zero mean and variance σ^2 and with the displacement vectors being iid from point to point.

3.1 Error Exponents

The definitions are given for a general setting. Let D be some stationary ergodic additive noise. For all stationary and ergodic point processes μ^n of normalized logarithmic intensity $-h(D) - ln(\alpha)$ and all jointly stationary decoding regions $C^n = \{C^n_k\}_k$, let $p^{pp}_e(n, \mu^n, C^n, \alpha, D)$ be the probability of error. Then

$$\pi(\mu, C, \alpha, D) = \lim_{n\to\infty} -\frac{1}{n} ln(p^{pp}_e(n, \mu^n, C^n, \alpha, D))$$

The same limit by plugging in the optimal probability of error across all stationary and ergodic point processes with NLD $-h(D) - ln(\alpha)$and all decoding regions will be called the error exponent η

$$\eta(\alpha, D) = \lim_{n\to\infty} -\frac{1}{n} ln(p^{pp}_{e,opt}(n, \alpha, D))$$

Suppose we take a particular point process μ (say Poisson) and a decoding rule C (say MLE), the lower bound on the reliability function we calculate in this setting

$\underline{\pi}(\mu, C, \alpha, D)$ will by definition be lesser than the optimal lower bound $\underline{\eta}(\alpha, D)$

$$\underline{\eta}(\alpha, D) \geq \underline{\pi}(\mu, C, \alpha, D)$$

And similarly

$$\overline{\eta}(\alpha, D) \leq \overline{\pi}(\mu, C, \alpha, D)$$

Then the problem of finding the error exponent η is just estimating the best π. If we take different point processes μ^n and find $\underline{\pi}$ for each of them, the greatest of these estimates will be our lower bound on the reliability function $\underline{\eta}$. We'll be doing exactly that in the following two sections taking Poisson and Matern point processes respectively.

3.2 Theorem

(Theorem 3-wgn in [5]) For all stationary isotropic and ergodic point processes $\mu^n$and all iid white gaussian displacement vectors $r\vec{v}$, the probability of error under MLE is

$$p_e(n) = \int_{r \geq 0} (1 - P_0^n(\mu^n(B^n(r\vec{v}, r)) = 0)) g_\sigma^n(r) dr$$

where P_0^n is the palm probability of μ^n,$B^n(r\vec{v}, r)$ is a ball of radius r around $r\vec{v}$, and g_σ^r is the gaussian noise term.

3.3 Lower bound for Error Exponent

The argument is same as Random Coding given by Shannon. The random codes are in this case conveniently taken as point process. Then the problem of finding error exponents from the probability of error is simply a minimization problem as shown below.

$$p_e(n) = \int_{r \geq 0} (1 - P_0^n(\mu^n(B^n(r\vec{v}, r)) = 0)) g_\sigma^n(r) dr$$

$$= \int_{v \geq 0} (1 - P_0^n(\mu^n(B^n(v\sigma\sqrt{n})) = 0)) g_1^n(v\sqrt{n}) \sqrt{n} dv$$

where B^n is a sphere of radius $v\sigma\sqrt{n}$ touching the Origin. Now if you are able to represent $(1 - P_0^n(\mu^n(B^n(v\sigma\sqrt{n})) = 0))$ (the probability of finding a point in a sphere of radius $v\sigma\sqrt{n}$ touching Origin) as $exp(-nb(v))$, that is

$$(1 - P_0^n(\mu^n(B^n(v\sigma\sqrt{n})) = 0)) \leq exp(-nb(v))$$

then the lower bound for error exponents are simply the result of the optimization problem- For each $\alpha > 1$

$$Minimize\, a(v) + b(v)\, over\, v \geq 0$$

3.3.1 A general result

A result true for any point process is the following

$$1 - P_0^n(\mu^n(B^n(r\vec{v}, r)) = 0) \leq min(1, \text{Expected no of pts in } B)$$

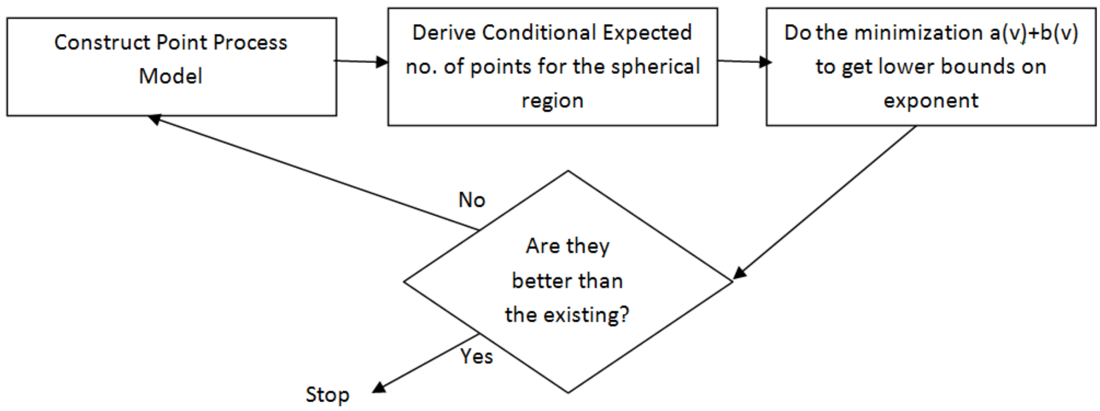
Note that the $Expected\, no\, of\, pts\, in\, B$ is a conditional expectation on the point at Origin. So for stationary point process finding $b(v)$ is equivalent to doing the integral

$$\text{Conditional Expected number of points in } B$$

$$= \int_B \frac{\rho_2(x, o)}{\rho_1(o)} dx$$

where B is a sphere of radius $v\sigma\sqrt{n}$ touching the Origin.

