On Convexity of Error Rates in Digital Communications

Sergey Loyka, Victoria Kostina, Francois Gagnon

Abstract

Convexity properties of error rates of a class of decoders, including the ML/min-distance one as a special case, are studied for arbitrary constellations, bit mapping and coding. Earlier results obtained for the AWGN channel are extended to a wide class of noise densities, including unimodal and spherically-invariant noise. Under these broad conditions, symbol and bit error rates are shown to be convex functions of the SNR in the high-SNR regime with an explicitly-determined threshold, which depends only on the constellation dimensionality and minimum distance, thus enabling an application of the powerful tools of convex optimization to such digital communication systems in a rigorous way. It is the decreasing nature of the noise power density around the decision region boundaries that insures the convexity of symbol error rates in the general case. The known high/low SNR bounds of the convexity/concavity regions are tightened and no further improvement is shown to be possible in general. The high SNR bound fits closely into the channel coding theorem: all codes, including capacity-achieving ones, whose decision regions include the hardened noise spheres (from the noise sphere hardening argument in the channel coding theorem) satisfies this high SNR requirement and thus has convex error rates in both SNR and noise power. We conjecture that *all* capacity-achieving codes have convex error rates. Convexity properties in signal amplitude and noise power are also investigated. Some applications of the results are discussed. In particular, it is shown that *fading is convexity-preserving and is never good in low dimensions under spherically-invariant noise*, which may also include any linear diversity combining.

Index Terms

Convexity/concavity, error rate, BER, pairwise probability of error, maximum-likelihood decoding, unimodal noise, spherically-invariant noise.

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, 161 Louis Pasteur, Ottawa, Ontario, Canada, K1N 6N5, E-mail: sergey.loyka@ieee.org.

V. Kostina is with the Department of Electrical Engineering, Princeton University, NJ, 08544, USA (e-mail: vkostina@princeton.edu).

F. Gagnon is with the Department of Electrical Engineering, Ecole de Technologie Suprieure, Montreal QC H3C 1K3, Canada (e-mail: francois. gagnon@etsmtl.ca).

I. INTRODUCTION

Convexity properties play a well-known and important role in optimization problems [1][2]: it is essentially the class of convex problems that are solvable numerically. Furthermore, significant analytical insights are available for this class, which cannot be said about the general class of nonlinear problems [2].

In the world of digital communications, various types of error rates often serve as objective or constraint functions during optimization [3]-[10]. Therefore, their convexity properties are of considerable importance. While, in some simple scenarios, the convexity can be established by inspection or differentiation of corresponding closed-form error probability expressions [3]-[10], this approach is not feasible not only in the general case, but also in most cases of practical importance (e.g. modulation combined with coding etc.), since such expressions are either not known or prohibitively complex [11].

A general approach (i.e. not relying on particular closed-form error probability expressions) to convexity analysis in binary detection problems has been developed in [12]. This approach has been later extended to arbitrary multidimensional constellations (which can also include coding) in [13][14]. In particular, it has been shown that the symbol error rate (SER) of the maximum-likelihood (ML) decoder operating in the AWGN channel is always convex in SNR in dimensions 1 and 2, and also in higher dimensions at high SNR and concave at low SNR (with explicitly specified boundaries of the high/low SNR regimes), for any modulation and coding. Bit error rate (BER) has been shown to be convex in the high SNR regime as well. These results have been also extended to fading channels. In particular, it was shown that "fading is never good in low dimensions". In a related but independent line of study, a log-concavity property of the SER in SNR [dB] for the multi-dimensional uniform square-grid constellations (M-QAM), including fading and diversity reception, has been established in [15] and a number of new local SER bounds have been obtained based on it.

In the present paper, the earlier results in [14] are expanded in several directions, including an extension to a class of decoders and a wide class of noise densities (significantly different from Gaussian), as well as tightening the high/low SNR bounds of the convexity/concavity regions reported in [14].

While the utility of the Gaussian noise model is well-known, there are a number of scenarios

where it is not adequate, most notably an impulsive noise [18][19][23][26]-[31] with tails much heavier than Gaussian. To address this, an important and natural generalization of the Gaussian random process has been developed, namely, the spherically-invariant random process (SIRP). It has found a wide range of applications in communications, information-theoretic and signal processing areas [20]-[27]. This class of processes, while having some of the important properties of the Gaussian process, significantly extends modeling flexibility and thus can be applied to a wider range of phenomena such as impulsive noise, radar clutter, radio propagation disturbances, bandlimited speech waveforms [23][26]. While the marginal PDF of a SIRP may be significantly different from Gaussian, this class of processes shares a number of important theoretical properties with the Gaussian process: it is closed under linear transformations, it is the most general class of processes for which the optimal MMSE estimator is linear, and the optimal (ML) decoding is still the minimum distance one (this may also include fading and correlated noise) [20]-[27]. The present paper will extend this list to include the convexity properties of SER under a SIRP noise, which turn out to be similar to those in the AWGN channel (see section III-B for further details). In addition, a general class of unimodal noise power densities will be considered and conditions on an arbitrary noise density will be formulated under which the SER is convex. In particular, the SER is convex in the SNR provided that the noise power density is decreasing around the decision region boundaries, regardless of its behavior elsewhere. It is convex at high SNR under a unimodal or a SIRP noise, and its is always convex (for any SNR) in low dimensions under SIRP noise. Similar results are obtained for convexity in signal amplitude and noise power (which are important for an equalizer design and a jammer optimization). All the results formulated for an i.i.d. noise are extended to the case of correlated noise as well.

In general, convexity of the SER does not say anything about convexity of the BER, since the latter depends on pairwise probabilities of error (PEP) and not on the SER [32][5]. Since the BER is an important performance indicator and thus appears as an objective in many optimization problems, its convexity properties are also studied here using the generic geometrical framework developed for the SER analysis. The setting is general enough so that the results apply to arbitrary constellations, bit mapping and coding. It turns out that the BER is convex at high SNR for a wide class of noise distributions and a class of decoders, where the high SNR boundary is determined by the constellation minimum distance and dimensionality, all other its details being irrelevant.

While the convexity of the PEP and the BER has been established at high SNR, the question remains: how relevant this high SNR regime is, i.e. does it correspond to realistic(practical) SNR values? This has significant impact on the result's importance and its utility when solving practically-relevant optimization problems. In this paper, we provide a positive answer: the high SNR is almost the same as that required by the channel coding theorem so that any code, including capacity-achieving ones, whose decision regions include the hardened noise spheres (from the sphere packing/hardening arguments in the channel coding theorem [3][33]), are in this range. In other words, the boundary of the high SNR regime is closely matched to that in the channel coding theorem so that arbitrary low probability of error implies its convexity and hence power/time sharing does not help to reduce it further. This complements the well-known result that the capacity cannot be increased by power/time sharing. Any practical code whose decision regions include the hardened noise spheres has also convex SER, PEP and BER. This opens up an opportunity to apply numerous and powerful tools of convex optimization to design of systems with such codes on a rigorous basis.

The main contributions are summarized as follows:

• New tighter high/low SNR bounds of the convexity/concavity regions are obtained and it is demonstrated that no further improvement is possible in the general case.

• While the earlier results in [14] were established for the ML (min-distance) decoders only, the same results are shown to apply to any decoder with center-convex decision regions (see section III-A for details), of which the min-distance one is a special case.

• While the earlier results in [14] were established for the AWGN channel only, the present paper considers a wide class of noise densities of which Gaussian is a special case (e.g. generic unimodal, SIRP etc.). In particular, the SER and the BER are shown to be convex at high SNR for this wider class as well; the SER turns out to be convex in low dimensions not only for the Gaussian, but also for an arbitrary SIRP noise. The constellation dimensionality and minimum distance appear as the main factors affecting the convexity properties.

• The boundary of the high SNR regime (where the SER/BER convexity is ensured) is shown to be closely linked to the channel coding theorem, so that error rates of capacity-achieving codes (with vanishingly-small probability of error) are convex.

• Any flat-fading and any linear diversity combining are shown to be convexity-preserving, so that fading is never good in low dimensions under spherically-invariant noise, including linear

diversity combining.

Tables 1, 2 and 3 summarize the results for convexity properties in the SNR/signal power, signal amplitude and noise power. Unless otherwise indicated, a non-fading channel, an arbitrary constellation and a decoder with center-convex decision regions are assumed.

TABLE I

CONVEXITY PROPERTIES OF THE SER/PEP/BER IN THE SNR/SIGNAL POWER.

Convexity/concavity: scenario	Where
AWGN: SER is convex at high and concave at low SNR; always convex in low dimensions ($n \leq 2$).	Theorem 2, (14), (15); Corollary 2.1
Arbitrary noise density: SER is convex if the power density is non-increasing at the boundaries of decision regions	Theorem 3, (18), (19)
Unimodal noise: SER is convex at high SNR; always convex if the noise power density is non-increasing.	Corollary 3.1, 3.2
SIRP noise: SER is convex at high and concave at low SNR; always convex if $n \leq 2$.	Theorem 5, (24), (25)
AWGN: BER/PEP are convex at high SNR	See [14], [16], [17] and Theorem 9
AWGN: BER is convex for capacity-approaching codes	(34)-(37); Conjecture 1
SIRP noise: BER/PEP are convex at high SNR	Theorem 10, (38), (39)
Fading + SIRP noise (AWGN is a special case): fading is never good in low dimensions, including linear combining	Propositions 4, 6
Fading channel: any flat-fading and any linear combining are convexity preserving (under any noise)	Proposition 1, (28); Propositions 5, 6

II. SYSTEM MODEL

The standard baseband discrete-time system model in an additive noise channel, which includes matched filtering and sampling, is

$$\mathbf{r} = \mathbf{s} + \boldsymbol{\xi} \tag{1}$$

where $\mathbf{s}$ and $\mathbf{r}$ are n-dimensional vectors representing transmitted and received symbols respectively, $\mathbf{s} \in \{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_M\}$, a set of M constellation points, $\boldsymbol{\xi}$ is an additive white noise. Several noise models will be considered, including the AWGN one, in which case $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \mathbf{I})$, and the corresponding probability density function (PDF) is

$$f_\xi(\mathbf{x}) = \left(2\pi\sigma_0^2\right)^{-n/2} e^{-|\mathbf{x}|^2/2\sigma_0^2} \tag{2}$$

TABLE II

CONVEXITY PROPERTIES OF THE SER IN SIGNAL AMPLITUDE

Convexity/concavity: scenario	Where
AWGN: SER in convex at high and concave at low SNR; always convex if $n = 1$.	Theorem 6, Corollary 6.1.
Arbitrary noise density: SER is convex if the noise amplitude density is non-increasing at the boundaries of decision regions	Theorem 7
Unimodal noise: SER is convex at high SNR; always convex if the noise amplitude density is non-increasing.	(30)
SIRP noise: SER is convex at high and concave at low SNR; always convex if $n = 1$.	Theorem 8; Corollary 8.1.

