ISIT 2008, Toronto, Canada, July 6 - 11, 2008

A Palm theory approach to error exponents

Venkat Anantharam
EECS Department
University of California, Berkeley
Berkeley, CA 94720, U.S.A.
ananth@eecs.berkeley.edu

François Baccelli
École Normale Supérieure/INRIA
45 rue d'Ulm
75005, Paris, FRANCE
Francois.Baccelli@ens.fr

***Abstract*— We define a class of problems in the theory of Euclidean point processes, motivated by the study of the error exponent (reliability function) for additive noise channels. For the case of Gaussian noise this gives an interesting perspective on the Poltyrev exponent. It also suggests an approach to attack the long standing gap between the best known upper and lower bounds on the reliability function of the traditional AWGN channel, using techniques from point process theory.**

I. INTRODUCTION

In this paper we define a class of problems in the theory of Euclidean point processes, motivated by the study of the error exponent (reliability function) for additive noise channels. We are given a sequence of stationary ergodic marked point processes in Euclidean space $\mathbb{R}^n$, one for each dimension $n \geq 1$. The mark of a point is a pair comprised of an $\mathbb{R}^n$-valued vector and a measurable subset of $\mathbb{R}^n$. For each $n \geq 1$, the subsets associated to the various points of the process form a decomposition of $\mathbb{R}^n$. The canonical example to keep in mind is when the vectors associated to the individual points of the process are independent and identically distributed (i.i.d.) mean zero Gaussian random vectors each with i.i.d. coordinates and independent of the points, and the set associated to a point is its Voronoi cell in the realization of the process.

Each point of the process is thought of as being perturbed by its associated vector. We propose to study the asymptotics in dimension of the *error probability*, i.e. the Palm stationary probability that a point is perturbed outside its associated set. Thus the set associated to a point may be thought of as a *decoding region*. This problem formulation is directly motivated by the study of error exponents for additive noise channels in information theory.

Studying this point process formulation, we are able to recover and to give a new perspective on the Poltyrev exponent [10]. In particular, it appears explicitly as corresponding to a well-studied class of Matérn point processes. However, the main value of presenting this work, as we see it, is to suggest new ideas to attack the long standing gap between the best known upper and lower bounds on the reliability function of the traditional AWGN channel. The viewpoint we propose raises the possibility of bringing techniques from point process theory to bear on this problem. Some thoughts along these lines are sketched in section VIII.

The structure of this paper is as follows. In section II we recall the basic facts about the capacity and the reliability function of the AWGN channel. In section III we recall the corresponding facts for the AWGN channel without restrictions, and the connections between this channel and the usual AWGN channel. Our stationary point process viewpoint is developed in the case of additive Gaussian noise in sections V and VI; more general additive noise is discussed in section VII. Some thoughts regarding how one might develop this viewpoint are in section VIII.

$B_n(x, r)$ will denote the ball with center x and radius r in n-dimensional Euclidean space $\mathbb{R}^n$, with its volume denoted by $V_n(r) = \frac{r^n \pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$. $S^{n-1}(r)$ will denote the sphere of radius r in $\mathbb{R}^n$ centered at the origin. A_{n-1} will denote the area of $S^{n-1}(1)$: $A_{n-1} = nV_n(1) = \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$. $Cube_n(a)$ will denote the cube of side length a centered at the origin in $\mathbb{R}^n$. The cardinality of a finite set $\mathcal{S}$ will be denoted by $\mid \mathcal{S} \mid$.

II. THE AWGN CHANNEL

A. AWGN channel model

We recall the AWGN channel model in a rather abstract way. For the relation to the underlying physically relevant waveform channel model, see the text of Gallager [6]. Consider communication over the AWGN channel with noise variance σ^2 per degree of freedom using codewords of block length n subject to the power constraint P.