3.3.2 Palm Approach for Lower Bound



3.4 Example: Poisson Point Process

3.4.1 Result

$$\underline{\pi}(\mu, C, \alpha, D) = \liminf_{n\to\infty} -\frac{1}{n} ln(p_e^{pp}(n, \mu^n, C^n, \alpha, D))$$

Let μ be a sequence of Poisson processes μ^n of rates $\lambda = e^{nR}$ where $R = \frac{1}{2} ln(\frac{1}{2\pi e \alpha^2 \sigma^2})$ for $\alpha > 1$. Decoding rule C is Maximum Likelihood Estimation denoted by $L(WGN)$. Noise is AWGN. Then

$$\underline{\pi}(Poi, L(WGN), \alpha, WGN) = \begin{cases} \frac{\alpha^2}{2} - \frac{1}{2} - ln(\alpha) & if\ 1 \leq \alpha < \sqrt{2} \\ \frac{1}{2} - ln(2) + ln(\alpha) & if\ \sqrt{2} \leq \alpha < \infty \end{cases}$$

3.4.2 Method

Conditional Expected number of points in B

$$= \int_B \frac{\rho_2(x, o)}{\rho_1(o)} dx$$

using Slivnyak's theorem [7]

$$\int_B \frac{\lambda^2}{\lambda} dx$$

$$\lambda V_B^n(v\sigma\sqrt{n})$$

where $V_B^n(r)$ is the volume of n dimensional ball of radius r.

$$\leq C(\frac{v}{\alpha})^n$$

Therefore

$$b(v) = max(0, \log(\alpha) - \log(v))$$

which will be denoted as

$$b(v) = (\log(\alpha) - \log(v))^+$$

Now we have the integrand in the e^{-nx} form required. The exponent of the integrand at v is of the form $a(v) + b(v)$ where

$$a(v) = \frac{v^2}{2} - \frac{1}{2} - ln(v)$$

pertains to the additive noise and

$$b(v) = (ln(\alpha) - ln(v))^+$$

to the ball having some point. Carrying out the optimization

$$Minimize\ a(v) + b(v)\ over\, v \geq 0$$

gives $\frac{\alpha^2}{2} - \frac{1}{2} - ln\alpha$ for $1 < \alpha < \sqrt{2}$and $\frac{1}{2} - ln2 + ln\alpha$ when $\alpha > \sqrt{2}$.

3.5 Example: Matern Point Process

3.5.1 Result

We consider the ground PPP to have rate $\lambda = e^{nR}$ where $R = \frac{1}{2}ln(\frac{1}{2\pi e\alpha^2\sigma^2})$ for $\alpha > 1$. The exclusion radius is $(\alpha - \epsilon)\sigma\sqrt{n}$. The intensity of this process is always lesser than the ground process and tends to λ as $n \to \infty$. So the infimum we calculate for the error probability of this point process has to be lesser than the ground process

$$\underline{\pi}(Mat, L(WGN), \alpha, WGN) \geq \frac{\alpha^2}{8}\ \ for\, all\, \alpha \geq 2$$

Together with Poisson, this gives the Polytrev bounds.

3.5.2 Method

Again consider Conditional Expected number of points in $B(v\sigma\sqrt{n})$

When $v < \frac{\alpha}{2}$ there can't be any point in the ball and hence $b(v) = \infty$

When $\frac{\alpha}{2} < v < \frac{\alpha}{\sqrt{2}}$ the figure [5] below illustrates how to find $b(v)$

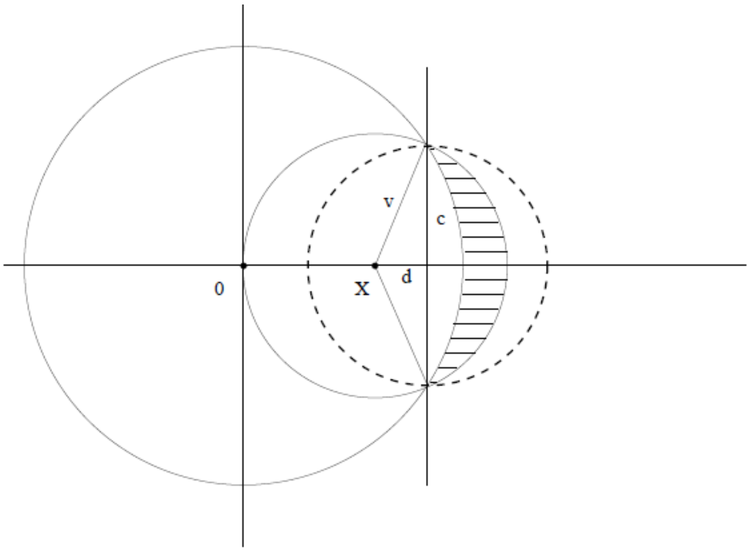


Figure : The origin of the plane is the tagged codeword. The large ball centered on the origin is the exclusion ball of the Matérn construction around the tagged codeword. Its radius is $(\alpha - \epsilon)\sigma\sqrt{n}$. The point X is the location of the noise added to the tagged codeword. Its norm is $v\sigma\sqrt{n}$. The ball centered on X with radius $v\sigma\sqrt{n}$ is the vulnerability region in the Poisson case. In the Matérn case, the vulnerability region is the shaded lune depicted on the figure. We are here in the case with $\frac{\alpha}{2} < v < \frac{\alpha}{\sqrt{2}}$. We upper bound the area of this lune by that of the ball of radius $c = \sqrt{n(v^2 - d^2)}\sigma$ with d as above. This ball is depicted by the dashed line disc.

When $v > \frac{\alpha}{\sqrt{2}}$ probability of error can be upper bounded by that of Poisson point process.

The minimization then looks like

$$minimize\ b(v) + a(v)$$

$$a(v) = \frac{v^2}{2} - \frac{1}{2} - ln(v)$$

$$b(v) = \begin{cases} \infty & if\, 0 < v < \frac{\alpha}{2} \\ ln\alpha\, - \frac{1}{2}ln(v^2 - (v - \frac{\alpha^2}{2v})^2) & if\, \frac{\alpha}{2} < v < \frac{\alpha}{\sqrt{2}} \\ (ln\alpha - lnv)^+ & if\, \frac{\alpha}{\sqrt{2}} < v \end{cases}$$

Considering only the case of $\alpha \geq 2$ we get $\frac{\alpha^2}{8}$ as the exponent.