TABLE III

CONVEXITY PROPERTIES OF THE SER/PEP/BER IN NOISE POWER.

Convexity/concavity: scenario	Where
AWGN: SER in convex at high(low) and concave at low(high) SNR(noise power).	Theorem 12, (43), (44); Corollary 12.1.
SIRP noise: SER is convex at high and concave at low SNR.	Theorem 13.
AWGN: PEP in convex at high and low SNR.	Theorem 14, (52), (53).
AWGN: BER in convex at high SNR.	Corollary 14.1, (54).
SIRP noise: PEP/BER are convex at high SNR	Corollary 14.2.

where σ_0^2 is the noise variance per dimension, and n is the constellation dimensionality[1]; lower case bold letters denote vectors, bold capitals denote matrices, x_i denotes i-th component of $\mathbf{x}$, $|\mathbf{x}|$ denotes L_2 norm of $\mathbf{x}$, $|\mathbf{x}| = \sqrt{\mathbf{x}^T\mathbf{x}}$, where the superscript T denotes transpose, $\mathbf{x}_i$ denotes i-th vector, $|\mathbf{A}|$ denotes the determinant of matrix $\mathbf{A}$. The average (over the constellation points) SNR is defined as $\gamma = 1/\sigma_0^2$, which implies the appropriate normalization, $\frac{1}{M}\sum_{i=1}^{M}|\mathbf{s}_i|^2 = 1$, unless indicated otherwise.

More general and distinctly different noise distributions will be considered as well, which include the SIRP and unimodal noise, see section III-A and III-B for further details.

---

[1]While we consider here a real-valued model, all the results extend to the complex-valued case as well by treating real and imaginary parts as two independent reals, so that n-D complex constellation corresponds to $2n$-D real one.

In addition to the maximum likelihood decoder (demodulator/detector), which is equivalent to the minimum distance one in the AWGN and some other channels [24][26],

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}_i} |\mathbf{r} - \mathbf{s}_i|,$$

a general class of decoders with center-convex decision regions (see Definition 1 and Fig. 1) will be considered, for which the min-distance one is a special case. The probability of symbol error P_{ei} (also known as symbol error rate, SER) given that $\mathbf{s} = \mathbf{s}_i$ was transmitted is

$$P_{ei} = \Pr\left[\hat{\mathbf{s}} \neq \mathbf{s}_i \middle| \mathbf{s} = \mathbf{s}_i\right] = 1 - P_{ci} \tag{3}$$

where P_{ci} is the probability of correct decision, and the SER averaged over all constellation points is

$$P_e = \sum_{i=1}^{M} P_{ei} \Pr\left[\mathbf{s} = \mathbf{s}_i\right] = 1 - P_c \tag{4}$$

Clearly, P_{ei} and P_{ci} possess the opposite convexity properties. P_{ei} can be expressed as

$$P_{ei} = 1 - \int_{\Omega_i} f_{\xi}(\mathbf{x})d\mathbf{x} \tag{5}$$

where Ω_i is the decision region (Voronoi region)[2], and $\mathbf{s}_i$ corresponds to $\mathbf{x} = 0$, i.e. the origin is shifted for convenience to the constellation point $\mathbf{s}_i$. For the min-distance decoder, Ω_i can be expressed as a convex polyhedron [1],

$$\Omega_i = \left\{\mathbf{x}\middle| \mathbf{A}\mathbf{x} \leq \mathbf{b}\right\}, \ \mathbf{a}_j^T = \frac{(\mathbf{s}_j - \mathbf{s}_i)}{|\mathbf{s}_j - \mathbf{s}_i|}, \ b_j = \frac{1}{2}|\mathbf{s}_j - \mathbf{s}_i| \tag{6}$$

where $\mathbf{a}_j^T$ denotes j-th row of $\mathbf{A}$, and the inequality in (6) is applied component-wise.

Another important performance indicator is the pairwise error probability (PEP) i.e. a probability $\Pr\left\{\mathbf{s}_i \to \mathbf{s}_j\right\} = \Pr\left[\hat{\mathbf{s}} = \mathbf{s}_j \middle| \mathbf{s} = \mathbf{s}_i\right]$ to decide in favor of $\mathbf{s}_j$ given that $\mathbf{s}_i$, $i \neq j$, was transmitted, which can be expressed as

$$\Pr\{\mathbf{s}_i \to \mathbf{s}_j\} = \int_{\Omega_j} f_{\xi}(\mathbf{x})d\mathbf{x} \tag{7}$$

where Ω_j is the decision region for $\mathbf{s}_j$ when the reference frame is centered at $\mathbf{s}_i$. The SER can now be expressed as

$$P_{ei} = \sum_{j \neq i} \Pr\left\{\mathbf{s}_i \to \mathbf{s}_j\right\} \tag{8}$$

---

[2]$\hat{\mathbf{s}} = \mathbf{s}_i$ if $\mathbf{r} \in \Omega_i$. If $\mathbf{r} \notin \Omega_i \forall i$, an error is declared.

and the BER can be expressed as a positive linear combination of PEPs [32]

$$\text{BER} = \sum_{i=1}^{M} \sum_{j \neq i} \frac{h_{ij}}{\log_2 M} \Pr\{\mathbf{s} = \mathbf{s}_i\} \Pr\{\mathbf{s}_i \to \mathbf{s}_j\} \tag{9}$$

where h_{ij} is the Hamming distance between binary sequences representing $\mathbf{s}_i$ and $\mathbf{s}_j$.

Note that the setup and error rate expressions we are using are general enough to apply to arbitrary multi-dimensional constellations, including coding (codewords are considered as points of an extended constellation). We now proceed to convexity properties of error rates in this general setting.

III. CONVEXITY OF SYMBOL ERROR RATES

Convexity properties of symbol error rates of the ML decoder in SNR and noise power have been established in [13][14] for arbitrary constellation/coding under ML decoding and AWGN noise and are summarized in Theorem 1 below for completeness and comparison purposes.

Theorem 1 (Theorems 1 and 2 in [14]): Consider the ML decoder operating in the AWGN channel. Its SER $P_e(\gamma)$ is a convex function of the SNR γ for any constellation/coding if $n \leq 2$,

$$d^2 P_e(\gamma)/d\gamma^2 = P_e(\gamma)'' \geq 0 \tag{10}$$

For $n > 2$, the following convexity properties hold:

- P_e is convex in the high SNR regime,

$$\gamma \geq (n + \sqrt{2n})/d_{\min}^2 \tag{11}$$

  where $d_{\min} = \min_i\{d_{\min,i}\}$ is the minimum distance from a constellation point to the boundary of its decision region over the whole constellation, and $d_{\min,i}$ is the minimum distance from $\mathbf{s}_i$ to its decision region boundary,

- P_e is concave in the low SNR regime,

$$\gamma \leq (n - \sqrt{2n})/d_{\max}^2 \tag{12}$$

  where $d_{\max} = \max_i\{d_{\max,i}\}$, and $d_{\max,i}$ is the maximum distance from $\mathbf{s}_i$ to its decision region boundary,

- there are an odd number of inflection points, $P_e(\gamma)'' = 0$, in the intermediate SNR regime,

$$(n - \sqrt{2n})/d_{\max}^2 \leq \gamma \leq (n + \sqrt{2n})/d_{\min}^2 \tag{13}$$

The same results can be extended to P_{ei} via the substitution $d_{\max(\min)} \to d_{\max,i(\min,i)}$ in the inequalities above.

A. Convexity in SNR/Signal Power

Since the high/low SNR bounds in Theorem 1 are only sufficient for the corresponding property, a question arises whether they can be further improved. Theorem 2 gives an affirmative answer and demonstrates that no further improvement is possible.

Theorem 2: Consider the ML decoder operating in the AWGN channel. Its SER $P_e(\gamma)$ has the following convexity properties: it is convex in the high SNR regime,

$$\gamma \geq (n-2)/d_{\min}^2 \tag{14}$$

it is concave in the low SNR regime,

$$\gamma \leq (n-2)/d_{\max}^2 \tag{15}$$

and there are an odd number of inflection points in-between. The high/low SNR bounds cannot be further improved without further assumptions on the constellation geometry.

Proof: See Appendix. ∎

Note that the high/low SNR bounds in Theorem 2 are tighter than those in Theorem 1, since

$$n - \sqrt{2n} < n - 2 < n + \sqrt{2n} \text{ for } n > 2.$$

Convexity of the SER for $n \leq 2$ is also obvious from this Theorem. In the case of identical spherical decision regions, a more definite statement can be made.

Corollary 2.1: Consider the case of Theorem 2 when all decision regions are spheres[3] of the same radius d. The following holds:

- The SER is strictly convex in γ in the high SNR regime:

$$P_e(\gamma)'' > 0 \text{ if } \gamma > (n-2)/d^2$$

- It is strictly concave in the low SNR regime:

$$P_i(\gamma)'' < 0 \text{ if } \gamma < (n-2)/d^2$$

- There is a single inflection point:

$$P_e(\gamma)'' = 0 \text{ if } \gamma = (n-2)/d^2$$

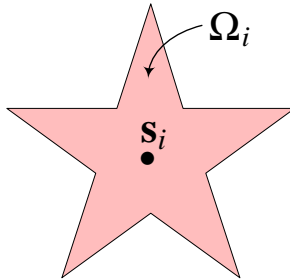[3]If the received signal does not belong to any of the decision regions, an error is declared.