We write A^2 for the *signal-to-noise ratio* P/σ^2. A *code* is a finite subset $\mathcal{C}$ of points in $S^{n-1}(\sqrt{nP})$. Elements of the code are called *codewords*. We call $R(\mathcal{C}) = \frac{1}{n} \ln \mid \mathcal{C} \mid$ the *rate* of the code. Consider the Voronoi decomposition of $\mathbb{R}^n$ determined by the codewords. For each codeword, the probability that a Gaussian random vector in $\mathbb{R}^n$ centered at the codeword with independent coordinates each having variance σ^2 lands outside the corresponding Voronoi cell is called the associated probability of error. By the average probability of error of the code, denoted $P_e(\mathcal{C})$, we mean the average over the codewords of the corresponding probability of error. Given $\mathcal{C}$, this only depends on the signal-to-noise ratio A^2. One of the central problems of coding for the AWGN channel is that of understanding, for each fixed $A > 0$, the set of achievable pairs $(R(\mathcal{C}), P_e(\mathcal{C}))$, as one varies $\mathcal{C}$.

B. Shannon capacity of the AWGN channel

For fixed $A > 0$, given $R > 0$, we may ask whether there exists a sequence of codes $\mathcal{C}_n$ in $\mathbb{R}^n$, $n \geq 1$, having rate at least R, asymptotically as $n \to \infty$, and such that $P_e(\mathcal{C}_n) \to 0$

978-1-4244-2571-6/08/$25.00 ©2008 IEEE

as $n \to \infty$. As proved by Shannon, this is possible if $R < \frac{1}{2}\ln(1+A^2)$ and not possible if $R > \frac{1}{2}\ln(1+A^2)$. Thus $\frac{1}{2}\ln(1+A^2)$ may be called the Shannon capacity of the AWGN channel.

C. The reliability function of the AWGN channel

Let $A > 0$ and $0 < R \le \frac{1}{2}\ln(1+A^2)$. Let $P_{e,opt}(n,R,A)$ denote the infimum of $P_e(\mathcal{C})$ over all codes in $\mathbb{R}^n$ of rate at least R when the signal-to-noise ratio is A^2. We write

$$E(n,R,A) = -\frac{1}{n}\log P_{e,opt}(n,R,A) \ .$$

$$\text{Let} \qquad \begin{aligned} \bar{E}(R,A) &= \limsup_n E(n,R,A), \text{ and} \\ \underline{E}(R,A) &= \liminf_n E(n,R,A). \end{aligned}$$

Assuming these are identical, we denote this common limit by $E(R,A)$. For fixed $A > 0$ the function $R \mapsto E(R,A)$ is called the *reliability function* or the *error exponent* of the AWGN channel when the signal-to-noise ratio is A^2.

The best currently known upper and lower bounds for the reliability function are given in [1], to which we refer the reader. [1] We write $R \mapsto E_{SG}(R,A)$ for the best known lower bound, since it is due to Shannon [11] and Gallager [6]. We write $R \mapsto E_{ABL}(R,A)$ for the best known upper bound, since it is due to Ashikhmin, Barg and Litsyn [1]. These bounds coincide, and thus $E(R,A)$ is known, for

$$\frac{1}{2}\ln\left(\frac{1}{2}+\frac{A}{4}+\frac{1}{2}\sqrt{1+\frac{A^2}{4}}\right) \le R \le \frac{1}{2}\ln(1+A^2) \ ,$$

but there is a gap between these bounds for all lower rates.

III. THE AWGN CHANNEL WITHOUT RESTRICTIONS

Motivated by the observation that many proposed practical constructions of codes for the AWGN channel were derived from lattices in Euclidean space (which may be thought of as countable codes whose codewords are not subject to a power constraint) Poltyrev [10] formulated a problem of coding for an AWGN channel without restrictions (i.e. without a power constraint). Let $\mathcal{S}$ denote a countable collection of points in $\mathbb{R}^n$. The quantity

$$\rho(\mathcal{S}) = \frac{1}{n}\ln\left(\limsup_{a\to\infty}\frac{|\ \mathcal{S}\cap Cube_n(a)\ |}{a^n}\right)$$

is called the *normalized logarithmic density* of $\mathcal{S}$. Consider the Voronoi decomposition of $\mathbb{R}^n$ determined by $\mathcal{S}$. Given $\sigma > 0$, for each point $s \in \mathcal{S}$, the corresponding probability of error, $\lambda(s)$, is defined as the probability that a Gaussian random vector in $\mathbb{R}^n$ centered at s with independent coordinates each having variance σ^2 lands outside the Voronoi cell of s. The quantity

$$\lambda(\mathcal{S}) = \limsup_{a\to\infty}\frac{1}{|\ \mathcal{S}\cap Cube_n(a)\ |}\sum_{s\in\mathcal{S}\cap Cube_n(a)}\lambda(s)$$

is called the average probability of error of $\mathcal{S}$. The problem of coding for an AWGN channel without restrictions, in the sense of [10], is that of understanding the set of achievable pairs $(\rho(\mathcal{S}),\lambda(\mathcal{S}))$ over the choice of $\mathcal{S}$.