Part II

MY WORK

CHAPTER 4

DETERMINANTAL POINT PROCESS

Determinnatal point process [8] arise in study of random matrices and fermions in quantum physics. They are natural choices for modelling repulsion between points. Incorrect decoding will be reduced when the code points have repulsion between them. So we decide to try them on our problem. The lower bound on error exponents got from random codes modelled as determinantal point process should be atleast equal to existing ones.

4.1 Introduction

4.1.1 Definition

A point process χ on $\mathbb{R}^n$ is said to be a determinantal point process with kernel K if it is simple and the joint intensities satisfy

$$\rho_k(x_1, \ldots x_k) = determinant \left[(K(x_i, x_j) \right]_{1 \leq i,j \leq k}$$

for every $k \geq 1$ and $x_1, \ldots x_k \in \mathbb{R}^n$

4.1.2 Examples

The most popular stationary model is

$$K(x, y) = \rho \exp \left[-\frac{||x||^2}{a^2} \right]$$

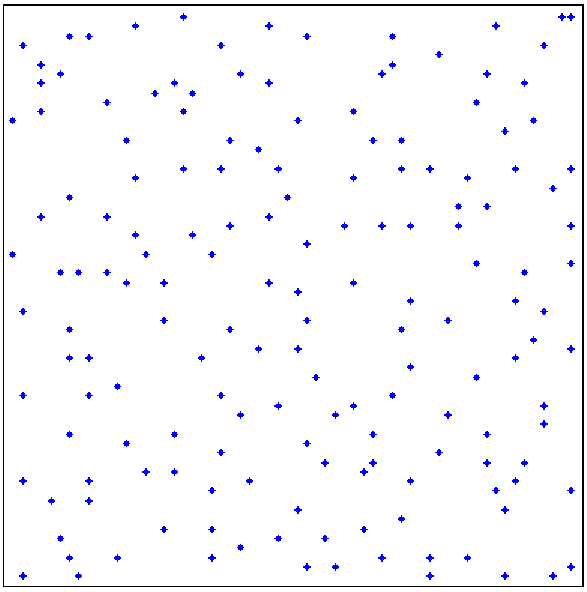
called the Gaussian kernel. An example of a non stationary determinantal point process is the following

$$K(z, w) = \frac{1}{\pi} \exp \left[-\frac{1}{2}(|z|^2 + |w|^2) + z\bar{w} \right]$$

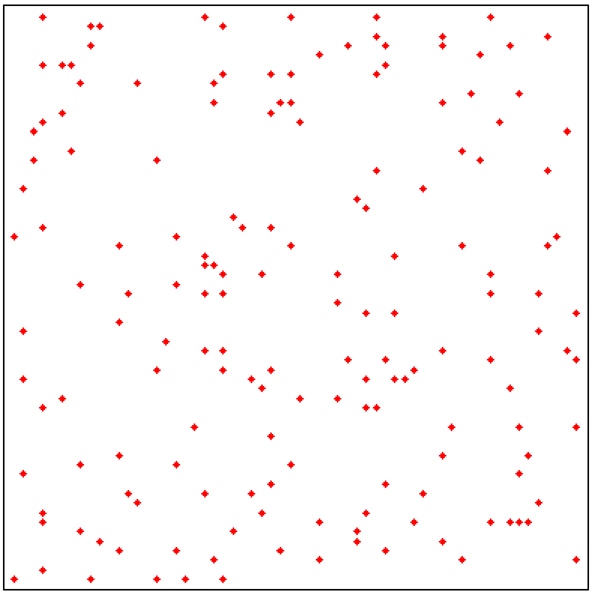
called Ginibre ensemble

4.1.3 Plot

The following is a two dimensional simulation of a determinantal point process with Gaussian kernel (algorithm in Appendix B). Below it is a sampling of same number of points as Poisson.



DPP



Independent

The repulsion between DPP points is clearly visible in the plots!

4.2 Conditions on Kernel

A kernel function $K(x, y)$ is valid for a stationary determinantal point process if the following are satisfied [9]

1. There exists $\kappa \in L^2(\mathbb{R}^n)$
$$K(x, y) = \kappa(x - y).$$
So $\kappa(x) = K(x, 0)$

2. $K(x, y)$ is continuous and positive definite. Continuity is straight forward to determine. Positive definiteness not so. Bochner's theorem says that a continuous positive definite function on a locally compact group corresponds to a finite positive measure on the dual group. In simple terms for our case this means the radial fourier transform of $K(x, y) = \kappa(r)$ needs to be non negative. The radial Fourier transform is given by
$$\mathcal{F}(\kappa)(s) = \int_0^\infty \int_0^\pi \exp(-jsr\cos(\theta))\kappa(r)V_{n-2}\sin(\theta)^{n-2}d\theta r^{n-1}dr$$
where V_n is the volume of n dimensional unit ball. Using the definition of Bessel J function
$$J_{\frac{n-2}{2}}(t) = \frac{t^{\frac{n-2}{2}}}{(2\pi)^{n/2}}V_{n-2}\int_0^\pi \exp(-jt\cos(\theta))\sin(\theta)^{n-2}d\theta$$
we have
$$\mathcal{F}(\kappa)(s) = \int_0^\infty \frac{(2\pi)^{n/2}V_{n-2}}{sr^{\frac{n-2}{2}}}J_{\frac{n-2}{2}}(sr)\kappa(r)r^{n-1}dr$$
This means the condition to be met for positive (semi)definiteness of K is
$$\int_0^\infty \frac{1}{sr^{\frac{n-2}{2}}}J_{\frac{n-2}{2}}(sr)K(r, 0)r^{n-1}dr \geq 0, \; \forall s > 0$$