Fig. 1. Center-convex decision region Ω_i centered on $\mathbf{s}_i$.

■

Note that this result cannot be obtained from Theorem 1 directly, as the bounds there are not tight. It also follows from this Corollary that the high/low SNR bounds of Theorem 2 cannot be further improved in general (without further assumptions on the constellation geometry).

The results above are not limited to the AWGN channel but can also be extended to a wide class of noise densities and a class of decoders, as Theorem 3 below demonstrates. We will need the following definition generalizing the concept of a convex region.

Definition 1: A decision region is *center-convex* if any of its points can be "seen" from the center (i.e. the corresponding line segment connecting the point to the center belongs to the region).

Note that any convex region (e.g. a convex polyhedron) is automatically center-convex but the converse is not necessarily true, so that ML/min-distance decoders are a special case of a generic decoder with center-convex decision regions. As an example, Fig. 1 illustrates such a decision region, which is clearly not convex.

To generalize the results above to a wide class of noise densities, we transform the Cartesian noise density $f_\xi(\mathbf{x})$ into the spherical coordinates $(p, \boldsymbol{\theta})$,

$$
\begin{aligned}
x_1 &= \sigma_0 \sqrt{p} \sin \theta_1 \\
x_2 &= \sigma_0 \sqrt{p} \cos \theta_1 \sin \theta_2 \\
&\vdots \\
x_{n-1} &= \sigma_0 \sqrt{p} \cos \theta_1 .. \cos \theta_{n-2} \sin \theta_{n-1} \\
x_n &= \sigma_0 \sqrt{p} \cos \theta_1 .. \cos \theta_{n-2} \cos \theta_{n-1}
\end{aligned}
\tag{16}
$$

where $\boldsymbol{\theta} = \{\theta_1, .., \theta_{n-1}\}$ are the angles, $-\pi/2 \leq \theta_i \leq \pi/2$ for $i = 1...n - 2$, $-\pi \leq \theta_{n-1} \leq \pi$, and p represents the normalized noise instant power $|\boldsymbol{\xi}|^2/\sigma_0^2$, so that

$$f_{p,\boldsymbol{\theta}}(p, \boldsymbol{\theta}) = |\partial\mathbf{x}/\partial(p, \boldsymbol{\theta})| f_\xi(\mathbf{x}) \tag{17}$$

where

$$J = |\partial\mathbf{x}/\partial(p, \boldsymbol{\theta})| = \sigma_0^n p^{n/2-1} \cos^{n-2} \theta_{n-2} \cos^{n-3} \theta_{n-3}.. \cos \theta_1$$

is the Jacobian of transformation from $\mathbf{x}$ to $(p, \boldsymbol{\theta})$, and $f_{p,\boldsymbol{\theta}}(p, \boldsymbol{\theta})$ is the noise power density in the spherical coordinates (see [35][24] for more on spherical coordinates and corresponding transformations). For simplicity of notations, we further drop the subscripts and use $f(p, \boldsymbol{\theta})$.

We are now in a position to generalize Theorem 2 to a wide class of noise densities and the class of center-convex decoders.

Theorem 3: Consider a decoder with center-convex decision regions operating in an additive noise channel of arbitrary density $f(p, \boldsymbol{\theta})$. The following holds:

$$P_e(\gamma)'' \geq 0 \text{ if } f_p'(p, \boldsymbol{\theta}) \leq 0 \;\forall\boldsymbol{\theta}, p \in [\gamma d_{\min}^2, \gamma d_{\max}^2], \tag{18}$$

where $f_p'(p, \boldsymbol{\theta}) = \partial f(p, \boldsymbol{\theta})/\partial p$. In particular, $P_e(\gamma)$ is convex in the interval $[\gamma_1, \gamma_2]$ if the noise density $f(p, \boldsymbol{\theta})$ is non-increasing in p in the interval $[\gamma_1 d_{\min}^2, \gamma_2 d_{\max}^2]$:

$$P_e(\gamma)'' \geq 0 \;\forall\gamma \in [\gamma_1, \gamma_2] \text{ if } f_p'(p, \boldsymbol{\theta}) \leq 0 \;\forall\boldsymbol{\theta}, p \in [\gamma_1 d_{\min}^2, \gamma_2 d_{\max}^2], \tag{19}$$

Proof: See Appendix. ∎

Note that it is the (non-increasing) behavior of the noise power density in the annulus $[\gamma_1 d_{\min}^2, \gamma_2 d_{\max}^2]$, i.e. around the boundaries of decision regions, that is responsible for the convexity of $P_e(\gamma)$; the behavior of the noise density elsewhere is irrelevant.

The inequalities in (18) and (19) can be reversed to obtain the corresponding concavity properties. The strict convexity properties can also be established by considering decoders with decision regions of non-zero measure in the corresponding SNR intervals. Convexity of individual SER P_{ei} can be obtained via the substitution $d_{\min(\max)} \to d_{\min,i(\max,i)}$. It is also straightforward to see that Theorem 2 is a special case of Theorem 3.

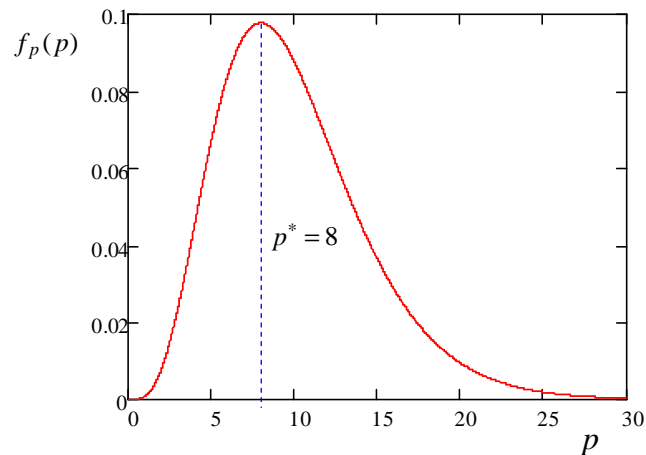Let us now consider more special cases of Theorem 3.

Fig. 2.   Gaussian noise power density for $n = 10$. It is unimodal with $p^* = 8$.

Corollary 3.1: Consider a decoder with center-convex decision regions operating in an additive noise channel of a unimodal noise power density[4],

$$f_p'(p, \boldsymbol{\theta}) \begin{cases} > 0, \ p < p^* \\ = 0, \ p = p^* \\ < 0, \ p > p^* \end{cases} \tag{20}$$

i.e. it has only one maximum at $p = p^*$; it is an increasing function on one side and decreasing on the other (see e.g Fig. 2, 3). Its SER is convex at high and concave at low SNR:

$$\begin{cases} P_e(\gamma)'' > 0, \ \gamma > p^*/d_{\min}^2 \\ P_e(\gamma)'' < 0, \ \gamma < p^*/d_{\max}^2 \end{cases} \tag{21}$$

∎

Corollary 3.2: Consider the case of monotonically-decreasing (in p) noise power density, $f_p'(p, \boldsymbol{\theta}) < 0 \ \forall p, \boldsymbol{\theta}$. Then, the SER is always convex: $P_e(\gamma)'' > 0 \ \forall \gamma$. ∎

Since the Gaussian noise power density is unimodal with $p^* = n - 2$ (see Fig. 2), Corollary 3.1 applies to the AWGN channel as well, thereby generalizing Theorem 2 to decoders with center-convex decision regions. The AWGN for $n = 1, 2$ is also a special case of Corollary 3.2.

---

[4]which is also quasi-concave [1]; many popular probability density functions are unimodal.
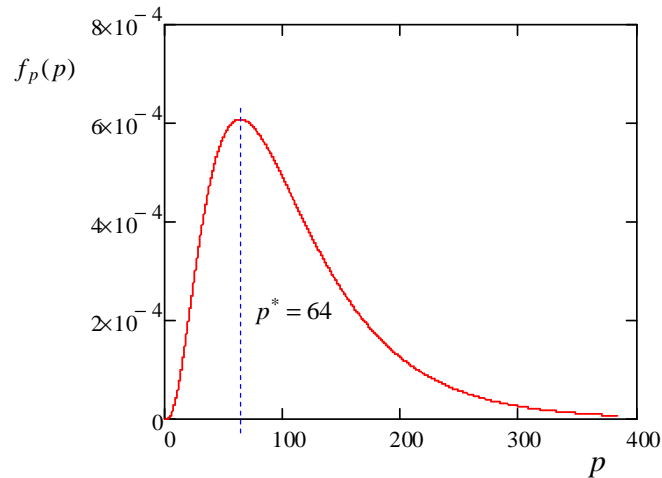
Fig. 3. Laplacian noise power density for $n = 10$. It is unimodal with $p^* = 64$.
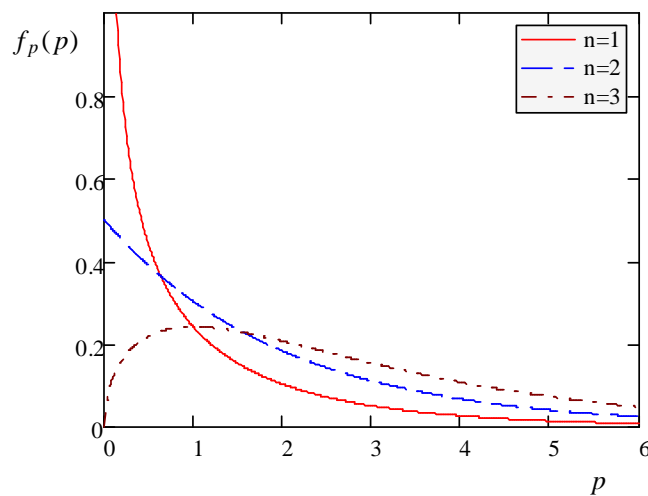

Fig. 4. The power density of Gaussian noise: while it is monotonically decreasing for $n = 1, 2$ and $\forall p$, it is unimodal for $n \geq 3$.

These Corollaries allow one to answer the question "Why is the SER in the AWGN channel always convex for $n = 1, 2$ but not for $n \geq 3$?" - the reason is the monotonically decreasing (in p) nature of the noise power density $f(p, \boldsymbol{\theta})$ for any p in the former but not the latter case, see Fig. 4.