A. The Poltyrev capacity of the AWGN channel without restrictions

Fix $\sigma > 0$. Given $\rho > 0$, we may ask whether there exists a sequence of countable sets $\mathcal{S}_n$ in $\mathbb{R}^n$, $n \ge 1$, having normalized logarithmic density at least ρ, asymptotically as $n \to \infty$, and such that $\lambda(\mathcal{S}_n) \to 0$ as $n \to \infty$. As proved in [10], this is possible if $\rho < \frac{1}{2}\ln(\frac{1}{2\pi e\sigma^2})$ and not possible if $\rho > \frac{1}{2}\ln(\frac{1}{2\pi e\sigma^2})$. Thus $\frac{1}{2}\ln(\frac{1}{2\pi e\sigma^2})$ may be called the Poltyrev capacity of the AWGN channel without restrictions when the noise variance per coordinate is σ^2.

B. The reliability function of the AWGN channel without restrictions

Fix $\sigma > 0$ and $0 < \rho \le \frac{1}{2}\ln(\frac{1}{2\pi e\sigma^2})$. Let $\lambda_{opt}(n,\rho,\sigma^2)$ denote the infimum of $\lambda(\mathcal{S})$ over all countable sets $\mathcal{S}$ in $\mathbb{R}^n$ of normalized logarithmic density at least ρ when the noise variance per coordinate is σ^2. We write

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

$$\text{Let} \qquad \begin{aligned} \bar{\eta}(\rho,\sigma^2) &= \limsup_n \eta(n,\rho,\sigma^2), \text{ and} \\ \underline{\eta}(\rho,\sigma^2) &= \liminf_n \eta(n,\rho,\sigma^2). \end{aligned}$$

Assuming these are identical, we denote this common limit by $\eta(\rho,\sigma^2)$. The function $\rho \mapsto \eta(\rho,\sigma^2)$ is called the *reliability function* or the *error exponent* of the AWGN channel without restrictions when noise variance per coordinate is σ^2.

For any $\alpha > 0$, if we replace σ^2 by $\alpha^2\sigma^2$, we may scale any $\mathcal{S}$ to one whose normalized logarithmic density changes by addition of $\ln\alpha$ in order to get a geometrically identical situation from the point of view of error probability calculations. Thus $\eta(\rho,\sigma^2)$ depends only on $\frac{e^{-2\rho}}{\sigma^2}$. We let α^2 denote $\frac{e^{-2\rho}}{2\pi e\sigma^2}$ and, abusing terminology and notation, denote the function $\alpha \mapsto \eta(\frac{1}{2}\ln\frac{1}{2\pi e\alpha^2\sigma^2},\sigma^2)$, defined for $\alpha \ge 1$, by $\alpha \mapsto \eta(\alpha)$ and call this the *reliability function* or the *error exponent* of the AWGN channel without restrictions [2]

C. The Poltyrev exponent

Upper and lower bounds for the reliability function of the AWGN channel without restrictions were proved in [10]. The lower bound of [10] is normally called the *Poltyrev exponent*. This is the function $\alpha \mapsto \pi(\alpha)$ defined for $\alpha \ge 1$ by

$$\pi(\alpha) = \begin{cases} \frac{\alpha^2}{2} - \frac{1}{2} - \ln\alpha & \text{if } 1 \le \alpha < \sqrt{2} \\ \frac{1}{2} - \ln 2 + \ln\alpha & \text{if } \sqrt{2} \le \alpha < 2 \\ \frac{\alpha^2}{8} & \text{if } \alpha \ge 2 \ . \end{cases}$$

See eqns. (32) and (36) of [10]. It is proved in [10] that $\underline{\eta}(\alpha) \ge \pi(\alpha)$ for all $\alpha \ge 1$.