3. The fourier transform of κ satisfies
$$\mathcal{F}(\kappa)(s) \leq 1.$$
Ofcourse you could do it by the steps used in the previous section by finding the exact radial fourier transform. We however used an easier method to get a sufficient condition.
$$\mathcal{F}(\kappa)(s) = \int_{\mathbb{R}^n} K(z, 0)e^{(-2\pi izs)}d(z)$$
$$\leq \int_{\mathbb{R}^n} |K(z, 0)e^{(-2\pi izy)}d(z)|$$
$$\int_{\mathbb{R}^n} |K(z, 0)|dz$$
$$\leq 1.$$

Thus it is sufficient that

$$\int_{\mathbb{R}^n} |K(z,0)|dz \le 1$$

4.3 Palm Procedure using DPP

- Find a function $K(x,y)$, $\{x,y\} \in \mathbb{R}^n$ satisfying the three conditions above
- From Thm1.7 in [10], we know that palm measure for point at origin o is again a determinantal point process with kernel

$$K^o(x,y) = \frac{1}{K(o,o)} \left(K(x,y).K(o,o) - K(x,o).K(o,y) \right)$$

- Since intensity is simply $K(x,x)$, the expected number of points in a borel set B ,conditioned on Origin will be

$$\int_B (K_n^0(x,x))d\mu(x) \le \exp(-n\, b(v))$$

- After obtaining $b(v)$ minimize $a(v)+b(v)$ for every α to get lower bound on error exponents.

4.4 Finding a suitable kernel

4.4.1 The Kernel that Gave Better Bounds

Consider the kernel

$$K(z,w) = \lambda \exp\left[-\frac{\|z-w\|^n}{(\tau\alpha\sigma\sqrt{n})^n} \right]$$

where τ is a positive parameter we introduce and the rest of the terms are as defined in Chapter 3. For $\tau > 1$ this determinantal point process gives lower bound $\frac{\alpha^2}{2} - \frac{1}{2} - ln(\alpha)$ which is equal to sphere packing upper bound. However, such a kernel is not valid as it's fourier transform satisfies Condition 3 only for $\tau \le 1$.

4.4.2 The Kernel that Matched the Existing Bounds

$$K(z,w) = \lambda \exp\left[-\frac{1}{(\alpha\sigma\sqrt{n})^{2kn}} (\|z-w\|^{2kn}) \right]$$

for $k \geq 0$. This kernel gives error exponents matching the existing lower bounds for $k = \frac{1}{2}(\frac{\alpha^2}{4} - 1)$ (see Appendix D). But this isn't positive definite for $kn > 1$. So invalid again!

4.4.3 Cauchy Kernel

A valid kernel covariance function for a stationary DPP is

$$C_o(x) = \frac{\rho}{(1 + (\frac{\|x\|}{\alpha})^2)^{v+n/2}}, \ x \in \mathbb{R}^n$$

and the condtion for fourier transform being less than unity is (as given in [9])

$$\rho \leq \frac{\Gamma[v + n/2]}{\Gamma[v](\sqrt{\pi}\alpha)^n}$$

Plug in $\rho \leftarrow \lambda = (\alpha\sigma\sqrt{2\pi e})^{-n}$, $\alpha \leftarrow \alpha\sigma\sqrt{n}$, $v \leftarrow \tau$ and we can prove that for τ large enough

$$(\alpha\sigma\sqrt{2\pi e})^{-n} \leq \frac{\Gamma[\tau + n/2]}{\Gamma[\tau](\sqrt{\pi}\alpha\sigma\sqrt{n})^n}$$

So the DPP in our hand is

$$K(x, y) = \frac{\lambda}{(1 + (\frac{\|x-y\|}{\alpha\sigma\sqrt{n}})^2)^{\tau+n/2}}$$

Though this kernel satisfies all three conditions and is of the same form, it didn't give good lower bounds. In fact the lower bounds go to zero for large n.

4.5 Conclusion

After trying to tailor the right kernel we realize there may not be a determinantal point process that suits us. In any case, it is a bad idea to restrict to dpps alone when what we want is any stationary point process with favourable first and second moment densities. This leads to the point process of next chapter.

CHAPTER 5

A NEW POINT PROCESS

A result on the sufficient condition for the existence of certain point processes is given in [11]. Using the existence theorem for this point process, we'll construct valid models that give good lower bounds.

5.1 Existence Condition

A symmetrical function $h(x,y)$ defined on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying

1. $0 \leq h(x,y) \leq 1$
2. $c = \sup_x \int_{\mathbb{R}^n} (1 - h(x,y))dy < \infty$

will describe a point process such that the intensity meansure ρ are

$$\begin{aligned} \rho(x_1) &= \lambda \\ \rho(x_1, x_2, \ldots x_m) &= \lambda^m \prod_{\{i,j\} \subset \{1,2,\ldots k\}} h(x_i, x_j) \end{aligned}$$

if

$$0 \leq \lambda < \frac{1}{ce}.$$

5.2 Choosing h

Take

$$h(x,y) = 1 - \exp\left[-\frac{|x-y|^\beta}{(\alpha\sigma\sqrt{n})^\beta} \right]$$

where $\beta = (\frac{\alpha^2}{4} - 1)n$ and the rest of the terms are as defined in Chapter 3. The condition $0 \leq h(x,y) \leq 1$ is obviously satisfied. The second condition

$$c = \sup_x \int_{\mathbb{R}^n} (1 - h(x,y))dy$$

$$= \int_{\mathbb{R}^n} (1 - h(0,y))dy$$

$$= \int_{\mathbb{R}^n} \exp\left[-\frac{|x-y|^\beta}{(\alpha\sigma\sqrt{n})^\beta}\right] dy$$

$$= n\frac{\pi^{n/2}}{\Gamma(1+(n/2))}\frac{(\alpha\sigma\sqrt{n})^n}{\beta}\Gamma(\frac{n}{\beta})$$

$$< \infty \ \forall n, \beta$$

is also satisfied. By 5.1 ,we are now guaranteed a point process of this form as long as intensity λ satisfies

$$0 \leq \lambda < \frac{1}{ce}.$$

It can be shown that $\lambda = (\alpha\sigma\sqrt{2\pi e})^{-n}$ satisfies the above constraint (see Appendix C). So in particular we've proved there exists a point process with first and second moment densities

$$\rho(x_1) = \lambda$$

$$\rho(x_1, x_2) = \lambda^2 h(x_1, x_2) = \lambda^2 \left[1 - \exp\left[-\frac{|x-y|^\beta}{(\alpha\sigma\sqrt{n})^\beta}\right]\right]$$