Other examples of unimodal densities include Laplacian noise with the Cartesian PDF $f_\xi(\mathbf{x}) = c \cdot e^{-|\mathbf{x}|}$, where c is a normalizing constant, so that the spherical one is

$$f(p, \boldsymbol{\theta}) = c \cdot p^{n/2-1} e^{-\sqrt{p}} f(\boldsymbol{\theta}),$$

where $f(\boldsymbol{\theta})$ is the angular density. It is straightforward to see that this power density is unimodal in p with $p^* = (n-2)^2$, see Fig. 3. Importance of this distribution for communication/information theoretic problems is discussed in [36]-[38]. A more general example is a power exponential distribution [37]-[39]

$$f_\xi(\mathbf{x}) = c \cdot e^{-|\mathbf{x}|^{2\beta}}, \ \beta > 0,$$

(also known as generalized Gaussian [26] or, in a slightly modified form, as Weibull distribution [40]-[42]) whose spherical density is

$$f(p, \boldsymbol{\theta}) = c \cdot p^{n/2-1} e^{-p^\beta} f(\boldsymbol{\theta}), \beta > 0,$$

which is also unimodal in p with

$$p^* = \left(\frac{n-2}{2\beta} \right)^{1/\beta}$$

This distribution has a heavier (for $\beta < 1$) or lighter (for $\beta > 1$) tail than the Gaussian one, so that it offers a significant flexibility in noise modeling. In fact, it was shown in [39] that Weibull distribution can be presented as a mixture of normal distributions, where the variance of normal distribution is treated as a random variable with an α-stable distribution. This fits well into a typical model of interference in random wireless networks, where the interference distribution also follows an α-stable law [28]-[31]: each node transmits a Gaussian (capacity-achieving) signal of a fixed transmit power; at the receiver, the noise power coming from each node is random (due to random distance to transmitting nodes) and follows an α-stable law, so that the composite noise instant power follows the power exponential distribution.

Some spherically-invariant random processes or vectors considered in [22]-[26] also belong to the classes considered in these Corollaries or in Theorem 3, as discussed next.

B. Convexity of SER under SIRP noise

In this section, we consider an additive noise channel when the noise distribution follows that of a SIRP. The characterization of the SIRP class is strikingly simple: any SIRP process

is conditionally Gaussian, i.e. a Gaussian random process whose variance is a random variable independent of it. In the context of wireless communications, this structure represents such important phenomena as channel fading, random distance between transmitter and receiver etc. Below, we establish the SER convexity properties under a SIRP noise thus generalizing further the results of the previous section.

The following is one of the several equivalent definitions of a SIRP [20]-[24].

Definition 2: A random process $\{X(t), t \in R\}$ is a SIRP if a vector of any of its n samples $\mathbf{x} = \{X(t_1), X(t_2)..X(t_n)\}$ has the PDF of the following form:

$$f_{\mathbf{x}}(\mathbf{x}) = c_n h_n(\mathbf{x}^T \mathbf{C}_n^{-1} \mathbf{x}) \tag{22}$$

where $\mathbf{C}_n$ is the covariance matrix, $h_n(r)$ is a non-negative function of the scalar argument $r \geq 0$, and c_n is a normalizing constant. [5] ∎

In fact, Definition 2 says that the PDF of SIRP samples depends only on the quadratic form $\mathbf{x}^T \mathbf{C}_n^{-1} \mathbf{x}$ rather than on each entry individually, so that any linear combinations of the entries of $\mathbf{x}$ having the same variance will also have the same PDF [20]. Distributions of the functional form as in (22) are also known as elliptically-contoured distributions [35]. The characterization of SIRP is as follows (the SIRP representation theorem) [22]-[24].

Theorem 4: A random process is a SIRP iff any set of its samples has a PDF as in (22) with

$$h_n(r) = \int_0^\infty \sigma^{-n} \exp\left\{-\frac{r}{2\sigma^2}\right\} f(\sigma)d\sigma, \ 0 < r < \infty, \tag{23}$$

where $h_n(r)$ is defined by continuity at $r = 0$, and $f(\sigma)$ is any univariate PDF. ∎

An equivalent representation is $X(t) = CY(t)$, where $Y(t)$ is the Gaussian random process of unit variance, and C is an independent random variable of PDF $f(\sigma)$, so that Theorem 4 basically says that any SIRP can be obtained by modulating the Gaussian random process by an independent random variable [25]. A number of PDFs that satisfy Theorem 4 and corresponding $f(\sigma)$ can be found in [26] (which include Laplacian and power exponential densities above).

It was shown in [26] that the optimal decoder under the SIRP noise is still the minimum distance one (which follows from the fact that $h_n(r)$ in (23) is monotonically decreasing in r).

---

[5]An equivalent definition in terms of the characteristic function is also possible. Note also that not any $h_n(r)$ will do the job, but only those satisfying the Kolmogorov consistency condition [22]-[24].

Using this, we are now in a position to establish the SER convexity properties under SIRP noise with $\mathbf{C} = \mathbf{I}$.

Theorem 5: Consider an additive SIRP noise channel, where the noise density is as in (22) and (23) with $\mathbf{C} = \mathbf{I}$. Assume that $f(\sigma)$ in (23) has bounded support: $f(\sigma) = 0 \ \forall \sigma \notin [\sigma_1, \sigma_2]$. Then, the SER of any decoder with center-convex decision regions (including the min-distance/ML one as a special case) operating in this channel is convex at high and concave at low SNR as follows:

$$P_e(p_s)'' \ \geq \ 0 \text{ if } p_s \geq (n-2)\sigma_2^2/d_{\min}^2 \tag{24}$$

$$P_e(p_s)'' \ \leq \ 0 \text{ if } p_s \leq (n-2)\sigma_1^2/d_{\max}^2 \tag{25}$$

where p_s is the signal power, and $d_{\min(\max)}$ is the minimum (maximum) distance in the normalized constellation (corresponding to $p_s = 1$).

Proof: See Appendix. ■

Note that the high/low SNR bounds are independent of a particular form of $f(\sigma)$, but depend only on the corresponding boundaries of its support set. A particular utility of this Theorem is due to the fact that closed-form expressions of $P_e(p_s)$ are not available in most cases so its convexity cannot be evaluated directly. The following Corollary is immediate.

Corollary 5.1: Consider a decoder with center-convex decision regions operating in the SIRP noise channel as in Theorem 5 without the bounded support assumption. Its SER $P_e(p_s)$ is always convex when $n \leq 2$: $P_e(p_s)'' \geq 0 \ \forall p_s$. ■

Thus, the SER is convex in low dimensions for all the noise densities in Table I in [26] (i.e. contaminated normal, generalized Laplace, Cauchy and Gaussian), which extends the corresponding result in Theorem 1 to a generic SIRP noise.

While Corollary 3.1 characterizes the SER convexity for the identical spherical decision regions in the AWGN channel, such a simple characterization is not possible in a SIRP channel in general (when $\sigma_1 \neq \sigma_2$), as Theorem 5 shows.

C. Non-negative Mixture is Convexity-Preserving

The next proposition generalizes further the results above and shows that any non-negative mixture of noise densities is convexity-preserving in terms of error rates of a given decoder for any variable of interest. We will need the following definition.

Definition 3: Let $\{f_i\}$ be a set of noise densities, $i = 1...m$. Its convex hull [1] is any non-negative liniear combination which is also a density,

$$\mathcal{H}\{f_i\} = \{f : f = \sum_i \alpha_i f_i, \ \alpha_i \geq 0, \sum_i \alpha_i = 1\}$$

Proposition 1: Let $P_e[f]$ be an error rate of a given decoder as a functional of noise density f and let all $P_e[f_i]$ be convex, $P_e[f_i]'' \geq 0 \ \forall i$, where the derivative is over any variable of interest (SNR, power/amplitude of signal/noise). Then,

$$P_e[f]'' \geq 0 \ \forall f \in \mathcal{H}\{f_i\} \tag{26}$$

Proof: A key observation here is to note, from (5), that that the probability of correct decision $P_c[f] = 1 - P_e[f]$ is a linear functional of the underlying noise density f,

$$P_c\left[\sum_i \alpha_i f_i\right] = \sum_k \Pr\{s = s_k\} \int_{\Omega_k} \sum_i \alpha_i f_i(\boldsymbol{x}) d\boldsymbol{x}$$

$$= \sum_i \alpha_i \sum_k \Pr\{s = s_k\} \int_{\Omega_k} f_i(\boldsymbol{x}) d\boldsymbol{x}$$

$$= \sum_i \alpha_i P_c[f_i] \tag{27}$$

Since $\alpha_i \geq 0$ and each $P_c[f_i]$ is concave, so is $P_c[f]$, from which the result follows. $\blacksquare$

Thus, a convexity of error rates of a given decoder for $\{f_i\}$ is sufficient to insure the convexity for any f in the convex hull of $\{f_i\}$.

The same preservation holds for concavity and also when the corresponding property is strict. It can be further extended to continuous mixtures as well,

$$P_e\left[\int \rho(\alpha) f_\alpha d\alpha\right]'' \geq 0 \text{ if } P_e[f_\alpha]'' \geq 0 \ \forall \alpha, \rho(\alpha) \geq 0, \int \rho(\alpha) d\alpha = 1 \tag{28}$$

where f_α is the noise density parameterized by a continuous mixture parameter α and $\rho(\alpha)$ is a (non-negative) density of α. Such a mixture can model a fading channel where α represents the channel (random) gain, so that (28) states in fact that (flat) fading is a convexity-preserving process. This will be elaborated in further details in Section VI.

Observe that the same convexity-preserving property holds in terms of the signal power/amplitude and noise power/amplitude, due to the linearity of $P_c[f]$. Note also that Proposition 1 and (28) extend the confexity/concavity properties to a very broad class of noise densities, which includes,

as a special case, the SIRP noise, and do not even assume convex or center-convex decision regions[6].

The following result is a direct consequence of Proposition 1.

Proposition 2: Let $\mathcal{C}$ be a set of all noise densities for which error rates of a given decoder are convex (in any variable of interest). $\mathcal{C}$ is a convex set.