---

[1] Our ideas are more directly related to the Poltyrev exponent, so there is no direct benefit from reproducing the expressions of the AWGN reliability function bounds here. Note that the upper bounds are on $\bar{E}(R,A)$ and the lower bounds are on $\underline{E}(R,A)$.

[2] More precisely, we define $\alpha \mapsto \bar{\eta}(\alpha)$ and $\alpha \mapsto \underline{\eta}(\alpha)$ and assume they are identical.

The following upper bound for all $\alpha \geq 1$ is also proved in [10]:

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

The upper and lower bounds coincide for $1 \leq \alpha \leq \sqrt{2}$, and so $\eta(\alpha)$ is known in this range. For all $\alpha > \sqrt{2}$ there is a gap between the upper and lower bounds of [10].

The upper bound of [10] comes from a naive sphere packing bound, and can be immediately improved for large α by appealing to the path breaking work of Kabatyanskii and Levenshtein [7] on Euclidean sphere packing, as mentioned in the next subsection. In this connection, see also the papers of Cohn and Elkies [2], [3].

D. Connections with error exponents for the AWGN channel

Given $A > 0$ and $\alpha \geq 1$, consider $E(\frac{1}{2} \ln \frac{1+A^2}{\alpha^2}, A)$, where the function $R \mapsto E(R, A)$ is defined in subsection II-C. Then one can show, for all $\alpha \geq 1$, that

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

One also has, for all $\alpha \geq 1$:

$$\lim_{A \to \infty} E_{SG}(\frac{1}{2} \ln \frac{1+A^2}{\alpha^2}, A) = \pi(\alpha) \ .$$

If one takes

$$\lim_{A \to \infty} E_{ABL}(\frac{1}{2} \ln \frac{1+A^2}{\alpha^2}, A)$$

one gets an upper bound on $\alpha \mapsto \eta(\alpha)$ that improves the upper bound proved in [10]. Indeed, the upper bound of [1] is derived by using techniques inspired by [7]. One can beat the upper bound of [10] more crudely by directly appealing to results in [7]. Let Δ_n denote the *maximum packing density* of spheres in $\mathbb{R}^n$ (for the definition see [2]). From Cor. 2 of [7] one has the estimate $\Delta_n \leq (0.663)^n$, valid for large n. From this it is easy to conclude that $\eta(\alpha)$ is bounded above by $\frac{1}{2}(0.663)^2\alpha^2$.

The main point of this subsection is that improvements in the gap between the upper and lower bounds for the function $\alpha \mapsto \eta(\alpha)$ will lead to improvements in the gap between the upper and lower bounds for the reliability function of the AWGN channel, at least for large signal-to-noise ratios. This observation motivates our main theme, to which we turn next.

IV. STATIONARY POINT PROCESSES

For basic facts regarding stationary point processes see, for instance, the book of Daley and Vere-Jones [5].

Let μ^n be a sequence of stationary ergodic marked point processes. We assume that for all n, μ^n is a simple point process in $\mathbb{R}^n$. Let λ_n denote the intensity of the point process μ^n. We denote by $\{T_k^n\}$ the points of μ^n. The mark of the point T_k^n is a pair (X_k^n, C_k^n), where X_k^n is an vector in $\mathbb{R}^n$ and C_k^n is a measurable subset of $\mathbb{R}^n$. The countable collection of sets $\{C_k^n\}$ is assumed to form a decomposition of $\mathbb{R}^n$. We call X_k^n the *displacement vector* associated to the point T_k^n and C_k^n the *decoding region* associated to the point T_k^n.

Each point T_k^n of μ^n may be thought of as a codeword. When this codeword is transmitted, the channel adds to it the displacement vector X_k^n, so that the received signal is $T_k^n + X_k^n$. We would want the received signal to land in the correct decoding region. With this interpretation, we define the *probability of error* as

$$p_e(n) = \lim_{W \to \infty} \frac{\sum_k \mathbb{1}_{T_n^k \in B(0,W)} \mathbb{1}_{T_k^n + X_k^n \notin C_k^n}}{\sum_k \mathbb{1}_{T_n^k \in B(0,W)}}. \qquad (1)$$

The *probability of success* is defined as $p_s(n) = 1 - p_e(n)$. The limit in (1) exists almost surely from the assumption that the marked point process μ^n with marks (X_k^n, C_k^n) is stationary and ergodic.