5.3 Error Exponents

Now we use the palm approach to get error exponents. The Existence Probability

$$1 - P_0^n(\mu^n(B^n(r\vec{v}, r)) = 0) \leq min(1, \text{Expected no of pts in}\, B)$$

And for a ball B of radius $v\sigma\sqrt{n}$ centered at $v\sigma\sqrt{n}\hat{x_1}$

$$\text{Palm Expected no of pts in}\, B = \int_B \lambda \left[1 - \exp\left[-\frac{|x|^\beta}{(\alpha\sigma\sqrt{n})^\beta}\right]\right] dx.$$

It can be shown that (Appendix D) this integral is

$$\leq \exp(-n\, b(v))$$

where

$$b(v) = (2k_\alpha + 1)\left\{\log(\alpha) - \log(v) + \log(\frac{(k_\alpha+1)^{\frac{k_\alpha+1}{2k_\alpha+1}}}{(2k_\alpha+1)^{\frac{1}{2}}.2^{\frac{k_\alpha}{2k_\alpha+1}}})\right\}$$

and $k_\alpha = \frac{1}{2}(\frac{\alpha^2}{4} - 1)$. Hence we have

$$1 - P_0^n(\mu^n(B^n(r\vec{v}, r)) = 0) \leq \exp[-n \max(b(v), 0)]$$

With $a(v) = \frac{v^2}{2} - \frac{1}{2} - \log(v)$, we have the minimization problem to obtain exponents

$$Minimize\, a(v) + (b(v))^+ \, over\, v \geq 0$$

The expression $a(v) + b(v)$ achieves minimum at $v = \sqrt{2k_\alpha + 2}$. So, under the condition that $\alpha > \frac{(2k_\alpha+1)^{\frac{1}{2}}.2^{\frac{k_\alpha}{2k_\alpha+1}}}{(k_\alpha+1)^{\frac{k_\alpha+1}{2k_\alpha+1}}}\sqrt{2k_\alpha + 2}$, we get the minimum to be $a(\sqrt{2k_\alpha + 2}) + b(\sqrt{2k_\alpha + 2})$ which is

$$(2k_\alpha+1)\log(\alpha) + k_\alpha + \frac{1}{2} - (k_\alpha+1)\log(2k_\alpha+2) + (2k_\alpha+1)\log(\frac{(k_\alpha+1)^{\frac{k_\alpha+1}{2k_\alpha+1}}}{(2k_\alpha+1)^{\frac{1}{2}}.2^{\frac{k_\alpha}{2k_\alpha+1}}})$$

$k_\alpha = \frac{1}{2}(\frac{\alpha^2}{4} - 1)$ possible when $\alpha > 2$. Then the expression above simplifies to

$$\frac{\alpha^2}{8}$$

true when $\alpha > 2$. This matches the best lower bound.

5.4 Conclusion

We have constructed a new stationary point process that gives good lower bounds in the Polytrev regime. This is the only continuous intensity point process model to give the best known bounds. The advantage of this technique is it offers a method to improve the lower bound by constructing better point process in terms of the ρ_1 and ρ_2 functions.

5.5 Research Suggestions

- Improving the lower bound will just be about finding the best possible ρ_1 and ρ_2 satisfying stationarity. Conjecture 5.4 in [13] regarding the existence of stationary point processes, if true would be more general than 5.1 . Using that, it might be possible to construct the optimal random codes.
- I concentrated majority of my research on lower bound because it looked easier to work with than the upper bound over the short time frame. Improving the upper bounds- both the sphere packing bound and the minimum distance bound would require work on the 100 year old sphere packing problem (Appendix E).
- The opinions on the various papers I've read on which bound is weaker have been mixed. I am leaning towards "the upper bound that has to be brought down" side. But what people feel can be misleading.

APPENDIX A

POINT PROCESS THEORY

For detailed theory of point process refer [7]. The following is a quick recap.

A.1 Basics

Point process is extension of random variable concept to n dimensions. It is defined as a random collection of points in $\mathbb{R}^n$ with the Borel sets B^n as the associated σ-algebra, equipped with the Lebesgue measure. We'll adopt the random measure formalism to describe the point process distribution. In this, Point processϕ is described fully by $\phi(B)$,the number of points of ϕ falling in arbitrary sets $B \subset \mathbb{R}^n$. The intensity measure of a point process ϕ in a Borel set B is defined as

$$\Lambda(B) := E[\phi(B)]$$

The intensity λ is defined as the spatial function such that $\Lambda(B) = \int_B \lambda dx$

A.2 Examples

A.2.1 Poisson Point Process

This is the most popular of all the point processes. We'll be considering only uniform PPPs in our discussion.