D. Convexity in Signal Amplitude

Convexity of the SER as a function of signal amplitude $A = \sqrt{\gamma}$, $P_e(A)$, is also important for some optimization problems (e.g. an equalizer design). For the ML decoder operating in the AWGN channel those properties have been established in [14], which are summarized in Proposition 3 for completeness.

Proposition 3: Consider the ML decoder in the AWGN channel. Its SER $P_{ei}(A)$ as a function of signal amplitude A has the following convexity properties:

- $P_{ei}(A)$ is always convex in A if $n = 1$,
- For $n > 1$, it is convex in the large SNR regime $A \geq \sqrt{\alpha_1}/d_{\min,i}$ and concave in the small SNR regime $A \leq \sqrt{\alpha_2}/d_{\max,i}$, where

$$\alpha_1 = (2n + 1 + \sqrt{8n + 1})/2, \ \alpha_2 = (2n + 1 - \sqrt{8n + 1})/2$$

  and there are an odd number of inflection points in-between.

- The same applies to $P_e(A)$ via the substitution $d_{\max(\min),i} \to d_{\max(\min)}$. ■

The next Theorem provides tighter high/low SNR bounds, which cannot be further improved in general, and also extends the result to any decoder with center-convex decision regions (of which the ML/min-distance one is a special case).

Theorem 6: Consider a decoder with center-convex decision regions operating in the AWGN channel. Its SER $P_{ei}(A)$ as a function of signal amplitude A has the following convexity properties for any n:

- The SER is convex in A in the large SNR regime:

$$P_{ei}(A)'' \geq 0 \text{ if } A \geq \sqrt{n - 1}/d_{\min,i}$$

---

[6]To the best of our knowledge, this is the most general known result about the convexity properties of error rates.

- It is concave in the small SNR regime

$$P_{ei}(A)'' \leq 0 \text{ if } A \leq \sqrt{n-1}/d_{\max,i}$$

- There are an odd number of inflection points in-between.

- The bounds cannot be further tightened in general (without further assumptions on the constellation geometry).

- The same applies to $P_e(A)$ via the substitution $d_{\max(\min),i} \to d_{\max(\min)}$.

Proof: See Appendix. ∎

Note that the convexity of $P_{ei}(A)$ for $n = 1$ and any A follows automatically from this Theorem. It is straightforward to see that the bounds of Theorem 6 are indeed tighter than those of Proposition 1, since

$$\alpha_2 \leq n - 1 < \alpha_1$$

with strict inequality for $n \geq 2$. The following is a direct consequence of Theorem 6, which cannot be obtained from Proposition 1.

Corollary 6.1: Consider the case of Theorem 6 when all decision regions are the spheres of same radius d. The following holds:

- The SER is strictly convex in A in the large SNR regime:

$$P_e(A)'' > 0 \text{ if } A > \sqrt{n-1}/d$$

- It is strictly concave in the small SNR regime:

$$P_e(A)'' < 0 \text{ if } A < \sqrt{n-1}/d$$

- There is a single inflection point:

$$P_e(A)'' = 0 \text{ if } A = \sqrt{n-1}/d$$

∎

Theorem 6 can also be extended to a wide class of noise densities following the same approach as in Theorem 3.

Theorem 7: Consider a decoder with center-convex decision regions operating in an additive noise channel of arbitrary density $f(r, \boldsymbol{\theta})$, where r represents the normalized noise amplitude

$|\boldsymbol{\xi}|/\sigma_0$. The SER $P_e(A)$ is convex in A in the interval $[A_1, A_2]$ if the noise density $f(r, \boldsymbol{\theta})$ is non-increasing in r in the interval $[A_1 d_{\min}, A_2 d_{\max}]$:

$$P_e(A)'' \geq 0 \ \forall A \in [A_1, A_2] \text{ if } f_r'(r, \boldsymbol{\theta}) \leq 0 \ \forall \boldsymbol{\theta}, r \in [A_1 d_{\min}, A_2 d_{\max}], \tag{29}$$

Proof: See Appendix. ∎

Similarly to Theorem 3, the inequalities can be reversed to obtain the concavity properties and unimodal densities are a special case. In particular, Corollary 3.1 holds with the substitution $p \to r$, $\gamma \to A$, and (21) reads as

$$\begin{cases} P_e(A)'' > 0 \text{ if } A > r^*/d_{\min} \\ P_e(A)'' < 0 \text{ if } A < r^*/d_{\max} \end{cases} \tag{30}$$

The Gaussian, Laplacian and exponential noise amplitude distributions are all unimodal, with $r^* = \sqrt{n-1}$, $r^* = n - 1$ and $r^* = ((n-1)/2\beta)^{1/2\beta}$ respectively, so that the SER is always convex if $n = 1$.

For the case of a SIRP noise as in Theorem 5, one obtains the following.

Theorem 8: Consider an additive SIRP noise channel with the density as in (22), (23) and $\mathbf{C} = \mathbf{I}$. Assume that $f(\sigma)$ has bounded support: $f(\sigma) = 0 \ \forall \sigma \notin [\sigma_1, \sigma_2]$. Then, the SER of any decoder with center-convex decision regions operating in this channel is convex at high and concave at low SNR as a function of signal amplitude A:

$$P_e(A)'' \ \geq \ 0 \text{ if } A \geq \sigma_2 \sqrt{n-1}/d_{\min} \tag{31}$$

$$P_e(A)'' \ \leq \ 0 \text{ if } A \leq \sigma_1 \sqrt{n-1}/d_{\max} \tag{32}$$

where $d_{\min(\max)}$ is the minimum (maximum) distance of the normalized constellation (i.e. the one that corresponds to $A = 1$).

Proof: See Appendix. ∎

The following is immediate.

Corollary 8.1: Consider the scenario in Theorem 8 for $n = 1$ without the bounded support assumption. The SER is always convex in A: $P_e(A)'' \geq 0 \ \forall A$. ∎

E. Extension to Correlated Noise

While Theorems 2, 5, 6, 8 and corresponding Corollaries apply to a channel with i.i.d. noise, a similar result can be established when noise is not i.i.d. (i.e. correlated or/and of non-identical

variance per dimension). Let us consider the model in (1), where the noise covariance is $\mathbf{R}_\xi = E\{\boldsymbol{\xi}\boldsymbol{\xi}^+\}$. Applying the sufficient statistics approach, one can use $\mathbf{r}' = \mathbf{R}_\xi^{-1/2}\mathbf{r}$ instead of $\mathbf{r}$ as decision variables without affecting the performance (i.e. a whitening filter). The equivalent channel

$$\mathbf{r}' = \mathbf{R}_\xi^{-1/2}(\mathbf{s} + \boldsymbol{\xi}) \tag{33}$$

has i.i.d. noise $\mathbf{R}_\xi^{-1/2}\boldsymbol{\xi}$ and the equivalent constellation is $\{\mathbf{s}_i'\} = \{\mathbf{R}_\xi^{-1/2}\mathbf{s}_i\}$, so that equivalent decision regions Ω_i' and corresponding minimum/maximum distances can be found to which Theorems 2, 5, 6, 8 apply. In particular, the SER is still convex at high SNR. Note that Theorems 3 and 7 do not require the noise to be i.i.d.

IV. CONVEXITY OF BER AND CAPACITY-ACHIEVING CODES

While the previous sections have established the convexity properties of the SER, it does not imply the corresponding convexity properties of the BER as the latter depends on the pairwise probability of error and not just the SER (see e.g. (9)). The PEP and the SER have somewhat different convexity properties. The convexity of the PEP has been established in [14] and, based on it, the following result was obtained.

Theorem 9: Consider the ML decoder operating in the AWGN channel. Its BER, SER and PEP are all convex functions of the SNR, for any constellation, bit mapping and coding, in the high SNR (small noise) regime, when

$$d_{\min}^2 \geq (n + \sqrt{2n})\sigma_0^2, \tag{34}$$

∎

Note that the lower bound in (34) has an interesting interpretation: $n\sigma_0^2$ is the mean of $|\boldsymbol{\xi}|^2$ and $\sqrt{2n}\sigma_0^2$ is its standard deviation, so that (34) requires that $d_{\min}^2$ be larger than the average noise power by at least its standard deviation, which is intuitively what is required for low probability of error. Thus, the condition in (34) should be satisfied when probability of error is small.

Below, we make this statement more precise and proceed to establish practical relevance of the high-SNR regime in (34) based on the channel coding theorem. Recall that the sphere hardening argument (from the channel coding theorem) states that the noise vector $\boldsymbol{\xi}$ is contained within the sphere of radius $\sqrt{n(\sigma_0^2 + \varepsilon)}$ with high probability (approaching 1 as $n \to \infty$) [33][3], where

$\varepsilon > 0$ is a fixed, arbitrary small number, so that the decision regions should have minimum distance to the boundary

$$d_{\min}^2 \geq n(\sigma_0^2 + \varepsilon), \tag{35}$$

i.e. to enclose the hardened noise sphere of radius $\sqrt{n(\sigma_0^2 + \varepsilon)}$, to provide arbitrary low probability of error as $n \to \infty$. For any code satisfying this requirement, it follows that

$$d_{\min}^2 \geq n\left(\sigma_0^2 + \varepsilon\right) > (n + \sqrt{2n})\sigma_0^2, \tag{36}$$

for sufficiently large n and $\forall \varepsilon > 0$. Thus, for any code whose decision regions enclose the hardened noise spheres, the condition of Theorem 9 is satisfied and therefore the error rates (SER, PEP, BER) of such codes are all convex.

On the other hand, for any code whose decisions regions are enclosed by the spheres of radius $\sqrt{n + \sqrt{2n}}\sigma_0$, i.e. $d_{\max}^2 \leq (n + \sqrt{2n})\sigma_0^2$, the symbol error rates are lower bounded as

$$P_{ei} \geq \Pr\left\{\frac{|\boldsymbol{\xi}| - n}{\sqrt{2n}} > 1\right\} \approx Q(1) \approx 0.16 > 0, \tag{37}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$ is the Q-function, so that arbitrary-low probability of error is not achievable. Based on these two arguments, we conjecture the following.