Let $\mathbb{P}_n^0$ denote the Palm probability of μ^n. Likewise, $\mathbb{E}_n^0$ denotes Palm expectation for μ^n. The pointwise ergodic theorem implies that

$$p_e(n) = \mathbb{P}_n^0 \left(X_0^n \notin C_0^n \right) \ .$$

V. ADDITIVE GAUSSIAN NOISE CAPACITY OF A STATIONARY POINT PROCESS

Let us restrict attention to the case where we start with point processes μ^n of intensity λ_n in $\mathbb{R}^n$, where the displacement vectors are independent of the points, each displacement vector being Gaussian with i.i.d. coordinates having mean zero and variance σ^2 and with the displacement vectors being i.i.d. from point to point. Further, C_k^n is the Voronoi cell of T_k^n in the point configuration μ^n. Write λ_n as e^{nR_n}. Then we can prove:

Theorem 1 For any subsequence $n_k \to \infty$ such that $\liminf_{k \to \infty} R_{n_k} > \frac{1}{2} \ln \frac{1}{2\pi e \sigma^2}$, we have $\lim_{k \to \infty} p_e(n_k) = 1$. □

Further, we can prove:

Theorem 2 Let μ^n be a Poisson process of intensity $\lambda_n = e^{nR_n}$. For any subsequence $n_k \to \infty$, if $\limsup_{k \to \infty} R_{n_k} < \frac{1}{2} \ln \frac{1}{2\pi e \sigma^2}$, we have $\lim_{k \to \infty} p_e(n_k) = 0$. □

Together these theorems are the analog of the Poltyrev capacity theorem for stationary ergodic point process in the case of Gaussian noise. Theorem 2 is a consequence of the evaluation of the reliability exponent of Poisson processes in additive Gaussian noise, which is sketched in the next subsection. We now sketch a proof of Theorem 1. The Palm stationary expectation of the volume of the Voronoi cell of the origin, i.e. $\mathbb{E}_n^0[Vol(C_0^n)]$, is precisely e^{-nR_n}, i.e. the reciprocal of the intensity. Under the assumptions, there is some $\sigma > 2c > 0$ such that, for all sufficiently large k, with $n = n_k$, we have

$$\mathbb{E}_n^0[Vol(C_0^n)] \leq e^{\frac{n}{2} \ln 2\pi e(\sigma - 2c)^2} \ .$$

Hence

$$\mathbb{P}_n^0(Vol(C_0^n) > e^{\frac{n}{2} \ln 2\pi e(\sigma - c)^2}) \leq \left(\frac{\sigma - 2c}{\sigma - c} \right)^n \ .$$

Now, $\mathbb{P}_n^0(X_0^n \notin C_0^n)$ is at least as big as the probability that X_0^n lands outside the ball centered at the origin with volume $Vol(C_0^n)$. The volume of the ball of radius $\sigma\sqrt{n}$ is

$$V_n(\sigma\sqrt{n}) = \frac{(n\sigma^2)^{n/2} \pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} = e^{\frac{n}{2}\left(\ln(2\pi e \sigma^2) + o(1)\right)} \ .$$

Hence

$$\mathbb{P}^0_n(X^n_0 \notin C^n_0) \geq \mathbb{P}^0_n(\sum_{k=1}^{n} X^n_0(k)^2 \geq (\sigma - c)^2) \left[1 - \left(\frac{\sigma - 2c}{\sigma - c}\right)^n\right],$$

which tends to 1 as $n \to \infty$, by the strong law of large numbers. □

If $X(1), \ldots, X(n)$ are i.i.d. Gaussian random variables with mean zero and variance σ^2, the density of $\sqrt{X(1)^2 + \cdots + X(n)^2}$ at r is $g^\sigma_n(r) = \frac{1}{\sigma} g^1(\frac{r}{\sigma})$, where

$$g^1_n(r) = 1_{r>0} e^{-r^2/2} \frac{1}{2^{n/2}} r^{n-1} \frac{2}{\Gamma(n/2)}.$$