Uniform PPP, with intensity λ is a point process in $\mathbb{R}^n$ such that

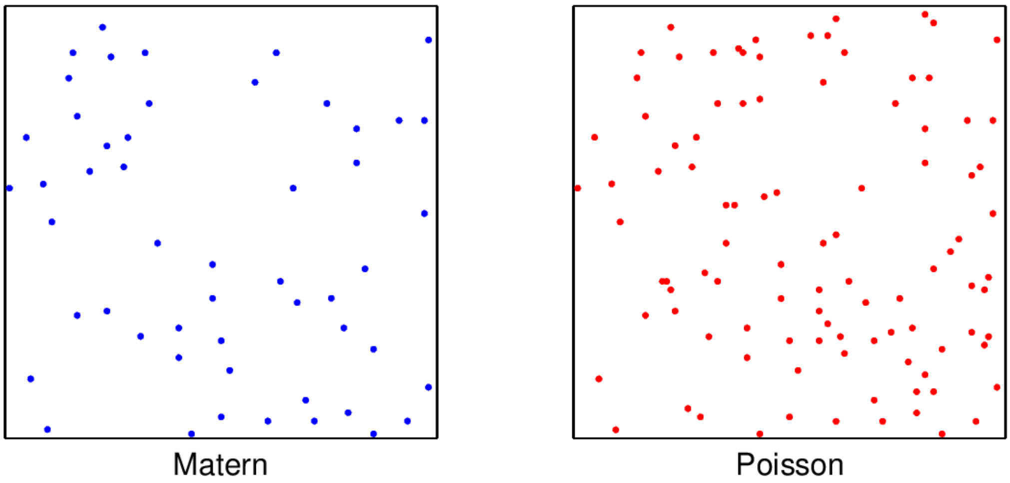
- for every bounded closed set B, $\phi(B)$ has a Poisson distribution with mean $\lambda \mid B \mid$.
- If $B_1, B_{2,} \ldots, B_m$ are disjoint bounded sets , then $\phi(B_1)$,$\phi(B_2)$,$\ldots \phi(B_m)$ are independent random variables

A.2.2 Matern Point Process

Matern Process ϕ (or Matern Hard-Core Process of type I) can be created by dropping points from a uniform PPP ϕ_b with intensity λ_b. Flag all points for removal that have a neighbour within distance r. Then remove all flagged points. Formally,

$$\phi = \{ \ x \in \phi_b \ : \ if \ \ | \ \phi_b \cap b(x,r) \setminus \{x\} \ |= 0 \ \}$$

where $b(x,r)$ is the n-dimensional ball centered at x of radius r. The figure below shows Matern Point Process with its parent Poisson point process.



A.3 Moment Densities

The moment densities (also called joint intensity functions) of a point process ϕ are functions $\rho_k : (\mathbb{R}^n)^k \to [0,\infty)$ for $k \geq 1$ such that for any family of mutually disjoint subsets $D_1, \ldots D_k$ of $\mathbb{R}^n$,

$$E\left[\prod_{i=1}^{k} \phi(D_i)\right] = \int_{\prod_i D_i} \rho_k(x_1, \ldots, x_k) dx_1 \ldots dx_k.$$

$\rho_k(x_1, \ldots x_k)$ are in a way a measure of the likelihood of $\{x_1, \ldots x_k\}$ being in a realization of point process ϕ. Further the intensity $\lambda(x)$ is simply $\rho_1(x)$.

A.4 Palm Distribution Intensity

Palm Distribution means conditional distribution. For stationary point process ϕ, condition on the occurence of point at Origin o. Then the resulting probability distribution $P(Y \,|o \in \phi)$ is called the reduced Palm distribution and its intensity is (section 7.6 of [7])

$$\frac{\rho_2(x, o)}{\rho_1(o)}$$

APPENDIX B

Simulating DPP

B.1 Determinantal Projection Process

A determinantal point process is a mixture of determinatal projection processes. Suppose χ is a determinantal point process with kernel

$$K(x,y) = \Sigma_{k=1}^{n} \lambda \phi_k(x) \bar{\phi_k}(y)$$

where ϕ_k are normalized eigenfunctions of K with $\lambda_k \in [0,1]$ the eigenvalues. Let I_k, $1 \leq k \leq n$ be independent random variables with $I_k \sim Bernoulli(\lambda_k)$. Define K_I the random analogue of kernel K as

$$K_I(x,y) = \Sigma_{k=1}^{n} I_k \phi_k(x) \bar{\phi_k}(y)$$

and the point process with this kernel as χ_I. Then

$$\chi = \chi_I$$

An important property of determinantal projection processes is the following:
A determinintal projection process with kernel $K(x,y) = \Sigma_{k=1}^{n} \phi_k(x) \bar{\phi_k}(y)$ where $\{\phi_k\}$ is an orthonormal set, has almost surely n number of points.

B.2 The Algorithm

(from [9])

1. The first step is to find the eigen functions ϕ and eigenvalues λ of the matrix of all possible values of $K(x,y)$. (Eigen in Matlab)
2. Sample the iid Bernoulli random variables $Bernoulli(\lambda_k)$ to get the determinantal projection process $K_I(x,y) = \Sigma_{k=1}^{N} \phi_k(x) \bar{\phi_k}(y)$ from $K(x,y)$

3. Define $v(x) = (\phi_1(x), \phi_2(x), \dots \phi_N(x))\backslash$

4. Sample point X_N from the distribution $p_N(x) = \frac{|v(x)|^2}{N}$

5. Set $e_1 = \frac{v(X_N)}{|v(X_N)|}$

6. For $i = N - 1$ to 1do
   Sample X_i from $p_i(x) = \frac{1}{i}\left[|v(x)|^2 - \sum_{j=1}^{N-i} |e_j^* v(x)|^2\right]$
   set $w_i = v(X_i) - \sum_{j=1}^{N-i}(e_j^* v(x))e_j$ and $e_{N-i+1} = \frac{w_i}{|w_i|}$
   end for
   return $X_1, ..., X_N$

APPENDIX C

Proving Intensity Limit

$\lambda = (\alpha\sigma\sqrt{2\pi e})^{-n}$ is non negative, so what we need to check is

$$\text{Is}\lambda < \frac{1}{ce}?$$

$$ce < \frac{1}{\lambda}$$

$$n\frac{\pi^{n/2}}{\Gamma(1+(n/2))}\frac{(\alpha\sigma\sqrt{n})^n}{\beta}\Gamma(\frac{n}{\beta})e < (\alpha\sigma\sqrt{2\pi e})^n$$

$$\frac{n}{\Gamma(1+(n/2))}\frac{(\sqrt{n})^n}{\beta}\Gamma(\frac{n}{\beta})e < (\sqrt{2e})^n$$