Conjecture 1: Consider a capacity-achieving code designed for SNR $= \gamma_0$. Error rates of *any* such code are convex for SNR $\geq \gamma_0$, i.e. when it provides an arbitrary low probability of error. ∎

This conjecture is stronger that our convexity statement above since the latter requires the decision regions to include the hardened noise spheres, which is only a sufficient condition for arbitrarily low probability of error, so that it is possible that a capacity-achieving code violates the condition in (36). The conjecture effectively states that, if present, such a violation is minor in nature and does not affect the convexity property.

As an application of this result, we note that power/time sharing cannot reduce error rates of any code for which (34) holds. This complements the well-known result that power/time sharing cannot increase the capacity.

In summary, any code respecting the noise sphere hardening and hence having low probability of error will also have convex error rates (SER, PEP and BER). This is the way convexity intimately enters into the channel coding theorem.

Theorem 9 can also be extended to a wide class of decoders with center-convex decision regions under a SIRP noise as follows. To separate the effects of noise power (σ_0^2) and the shape of its PDF, let us consider the normalized noise $\boldsymbol{\xi}/\sigma_0$ and assume that it has the PDF as in (22), (23), where σ is now a normalized conditional standard deviation and $f(\sigma)$ is its PDF.

Theorem 10: Consider a channel with an additive SIRP noise as in Theorem 4 when the PDF $f(\sigma)$ of conditional normalized standard deviation has bounded support:

$$f(\sigma) = 0 \ \forall \sigma \notin [\sigma_1, \sigma_2] \tag{38}$$

The PEP, SER and BER of any decoder with center-convex decision regions (e.g. min-distance decoder for any constellation, bit mapping and coding) operating in this channel is a convex function of the SNR in the high SNR/low noise regime,

$$d_{\min}^2 \geq (n + \sqrt{2n})(\sigma_0 \sigma_2)^2 \tag{39}$$

Proof: See Appendix. ∎

Theorem 10 essentially says that Theorem 9 also applies to a SIRP noise channel provided the maximum conditional noise variance is used in (34). We remark that it is only the constellation dimensionality and the minimum distance that determine its BER convexity and only via the bound in (39), all its other details being irrelevant. As far as the noise is concerned, it is only the maximum conditional variance $(\sigma_0 \sigma_2)^2$ that matters and only via the same bound. A particular functional form of $f(\sigma)$ is irrelevant, i.e. many different unconditional noise distributions will induce the same convexity properties in the high-SNR regime.

V. CONVEXITY OF ERROR RATES IN NOISE POWER

In a jammer optimization problem, it is convexity properties in noise power that are important [12]. Motivated by this fact, we study below convexity of the SER, the PEP and the BER in the noise power.

The following result has been established in [14].

Theorem 11 (Theorem 4 in [14]): Consider the ML decoder operating in the AWGN channel. Its symbol error rates have the following convexity properties in the noise power σ_0^2, for any constellation/coding,

- P_e is concave in the large noise regime,

$$a_n \sigma_0^2 \geq d_{\max}^2 \tag{40}$$

where $a_n = n + 2 - \sqrt{2(n+2)}$,

- P_e is convex in the small noise regime,

$$b_n \sigma_0^2 \leq d_{\min}^2 \tag{41}$$

where $b_n = n + 2 + \sqrt{2(n+2)}$,

- there are an odd number of inflection points for intermediate noise power,

$$d_{\min}^2 b_n^{-1} \leq \sigma_0^2 \leq d_{\max}^2 a_n^{-1} \tag{42}$$

■

The following Theorem tightens the high/low SNR bounds above and also shows that the new bounds cannot be further improved in general.

Theorem 12: Consider a decoder with center-convex decision regions operating in the AWGN channel. Its symbol error rates have the following convexity properties in the noise power:

- P_e is concave in the large noise regime,

$$(n + 2)\sigma_0^2 \geq d_{\max}^2 \tag{43}$$

- P_e is convex in the small noise regime,

$$(n + 2)\sigma_0^2 \leq d_{\min}^2 \tag{44}$$

- there are an odd number of inflection points for intermediate noise power,

$$d_{\min}^2 \leq (n + 2)\sigma_0^2 \leq d_{\max}^2 \tag{45}$$

- These bounds cannot be improved in the general case.

Proof: See Appendix. ■

Note that the bounds of Theorem 12 are indeed tighter than those of Theorem 11, since

$$a_n < n + 2 < b_n \tag{46}$$

We further remark that similar results apply to $P_{e,i}$ via the substitution $d_{\max(\min)} \rightarrow d_{\max(\min),i}$.

The following Corollary, which cannot be obtained from Theorem 11, follows immediately from Theorem 12.

Corollary 12.1: Consider the scenario of Theorem 12 when all decision regions are the spheres of same radius d $(= d_{\max} = d_{\min})$. The SER has the following convexity properties in noise power σ_0^2:

- $P_e(\sigma_0^2)$ is strictly concave in the large noise regime,

$$\sigma_0^2 > d^2/(n+2) \tag{47}$$

- It is strictly convex in the small noise regime,

$$\sigma_0^2 < d^2/(n+2) \tag{48}$$

- There is a single inflection point,

$$P_e(\sigma_0^2)'' = 0 \text{ iff } \sigma_0^2 = d^2/(n+2) \tag{49}$$

■

Note that unlike $P_e(\gamma)$, which is convex in low dimensions ($n = 1, 2$) so that the transmitter cannot employ power/time sharing to reduce error rate, $P_e(\sigma_0^2)$ does not possess this property so that the jammer can increase error rate by power/time sharing even in low dimensions (in the low noise regime). In this respect, the jammer is in a more advantageous position compared to the transmitter in the AWGN channel. It is also clear from this Corollary that the high/low SNR bounds of Theorem 12 cannot be improved in general.

Theorem 12 can be also extended to a wider class of SIRP noise. As in Theorem 10, we consider the normalized noise $\boldsymbol{\xi}/\sigma_0$ to separate the effects of the noise power (σ_0^2) and the shape of its PDF, and assume that the normalized noise power has the PDF as in (22), (23), where $f(\sigma)$ has bounded support as in (38).

Theorem 13: Consider a decoder with center-convex decision regions operating in a SIRP noise channel channel under the stated-above conditions. Its symbol error rates have the following convexity properties in the noise power:

- P_e is concave in the large noise regime,

$$(n+2)(\sigma_1\sigma_0)^2 \geq d_{\max}^2 \tag{50}$$

- P_e is convex in the small noise regime,

$$(n+2)(\sigma_2\sigma_0)^2 \leq d_{\min}^2 \tag{51}$$

- there are an odd number of inflection points for intermediate noise power.

   Proof: See Appendix. ■

Let us study now the convexity/concavity properties of the PEP as a function of noise power.

Theorem 14: Consider a center-convex decoder operating in the AWGN channel. Its PEP $\Pr\{\mathbf{s}_i \to \mathbf{s}_j\}$ is a convex function of the noise power σ_0^2, for any n, in the low noise (high SNR) regime,

$$b_n \sigma_0^2 \leq d_{\min,i}^2 \tag{52}$$

and in the high noise (low SNR) regime,

$$a_n \sigma_0^2 \geq (d_{ij} + d_{\max,j})^2 \tag{53}$$

where $d_{ij} = |\mathbf{s}_i - \mathbf{s}_j|$, and has an even number of inflection points in-between.

Proof: See Appendix. ∎

Note that unlike the SER, the PEP is convex in the low SNR regime if $d_{\max,j} < \infty$. Based on this Theorem, a convexity property of the BER follows.

Corollary 14.1: For any constellation, bit mapping and coding, the BER of a center-convex decoder operating in the AWGN channel is a convex function of the noise power in the low noise (high SNR) regime:

$$b_n \sigma_0^2 \leq d_{\min}^2 \tag{54}$$

where the specifics of the constellation/code determine only the high-SNR boundary via $d_{\min}$. ∎

We remark that for any code respecting the sphere hardening argument,

$$d_{\min}^2 \geq n\left(\sigma_0^2 + \varepsilon\right) > b_n \sigma_0^2, \tag{55}$$

for sufficiently large n, so that the BER is a convex function of the noise power. For such codes, power/time sharing does not help to decrease the BER, but it is always helpful for a jammer whose objective is to increase the BER. A jammer transmission strategy to maximize the SER via a time/power sharing has been presented in [14] and, with some modifications, it can also be used to maximize the BER, following the convexity result in Corollary 14.1.

These results can also be extended to a SIRP noise channel.

Corollary 14.2: Consider a SIRP noise channel, where the conditional noise power σ^2 has bounded support

$$f(\sigma) = 0 \ \forall \sigma \notin [\sigma_1, \sigma_2]$$

The results of Theorem 14 and Corollary 14.1 apply with the substitutions $\sigma_0 \to \sigma_2$ for (52) and (54), and $\sigma_0 \to \sigma_1$ for (53). ∎

VI. CONVEXITY IN FADING CHANNELS

The convexity properties of error rates in non-fading channels can also be extended to fading channels. Let us consider the following standard flat-fading channel model, which is a generalization of (1),

$$\boldsymbol{r} = h\boldsymbol{s} + \boldsymbol{\xi} \tag{56}$$

where h is a (scalar) fading channel gain, so that the instantaneous SNR is $\gamma = |h|^2 \gamma_0$, and the instantaneous error rate is $P_e(\gamma) = P_e(|h|^2 \gamma_0)$. The average error rate $\overline{P}_e(\gamma_0)$ as a function of the average SNR $\overline{\gamma} = \gamma_0 = 1/\sigma_0^2$ is obtained by the expectation over the fading distribution,

$$\overline{P}_e(\gamma_0) = \overline{P_e(\gamma)} = \int P_e(|h|^2 \gamma_0) f(h) dh \tag{57}$$

where $f(h)$ is the PDF of h, and where $\overline{(\cdot)}$ denotes the expectation over the fading distribution.

If the instantaneous SER $P_e(\gamma)$ is convex for any SNR γ, the following result is immediate.