The proof of Theorem 1 was quite straightforward. If we want to do a more careful analysis of the probability of error we will need to start with the formula

$$p_s(n) = \int\limits_{r \geq 0} \int\limits_{\vec{v} \in \mathbb{S}^{n-1}} \mathbb{P}^0_n(\mu^n(B(r\vec{v}, r)) = 0) \frac{g^\sigma_n(r)}{A_{n-1}} d\vec{v} dr \ . \quad (2)$$

VI. ADDITIVE GAUSSIAN NOISE RELIABILITY OF A STATIONARY POINT PROCESS

In this section we will show how the point process viewpoint recovers the Poltyrev exponent. We consider first a sequence of Poisson processes μ^n of rates $\lambda_n = e^{nR}$ where $R = \frac{1}{2} \ln \frac{1}{2\pi e \alpha^2 \sigma^2}$ for $\alpha > 1$. We show that the limit $\lim_{n \to \infty} -\frac{1}{n} \ln p_e(n)$ (i.e. the error exponent) in the case of independent Gaussian displacements and the Voronoi decomposition associated to the points, exists and equals the *random coding* part of the Poltyrev exponent (see [10] for a definition). We then consider sequences of Matérn I point processes with parameters as in the Poisson case and show that the error exponent in this case equals the Poltyrev exponent.

A. Reliability of Poisson processes in white Gaussian noise

From (2), if μ^n is a Poisson point process with intensity λ_n, then

$$p_e(n) = \int_0^\infty \left(1 - e^{-\lambda_n V_n(r)}\right) g^\sigma_n(r) dr.$$

This is because Slyvniak's theorem tells us that for Poisson processes we have $\mathbb{P}^0_n(\mu^n(B(r\vec{v}, r)) = 0) = \mathbb{P}(\mu^n(B(r\vec{v}, r)) = 0) = e^{-\lambda_n V_n(r)}$

We have

$$g^\sigma_n(v\sigma\sqrt{n}) = e^{-n\left(\frac{v^2}{2} - \frac{1}{2} - \ln(v) + o(1)\right)}.$$

With $\lambda_n = e^{\frac{n}{2} \ln 2\pi e \sigma^2 \alpha^2}$, where $\alpha \geq 1$, we have

$$1 - e^{-\lambda_n V_n(v\sigma\sqrt{n})} = e^{-n\left((\ln \alpha - \ln v)^+ + o(1)\right)} \ .$$

We may therefore write the probability of error as

$$p_e(n) = \int_0^\infty e^{-n(\frac{v^2}{2} - \frac{1}{2} - \ln v + (\ln \alpha - \ln v)^+ + o(1))} dv \ .$$

For any $\alpha \geq 1$, we need to consider the minimum over all $v > 0$ of the corresponding contribution (i.e. from points at $v\sigma\sqrt{n}$). Let

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

We observe that exponent of the integrand at v is of the form $G(v) + b(v)$, where

$$G(v) = (\ln \alpha - \ln v)^+ \ .$$

Carrying out the minimization over v of $G(v) + b(v)$ gives $\frac{\alpha^2}{2} - \frac{1}{2} - \ln \alpha$ when $1 < \alpha < \sqrt{2}$ and $\frac{1}{2} - \ln 2 + \ln \alpha$ when $\alpha > \sqrt{2}$. The lower bound to $\eta(\alpha)$ comprised of these two portions is called the random coding error exponent, see [10].

B. Reliability of a Matérn process in white Gaussian noise

A Matérn point process can be created by dropping points from a Poisson process as follows. Choose some positive radius. If a point is within this fixed radius of another point, it is dropped. For an information theorist, this should be reminiscent of expurgation. A generalization, also called a Matérn process, is created from a Poisson processes as follows. Mark each point of the initial Poisson process with an independent real number chosen uniformly over the unit interval. Choose some positive radius. Every point of the original Poisson process examines the ball around it of this radius and will survive only if its associated mark is strictly larger than the marks of all the other points in this ball [3]. These processes were introduced by Matérn [9].