We've seen this form before and can realize that it will go through only for $\beta = o(n)$. Lets put $\beta = n$

$$\frac{n}{\Gamma(1+(n/2))}\frac{(\sqrt{n})^n}{n}\Gamma(\frac{n}{n})e < (\sqrt{2e})^n$$

$$\frac{(\sqrt{n})^n}{\Gamma(1+(n/2))}e < (\sqrt{2e})^n$$

Using Stirling's Approximation for Gamma for large n

$$\frac{(\sqrt{n})^n}{\sqrt{\frac{2\pi}{(1+(n/2))}}(\frac{1+(n/2)}{e})^{(1+(n/2))}}e < (\sqrt{2e})^n$$

Use the limit $\lim_{n\to\infty}(1+\frac{2}{n})^{-n/2} = e^{-1}$

$$\frac{e}{\sqrt{2\pi}\sqrt{(1+(n/2))}} < 1$$

which is true for suffieciently large n

APPENDIX D

Deriving Error Exponents

Consider the integral over a ball B of radius $v\sigma\sqrt{n}$ centered at $v\sigma\sqrt{n}\hat{x}_1$

$$\int_B \lambda(1-\exp(-\frac{\|x\|^{2k_\alpha n}}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}}))\,d\mu(x)$$

$$\leq \int_B \lambda\frac{\|x\|^{2k_\alpha n}}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}}\,d\mu(x)$$

Splitting x into the cartesian coordinates $(x_1, x_2, x_{3,}, ...x_n)$, we get the integral as

$$\int_{x_1=0}^{2v} \frac{\lambda}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \int_{x_2^2+x_3^2+...x_n^2<v^2-(x_1-v)^2} (x_1^2+x_2^2+...x_n^2)^{k_\alpha n} dx_2...dx_n dx_1$$

which using binomial expansion becomes

$$\frac{\lambda}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \sum_{i=0}^{k_\alpha n} \binom{k_\alpha n}{i} \int_{x_1=0}^{2v} x_1^{2(nk_\alpha - i)} \left\{ \int_{x_2^2+x_3^2+...x_n^2<v^2-(x_1-v)^2} (x_2^2+...x_n^2)^i dx_2...dx_n \right\} dx_1$$

$$= \frac{\lambda}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \sum_{i=0}^{k_\alpha n} \binom{nk_\alpha}{i} \int_{x_1=0}^{2v} x_1^{2(nk_\alpha - i)} \left\{ V_B^{n-1}(1) \int_{r=0}^{\sqrt{v^2-(x_1-v)^2}} r^{2k_\alpha} r^{n-2} dr \right\} dx_1$$

$$= \frac{\lambda V_B^{n-1}(1)}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \sum_{i=0}^{k_\alpha n} \binom{k_\alpha n}{i} \int_{x_1=0}^{2v} x_1^{2(nk_\alpha - i)} (2vx_1 - x_1^2)^{\frac{2i+nk_\alpha-1}{2}} dx_1$$

$$= \frac{\lambda V_B^{n-1}(1)}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \sum_{i=0}^{k_\alpha n} \binom{k_\alpha n}{i} 2^{2nk_\alpha+n-1} v^{2nk_\alpha+n} \frac{\Gamma(i+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n - i + \frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$$

Let i_{max} be i value that maximizes $\binom{k_\alpha n}{i}\frac{\Gamma(i+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n-i+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$. The value of i_{max} can be found by differentiating wrt to k_α and solving for root.

$$\frac{\partial}{\partial k_\alpha}\left[\binom{k_\alpha n}{i}\frac{\Gamma(i+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n - i+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}\right] = 0$$

Doing the differentiation, we get in the numerator (where PG is the $PolyGamma$ function)

$$PG(0, 1+k_\alpha n - i) + PG(0, \frac{1}{2}+\frac{n}{2}+i) - PG(0, 1+i) - PG(0, \frac{1}{2}(1+4nk_\alpha+n-2i))$$

Using $PG(1+x) \approx \log(1+x)$ we have

$$(1+k_\alpha n - i_{max})(\frac{1}{2}+\frac{n}{2}+i_{max}) = (1+i_{max})\frac{1}{2}(1+4nk_\alpha+n-2i_{max})$$

$$i_{max} = \frac{nk_\alpha(n-3)}{2(n+nk_\alpha-1)}$$

which for large n is

$$i_{max} = \frac{k_\alpha n}{2(k_\alpha+1)}$$

The correspoding maximum value is

$$\binom{k_\alpha n}{i_{max}} \frac{\Gamma(i_{max}+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n - i_{max}+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$$

Using $\binom{x}{y} = \frac{\Gamma(1+x)}{\Gamma(1+y)\Gamma(x-y+1)}$

$$\frac{\Gamma(1+k_\alpha n)}{\Gamma(1+\frac{k_\alpha n}{2(k_\alpha+1)})\Gamma(k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)}+1)} \frac{\Gamma(\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$$

We are interested only in the upper bound for the exponential terms. Taking $\Gamma(x) \leq c(\frac{x}{e})^x$and for large n

$$\leq \frac{(k_\alpha n)^{k_\alpha n}(\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2})^{\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2}}(2k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2})^{2k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)}+\frac{n}{2}}}{(\frac{k_\alpha n}{2(k_\alpha+1)})^{\frac{k_\alpha n}{2(k_\alpha+1)}}(k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)})^{k_\alpha n-\frac{k_\alpha n}{2(k_\alpha+1)}}(2k_\alpha n+n)^{2k_\alpha n+n}}$$

$$\leq \frac{(2k_\alpha+1)^{\frac{2k_\alpha+1}{2}n}}{2^{(k_\alpha+1)n}(k_\alpha+1)^{(k_\alpha+1)n}}$$