Proposition 4: Consider a fading channel under additive noise with monotonically-decreasing power density, e.g. a SIRP noise for $n \leq 2$. The average SER of a decoder with center-convex decision regions operating in this channel is lower bounded by the non-fading SER at the same (average) SNR for *any* fading distribution:

$$\overline{P}_e(\gamma_0) \geq P_e(\gamma_0) \tag{58}$$

i.e. *fading is never good in low dimensions under a SIRP noise.*

Proof: Follows from Jensen inequality [1] by observing that $P_e(\gamma)$ is convex in γ under the stated assumptions. ∎

Let us now consider the average error rate $\overline{P}_e(\gamma_0)$ as a function of the average SNR γ_0.

Proposition 5: Consider a fading channel in (56) and assume that the instantaneous SER $P_e(\gamma)$ is convex for any SNR γ (e.g. a SIRP noise for $n \leq 2$ or any noise with monotonically-deceasing power density under a center-convex decoder), then the average SER $\overline{P}_e(\gamma_0)$ is also convex in the average SNR γ_0 in such channel, i.e. *flat-fading is a convexity-preserving process.*

Proof: Follows from (57) since non-negative linear combination preserves convexity [1] or, equivalently, by using the convexity-preserving property in (28). ∎

We note that Propositions 4 and 5 extend the corresponding results in [14] obtained for the Gaussian noise and the ML decoder to a broad class of noise distributions and decoders. It

appears that it is the constellation dimensionality that has a major impact on convexity of the SER, rather than the specifics of the noise or the fading distribution.

These results can be further extended to diversity combining systems over such channel, which is a popular way to combat the detrimental effects of fading [3]-[5].

A. Convexity under diversity combining

Consider a maximum ratio combiner (MRC) operating over an m-branch fading channel as in (56),

$$\boldsymbol{r}_i = h_i \boldsymbol{s} + \boldsymbol{\xi}_i \tag{59}$$

where $\boldsymbol{r}_i$, h_i and $\boldsymbol{\xi}_i$, are the received signal, channel (voltage) gain and noise in i-th branch, $i = 1..m$. The i-th branch SNR is $\gamma_i = |h_i|^2 \gamma_0$ and the combiner's output SNR is $\gamma_{out} = \sum_i \gamma_i = \gamma_0 |\boldsymbol{h}|^2$ [4], where $\boldsymbol{h} = [h_1, .., h_m]^T$ is the vector of channel gains. The combiner's instantaneous error rate is $P_e(|\boldsymbol{h}|^2 \gamma_0)$ and the average error rate is

$$\overline{P}_e(\gamma_0) = \int P_e(|\boldsymbol{h}|^2 \gamma_0) f(\boldsymbol{h}) d\boldsymbol{h} \tag{60}$$

Using the same argument as in Proposition 5, this error rate is convex in the average SNR γ_0 provided the instantaneous SER is convex for any SNR.

This result can be now extended to an arbitrary linear combining of the form $\sum_i w_i \boldsymbol{r}_i$, where $\boldsymbol{w} = [w_1, ..., w_m]^T$ are the combining weights (which depend on the channel gains). The output SNR of this combiner is $\gamma_{out} = \gamma_0 |\boldsymbol{w}^T \boldsymbol{h}|^2$, assuming proper normalization $|\boldsymbol{w}| = 1$ (note that normalization does not affect the SNR), so that its average error rate is

$$\overline{P}_e(\gamma_0) = \int P_e(|\boldsymbol{w}^T \boldsymbol{h}|^2 \gamma_0) f(\boldsymbol{h}) d\boldsymbol{h} \tag{61}$$

which is also convex in the average SNR provided $P_e(\gamma)$ is convex, i.e. any linear combing is convexity-preserving. Note that the MRC is a special case of the general linear combining, with $\boldsymbol{w} = \boldsymbol{h}/|\boldsymbol{h}|$. Other special cases are the other 2 popular combining techniques: selection combining (SC), which selects the strongest branch with only one non-zero weight corresponding to that brach, and equal-gain combing, which adds coherently the required signals with unit gain [4]. All of them preserve the convexity of error rates, which is summarized below.

Proposition 6: Any linear diversity combining over *any* flat-fading channel as in (59) is convexity-preserving, i.e. given that the instantaneous SER $P_e(\gamma)$ is convex for any SNR γ,

the average SER $\overline{P}_e(\gamma_0)$ of the combiner is also convex in the average SNR γ_0 in such channel. Special cases include the maximum ratio, selection and equal gain combining. The lower bound in Proposition 4 also holds under any linear combining.

The utility of this convexity-preserving property is coming from the fact that most error rate expressions in fading channels and under diversity combining are prohibitively complex so that the straightforward evaluation of convexity via differentiation is not possible, while the results above establish the convexity indirectly and without evaluating the integrals (the most difficult part).

VII. CONCLUSION

Convexity/concavity properties of the error rates (SER, PEP, BER) in an additive noise channel have been considered. The earlier results obtained for the AWGN channel under ML (min-distance) decoder [14] have been improved and have also been extended to a class of decoders with center-convex decision regions and to a wide class of noise densities (unimodal and SIRP noise processes). In particular, the SER is shown to be a convex function of the SNR for any noise with monotonically-decreasing power density (e.g. SIRP, Laplacian, Weibull, power-exponential or AWGN noise in low dimensions). In higher dimensions, this property holds in the high SNR regime, for which the boundary has been explicitly given. The latter is such that any code that respects the sphere hardening condition of the channel coding Theorem also meets the high SNR condition so that all such codes have convex error rates (SER, PEP, BER). Fading is shown to be a convexity-preserving process, including any linear combining, and is never good in low dimensions under a SIRP noise.

All the applications discussed earlier in [14] (e.g. optimization of a spatial multiplexing system, optimum power/time sharing for a jammer and transmitter, optimal unitary precoding for an OFDM system) for the AWGN channel also hold under the general SIRP or unimodal noise and a convex-center decoder, based on the convexity properties established here.

VIII. APPENDIX

A. Proof of Theorem 2

First, we transform the Cartesian noise density $f_\xi(\mathbf{x})$ into the spherical coordinates $(p, \boldsymbol{\theta})$, where p represents the normalized noise instant power $|\mathbf{x}|^2 / \sigma_0^2$, and $\boldsymbol{\theta} = \{\theta_1, .., \theta_{n-1}\}$ are the

angles (see [35][24] for more on spherical coordinates):

$$f_{p,\boldsymbol{\theta}}(p,\boldsymbol{\theta}) = f_{\boldsymbol{\theta}}(\boldsymbol{\theta})f_p(p), \tag{62}$$

$$-\pi/2 \leq \theta_i \leq \pi/2, i = 1...n-2, -\pi \leq \theta_{n-1} \leq \pi$$

where $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $f_p(p)$ are the angular and normalized noise power densities,

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \Gamma(n/2)\pi^{-n/2}\cos^{n-2}\theta_{n-2}\cos^{n-3}\theta_{n-3}..\cos\theta_1 \tag{63}$$

$$f_p(p) = \frac{p^{n/2-1}e^{-p/2}}{2^{n/2}\Gamma(n/2)} \tag{64}$$

where $\Gamma(\cdot)$ is Gamma function. Using this, the probability of correct decision P_{ci} can be expressed as

$$P_{ci}(\gamma) = \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \int_0^{\gamma R_i^2(\boldsymbol{\theta})} f_p(p)dpd\boldsymbol{\theta} \tag{65}$$

where $D\boldsymbol{\theta}$ is the range of angles in (62) and $R_i(\boldsymbol{\theta})$ is the boundary of the normalized decision region (corresponding to $\gamma = 1/\sigma_0^2 = 1$). One can now obtain the second derivative in γ:

$$P_{ci}(\gamma)'' = \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta})f_p'(\gamma R_i^2(\boldsymbol{\theta}))R_i^4(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{66}$$

where

$$f_p'(p) = \frac{(n-2-p)p^{n/2-2}e^{-p/2}}{2^{n/2+1}\Gamma(n/2)} \tag{67}$$

so that $f_p'(\gamma R_i^2(\boldsymbol{\theta})) \leq 0$ if $\gamma R_i^2(\boldsymbol{\theta}) \geq n-2$. When the latter condition holds for any $\boldsymbol{\theta}$, i.e. when $\gamma \geq (n-2)/d_{\min,i}^2$, then the integrand in (66) is non-positive, since $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})R_i^4(\boldsymbol{\theta}) \geq 0$, and (14) follows. Reversing the inequalities, one obtains (15). To prove that the bound in (14) cannot be further improved in general (i.e. without further assumptions on the constellation geometry), consider the case when all Ω_i are spheres of the same radius $r = d_{\min} = d_{\max} < \sqrt{(n-2)/\gamma}$, so that $\gamma < (n-2)/d_{\min}^2 = (n-2)/d_{\max}^2$ and therefore $P_e(\gamma)'' < 0$, i.e. (14) is necessary for the convexity of the SER in general. The bound in (15) can be handled in the same way. The case of identical spherical decision regions follows in a straightforward way. As a side remark, we note that this proof is a significant simplification over those of Theorems 1 and 2 in [14]. ∎

B. Proof of Theorem 3

In the case of generic noise density $f(p,\boldsymbol{\theta})$, i.e. when (62) does not hold, (65) and (66) are generalized to

$$P_{ci}(\gamma) = \int_{D\boldsymbol{\theta}} \int_0^{\gamma R_i^2(\boldsymbol{\theta})} f(p,\boldsymbol{\theta})dpd\boldsymbol{\theta} \tag{68}$$

$$P_{ci}(\gamma)'' = \int_{D\boldsymbol{\theta}} f_p'(\gamma R_i^2(\boldsymbol{\theta}), \boldsymbol{\theta}) R_i^4(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{69}$$

Now observe that $P_{ci}(\gamma)'' \leq 0$ if $f_p'(\gamma R_i^2(\boldsymbol{\theta}), \boldsymbol{\theta}) \leq 0 \; \forall \boldsymbol{\theta}$, which holds if $f_p'(p, \boldsymbol{\theta}) \leq 0$ for $\gamma d_{\min}^2 \leq p \leq \gamma d_{\max}^2$, and all $\boldsymbol{\theta}$, so that (18) follows. (19) follows by observing that its condition insures that the condition in (18) is satisfied for $\gamma_1 \leq \gamma \leq \gamma_2$. ∎