Fix $\epsilon > 0$. Consider a sequence of Matérn processes μ^n constructed by first starting with a sequence of Poisson processes of rates $\lambda_n = e^{nR}$ where $R = \frac{1}{2} \ln \frac{1}{2\pi e \alpha^2 \sigma^2}$ for $\alpha > 1$, and then only retaining the points in the Poisson process that have no other points in a ball of radius $(\alpha - \epsilon)\sigma\sqrt{n}$ around them. We think of ϵ as being arbitrarily small. Irrespective of how small ϵ is, the intensity of the process of surviving points is $e^{n(R+o(1))}$ because the Palm probability (relative to the original Poisson process) that any point is dropped is asymptotically vanishing. When one considers the formula (2) to compute the error probability, one has to work with the Palm probability in the Matérn process, i.e. with respect to the surviving points. But around any surviving point there is no point within a radius of $(\alpha - \epsilon)\sigma\sqrt{n}$. This brings in the following phenomena, in the limit as $\epsilon \to 0$: (i) if $v < \frac{\alpha}{2}$ then the contribution from the probability that there is a point in the ball of radius $v\sigma\sqrt{n}$ around a point at distance $v\sigma\sqrt{n}$ from the origin becomes ∞ (at the level of the exponent, i.e. where we used $\ln \alpha - \ln v$ in the Poisson process calculation, we can now replace it by ∞ for such v); and (ii) for $\frac{\alpha}{2} < v < \frac{\alpha}{\sqrt{2}}$ the contribution $\ln \alpha - \ln v$ can be replaced by $\ln \alpha - \frac{1}{2} \ln(v^2 - (v - \frac{\alpha^2}{2v})^2)$. This is because we can replace the volume of the ball of radius $v\sigma\sqrt{n}$ around the point at distance $v\sigma\sqrt{n}$ from the origin (along some ray) by the volume of the portion of this ball that is cut out by a hyperplane perpendicular to the ray and at a distance $b\sigma\sqrt{n}$ from it (i.e. a distance of $(v + b)\sigma\sqrt{n}$ along this ray from the origin) where $b = \frac{\alpha^2}{2v} - v$ by elementary geometry. Actually we only need the volume of the portion of this set that is not in the sphere of radius $\alpha\sigma\sqrt{n}$ around the origin, but our generosity won't affect the asymptotics. The

---

[3] This generalization will not be analyzed in this document, but is mentioned because it is interesting.

probability of finding a surviving point in this set is being bounded above by the probability of finding any point at all in this set. This again won't affect the asymptotics, because the thinning is irrelevant on a logarithmic scale.

With these observations, one recovers the Poltyrev exponent, which can be thought of as resulting from adding the function $G(v)$ given by

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

to the "basic" function $b(v) = \frac{v^2}{2} - \frac{1}{2} - \ln v$ (which comes from the Gaussian), and then minimizing over v for each $\alpha \geq 1$. The $G(v)$ above comes from the Matérn I process, replacing $(\ln\alpha - \ln v)^+$, which came from the Poisson process.

VII. ADDITIVE NOISE CAPACITY AND RELIABILITY OF A STATIONARY POINT PROCESS

Starting with point processes μ^n of rate $\lambda_n = e^{nR_n}$ in $\mathbb{R}^n$, with the displacement vectors independent of the points, each displacement vector having i.i.d. coordinates having some density with differential entropy h, the displacement vectors being i.i.d. from point to point, we can prove the following:

Theorem 3 For any subsequence $n_k \to \infty$ such that $\liminf_{k\to\infty} R_{n_k} + h > 0$, and *any* choice C^n_k for the points T^n_k of the process (that are subsets of $\mathbb{R}^n$ jointly stationary with the points and the displacements, forming a decomposition of $\mathbb{R}^n$), we have $\lim_{k\to\infty} p_e(n_k) = 1$. □

Further, we can prove:

Theorem 4 Let μ^n be a Poisson process of intensity $\lambda_n = e^{nR_n}$. For any subsequence $n_k \to \infty$, if $\limsup_{k\to\infty} R_{n_k} + h < 0$, it is possible to choose C^n_k for the points T^n_k of the process (that are subsets of $\mathbb{R}^n$ jointly stationary with the points and the displacements, forming a decomposition of $\mathbb{R}^n$), such that $\lim_{k\to\infty} p_e(n_k) = 0$. □

The proofs of both theorems rely on the asymptotic equipartition property for random variables having a density, see e.g. the text of Cover and Thomas [4]. The proof of Theorem 3 is quite similar to that of Theorem 1, based on volume estimates. In the proof of Theorem 4, the marks C^k_n are constructed to capture the typical sequences, and the only issue is to show that they can be chosen to form a decomposition of $\mathbb{R}^n$. The details are omitted due to lack of space.

Together these results give a kind of capacity theorem for stationary point process perturbed by additive noise. This result is closely related to the concept of *information theoretic sphere packing* coined by Loeliger [8].

VIII. SUGGESTIONS FOR FURTHER WORK

The viewpoint described here suggests an approach to attack the gap between the best currently known upper and lower bounds on the error exponent of the AWGN channel. For sequences of stationary point processes μ^n in $\mathbb{R}^n$ with intensity $\lambda_n = e^{nR}$ for fixed $R = \frac{1}{2}\ln 2\pi e\alpha^2\sigma^2$, with $\alpha \geq 1$, specifically those having a repulsive structure between the points, one would attempt to describe the Palm stationary probability of finding a point at distance $v\sigma\sqrt{n}$ from the origin by a function $G(v)$ at the level of the exponent. To attack the lower bound, the aim would be to construct examples, if possible, where the exponent for the Palm probability of finding a point in a ball of radius $v\sigma\sqrt{n}$ around a point at distance $v\sigma\sqrt{n}$ from the origin gives a better minimum over v of $G(v) + b(v)$ for some value of α – of course, this can only happen, if at all, for $\alpha > \sqrt{2}$, since the true exponent is already known for $1 < \alpha < \sqrt{2}$. To attack the upper bound, one would attempt to exclude the existence of such sequences of processes with $G(v)$ for which the minimum over v of $G(v) + b(v)$ for some value of α exceeds a given value. This research program appears to offer a novel viewpoint to attack a classical problem in information theory.

IX. CONCLUDING REMARKS

We defined a class of asymptotic problems in high dimensional point process theory directly motivated by the study of error exponents in information theory. We are able to recover and to give an interesting perspective on the Poltyrev exponent. Our viewpoint also suggests a research program for an attack on the gap that exists between the best known upper and lower bounds on the reliability function of additive noise channels.

ACKNOWLEDGMENT

The work of the first author is supported by NSF grants CCF-0500234, CNS-0627161 and CCF-0635372, by Marvell Semiconductor, and by the University of California MICRO program. The collaboration of the two authors was supported by the Thomson Paris Research Lab through a grant to the second author. The work of the second author is supported by the EuroFGI Network of Excellence of the European Union and by ARC Ifany of INRIA.

REFERENCES

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

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

[3] H. Cohn and N. Elkies, "New Upper Bounds on Sphere Packings II". *Geometry and Topology*, Vol. 6, pp. 329 -353, 2002.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons Inc., New York, 1991.

[5] D. J. Daley and D. Vere-Jones *An Introduction to the Theory of Point Processes*, Springer, New York, 1988.

[6] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons Inc., New York, 1968.

[7] G. A. Kabatyanskii and V. I. Levenshtein, "Bounds for Packings on a Sphere and in Space". *Problems of Information Transmission*, Vol. 14, No. 1, pp. 1 -17, Jan. -Mar. 1978. (Translation from *Problemy Peredachi Informatsii*. Vol. 14, No. 1, pp. 3 -25, Jan. -Mar. 1978.)

[8] H.-A. Loeliger, "Averaging Bounds for Lattices and Linear Codes", *IEEE Trans. on Inform. Theory*, Vol. 43, No. 6. pp. 1767 -1773, Nov. 1997.

[9] B. Matérn. *Spatial Variation*. Meddelanden Statens Skogforkningsinstitut, Vol. 49 (1960). Second edition, Springer, Berlin, 1986.

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

[11] 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.