Getting back to our original expression

$$\frac{\lambda V_B^{n-1}(1)}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}} \sum_{i=0}^{k_\alpha n} \binom{k_\alpha n}{i} 2^{2nk_\alpha+n-1} v^{2nk_\alpha+n} \frac{\Gamma(i+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n-i+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$$

$$\leq \frac{\lambda V_B^{n-1}(1)}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}}.k_\alpha n.\binom{k_\alpha n}{i_{max}} 2^{2nk_\alpha+n-1} v^{2nk_\alpha+n} \frac{\Gamma(i_{max}+\frac{n}{2}+\frac{1}{2})\Gamma(2k_\alpha n-i_{max}+\frac{n}{2}+\frac{1}{2})}{\Gamma(1+2k_\alpha n+n)}$$

which simplifies to

$$\leq \frac{\lambda V_B^{n-1}(1)}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}}.k_\alpha n.\frac{(2v)^{2k_\alpha n+n}(2k_\alpha+1)^{\frac{2k_\alpha+1}{2}n}}{2^{(k_\alpha+1)n}(k_\alpha+1)^{(k_\alpha+1)n}}$$

now using expression for volume of unit ball

$$\frac{c(\alpha\sigma\sqrt{2\pi e})^{-n}(\frac{2\pi e}{n-1})^{\frac{n-1}{2}}}{(\alpha\sigma\sqrt{n})^{2k_\alpha n}}.k_\alpha n.\frac{(2v)^{2k_\alpha n+n}(2k_\alpha+1)^{\frac{2k_\alpha+1}{2}n}}{2^{(k_\alpha+1)n}(k_\alpha+1)^{(k_\alpha+1)n}}$$

Collecting only the $n-$exponetial terms

$$=\left\{\frac{(2k_\alpha+1)^{\frac{1}{2}}.2^{\frac{k_\alpha}{2k_\alpha+1}}}{(k_\alpha+1)^{\frac{k_\alpha+1}{2k_\alpha+1}}}.\frac{v}{\alpha}\right\}^{(2k_\alpha+1)n}$$

This means

$$b(v)=(2k_\alpha+1)\left\{log(\alpha)-log(v)+log(\frac{(k_\alpha+1)^{\frac{k_\alpha+1}{2k_\alpha+1}}}{(2k_\alpha+1)^{\frac{1}{2}}.2^{\frac{k_\alpha}{2k_\alpha+1}}})\right\}$$

APPENDIX E

Sphere Packing Problem

The following table from [13] gives the best known density of sphere packing bounds in d dimensional space

$(2)2^{-d}$	Minkowski (1905)
$[\ln(\sqrt{2})d]2^{-d}$	Davenport and Rogers (1947)
$(2d)2^{-d}$	Ball (1992)

TABLE 1. Dominant asymptotic behavior of lower bounds on $\phi_{\max}^{L}$ for large d.

$(d/2)2^{-0.5d}$	Blichfeldt (1929)
$(d/e)2^{-0.5d}$	Rogers (1958)
$2^{-0.5990d}$	Kabatiansky and Levenshtein (1978)

TABLE 2. Dominant asymptotic behavior of upper bounds on $\phi_{\max}$ for large d.

[3] is the best reference for sphere packing and contains the best known upper bounds of today. [12] provides new upper bounds for certain dimension spaces. [13] shows how to use point process for improving lower bounds of sphere packing.

REFERENCES

[1] C. E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel". *Bell System Technical Journal*, Vol. 38, No. 3, pp. 611-656, May 1959.

[2] A. E. Ashikhmin, A. Barg, and S. N. Listyn, "A New Upper Bound on the Reliability Function of the Gaussian Channel". *IEEE Trans. on Inform. Theory*, Vol. 46, No. 6, pp. 1945-1961, May 2000.

[3] G. A. Kabatiansky and V. I. Levenshtein, "Bounds for Packings on a Sphere and in Space". *Problems of Information Transmission*, 14, pp. 1-17, 1978.

[4] G. Polytrev, "On Coding Without Restrictions for the AWGN Channel". *IEEE Trans. on Inform. Theory*, Vol. 40, No. 2, pp. 409-417, Mar. 1994.

[5] V. Anantharaman and F. Baccelli, "Information-Theoretic Capacity and Error Exponents of Stationary Point Processes under Random Additive Displacements", Dec. 2010.

[6] V. Anantharaman and F. Baccelli, "A Palm theory approach to error exponents". *Proceedings of the IEEE Symposium on Information Theory*, Toronto, July 2008.

[7] Martin Haenggi, "Stochastic Geometry for Wireless Networks".

[8] Hough, J. B., M. Krishnapur, Y. Peres and B. Virag, "Determinantal processes and independence". *Probability Surveys 3*, pp. 206-229 , 2006.

[9] F. Lavancier, J. Moller and E. Rubak, "Statistical aspects of determinantal point processes". *Centre pour la Communication Scientifique Directe*, pp. 1-46 , May 2012.

[10] T. Shirai and Y. Takahashi, "Random point fields associated with certain Fredholm determinants II: Fermion shifts and their ergodic and Gibbs properties". *The Annals of Probability*, Vol. 31, No. 3, pp. 1533-1564 , 2003.

[11] R. Ambartzumian and H. Sukiasian, "Inclusion-Exclusion and Point Processes". *Acta Applicandae Mathematicae*, 22, pp. 15-31 , 1991.

[12] H. Cohn and N. Elkies, "New Upper Bounds on Sphere Packings I ". *Annals of Mathematics*, 157, pp. 689-714 , 2003.

[13] S. Torquato and F. H. Stillinger, "New Conjectural Lower Bounds on the Optimal Density of Sphere Packings". *Experimental Mathematics*, Vol. 15, pp. 307-331 , 2006.