C. Proof of Theorem 5

Using Theorem 4, the noise power density can be written as

$$f(p) = \int_0^\infty f(p|\sigma) f(\sigma) d\sigma \tag{70}$$

where $f(p|\sigma)$ is the conditional power density,

$$f(p|\sigma) = \frac{1}{2^{n/2} \Gamma(n/2) \sigma^2} \left(\frac{p}{\sigma^2} \right)^{n/2-1} \exp \left\{ -\frac{p}{2\sigma^2} \right\} \tag{71}$$

(68) and (69) can be written as

$$P_{ci}(p_s) = \int_{D\boldsymbol{\theta}} \int_{\sigma_1}^{\sigma_2} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f(\sigma) \int_0^{p_s R_i^2(\boldsymbol{\theta})} f(p|\sigma) dp\, d\sigma\, d\boldsymbol{\theta} \tag{72}$$

$$P_{ci}(p_s)'' = \int_{D\boldsymbol{\theta}} \int_{\sigma_1}^{\sigma_2} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f(\sigma) f_p'(p_s R_i^2(\boldsymbol{\theta})|\sigma) R_i^4(\boldsymbol{\theta}) d\sigma\, d\boldsymbol{\theta} \tag{73}$$

where now $R_i(\boldsymbol{\theta})$ is the boundary of normalized decision region corresponding to $p_s = 1$, and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is as in (63). Observe from (71) that

$$f_p'(p|\sigma) \geq 0 \text{ if } p \leq (n-2)\sigma^2 \tag{74}$$

$$f_p'(p|\sigma) \leq 0 \text{ if } p \geq (n-2)\sigma^2 \tag{75}$$

i.e. $f(p|\sigma)$ is unimodal in p with $p^* = (n-2)\sigma^2$. Using these properties in (73) and observing that $d_{\min,i} \leq R_i(\boldsymbol{\theta}) \leq d_{\max,i}$, $\sigma_1 \leq \sigma \leq \sigma_2$, the integrand in (73) is non-negative/non-positive if

$$p_s d_{\max,i}^2 \leq (n-2)\sigma_1^2 \tag{76}$$

$$p_s d_{\min,i}^2 \geq (n-2)\sigma_2^2 \tag{77}$$

so that $P_{ci}(p_s)'' \geq 0$ or $P_{ci}(p_s)'' \leq 0$ from which (24) and (25) follow.

D. Proof of Theorem 6

Using (63)-(66), $P_{ci}(A)$ can be written as

$$P_{ci}(A) = \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \int_0^{AR_i(\boldsymbol{\theta})} f(r) dr d\boldsymbol{\theta} \tag{78}$$

where $f_r(r)$ is the normalized noise amplitude density,

$$f(r) = \frac{r^{n-1} e^{-r^2/2}}{2^{n/2-1} \Gamma(n/2)} \tag{79}$$

and $R_i(\boldsymbol{\theta})$ is the decision region boundary of the normalized ($A = 1$) constellation. Therefore,

$$P_{ci}(A)'' = \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f_r'(AR_i(\boldsymbol{\theta})) R_i^2(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{80}$$

where

$$f_r'(r) = \frac{r^{n-2} e^{-r^2/2}(n-1-r^2)}{2^{n/2-1} \Gamma(n/2)}$$

so that the integrand in (80) is non-negative when $AR_i(\boldsymbol{\theta}) \leq \sqrt{n-1} \ \forall \boldsymbol{\theta}$, which is case when $Ad_{\max,i} \leq \sqrt{n-1}$ so that $P_{ei}(A)'' \leq 0$. The opposite case is similar. An odd number of inflection points follows from the continuity argument. The fact that the high/low SNR bounds cannot be tightened in general is clear from Corollary 6.1.

E. Proof of Theorem 7

In this case, (68) and (69) become

$$P_{ci}(A) = \int_{D\boldsymbol{\theta}} \int_0^{AR_i(\boldsymbol{\theta})} f(r, \boldsymbol{\theta}) dr d\boldsymbol{\theta} \tag{81}$$

$$P_{ci}(A)'' = \int_{D\boldsymbol{\theta}} f_r'(AR_i(\boldsymbol{\theta}), \boldsymbol{\theta}) R_i^2(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{82}$$

and the argument of section VIII-B goes thorough with the substitution $\gamma \to A$, $p \to r$.

F. Proof of Theorem 8

In this case, (70)-(73) are modified to

$$f(r) = \int_0^\infty f(r|\sigma) f(\sigma) d\sigma \tag{83}$$

$$f(r|\sigma) = \frac{1}{2^{n/2-1} \Gamma(n/2)\sigma} \left(\frac{r}{\sigma}\right)^{n-1} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} \tag{84}$$

$$P_{ci}(A) = \int_{D\boldsymbol{\theta}} \int_{\sigma_1}^{\sigma_2} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f(\sigma) \int_0^{AR_i(\boldsymbol{\theta})} f(r|\sigma) dr d\sigma d\boldsymbol{\theta} \tag{85}$$

$$P_{ci}(A)'' = \int_{D\boldsymbol{\theta}} \int_{\sigma_1}^{\sigma_2} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f(\sigma) f_r'(AR_i(\boldsymbol{\theta})|\sigma) R_i^2(\boldsymbol{\theta}) d\sigma d\boldsymbol{\theta} \tag{86}$$

and the rest of the proof in section VIII-C goes through with appropriate modifications.

G. Proof of Theorem 10

The SIRP noise Cartesian PDF can be expressed as:

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\sigma_1}^{\sigma_2} \sigma^{-n} g(\gamma, \sigma, \boldsymbol{x}) f(\sigma) d\sigma \tag{87}$$

where

$$g(\gamma, \sigma, \boldsymbol{x}) = \gamma^{n/2} \exp\left\{-\frac{\gamma |\boldsymbol{x}|^2}{2\sigma^2}\right\} \tag{88}$$

and the SNR $\gamma = 1/\sigma_0^2$, so that

$$\Pr\{\mathbf{s}_i \to \mathbf{s}_j\}_\gamma'' = \int_{\Omega_j} f(\boldsymbol{x})_\gamma'' \, d\boldsymbol{x} \tag{89}$$

where Ω_j is the decision region for $\mathbf{s}_j$ while the reference frame is centered on $\mathbf{s}_i$. Now observe that

$$g(\gamma, \sigma, \boldsymbol{x})_\gamma'' = \frac{1}{4} \gamma^{n/2} (w - w_1)(w - w_2) \exp(-\gamma w/2) \tag{90}$$

where $w = |\boldsymbol{x}|^2 \sigma^{-2}$ and $w_{1(2)} = (n \pm \sqrt{2n})/\gamma$, so that

$$g(\gamma, \sigma, \boldsymbol{x})_\gamma'' \geq 0 \text{ if } w \geq w_1 \tag{91}$$

and hence

$$f(\boldsymbol{x})_\gamma'' = \frac{1}{(2\pi)^{n/2}} \int_{\sigma_1}^{\sigma_2} \sigma^{-n} g(\gamma, \sigma, \boldsymbol{x})_\gamma'' f(\sigma) d\sigma \geq 0 \; \forall \boldsymbol{x} \in \Omega_j \text{ if } d_{\min,i}^2 \geq w_1 \sigma_2^2 \tag{92}$$

from which it follows that

$$\Pr\{\mathbf{s}_i \to \mathbf{s}_j\}_\gamma'' \geq 0 \; \forall i, j \text{ if } d_{\min}^2 \geq w_1 \sigma_2^2 \tag{93}$$

and, from (9), $\mathsf{BER}_\gamma'' \geq 0$ under the same condition, so that (39) follows, where $\gamma = 1/\sigma_0^2$.

H. Proof of Theorem 12

The proof follows along the same lines as that of Theorem 2, with the substitution $\gamma = 1/\sigma_0^2$. In particular, (66) is modified to

$$P_{ci}(\sigma_0^2)'' = \sigma_0^{-6} \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta})\{f_p'(\sigma_0^{-2} R_i^2(\boldsymbol{\theta}))R_i^2(\boldsymbol{\theta}) + 2\sigma_0^2 f_p(\sigma_0^{-2}R_i^2(\boldsymbol{\theta}))\}d\boldsymbol{\theta} \tag{94}$$

where the derivative of the noise power density $f_p'(p)$ is as in (67). The integrand in (94) is non-negative when $d_{\max,i}^2 \leq (n+2)\sigma_0^2$ and non-positive when $d_{\min,i}^2 \geq (n+2)\sigma_0^2$ from which (43) and (44) follow. The inflection points follow from the continuity argument. The fact that the bounds cannot be improved is clear from the equal spherical decision regions of Corollary 12.1.

I. Proof of Theorem 13

The proof is essentially a generalized version of the previous proof. Under the stated conditions, $P_{ci}(\sigma_0^2)$ can be written as

$$P_{ci}(\sigma_0^2) = \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \int_{\sigma_1}^{\sigma_2} f(\sigma) \int_0^{\sigma_0^{-2} R_i^2(\boldsymbol{\theta})} f(p|\sigma)dp d\sigma d\boldsymbol{\theta} \tag{95}$$

where $R_i(\boldsymbol{\theta})$ is the boundary of normalized decision region corresponding to $\sigma_0 = 1$ and $f(p|\sigma)$ is as in (71), so that its second derivative in σ_0^2 is

$$P_{ci}(\sigma_0^2)'' = \sigma_0^{-6} \int_{D\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \int_{\sigma_1}^{\sigma_2} f(\sigma)\{f_p'(\sigma_0^{-2}R_i^2(\boldsymbol{\theta})|\sigma)R_i^2(\boldsymbol{\theta}) + 2\sigma_0^2 f_p(\sigma_0^{-2}R_i^2(\boldsymbol{\theta})|\sigma)\}d\sigma d\boldsymbol{\theta} \tag{96}$$

and the integrand in (96) is non-negative when $d_{\max,i}^2 \leq (n+2)(\sigma_1\sigma_0)^2$ and non-positive when $d_{\min,i}^2 \geq (n+2)(\sigma_2\sigma_0)^2$ from which the Theorem follows.

J. Proof of Theorem 14

The PEP can be expressed as

$$\Pr\{\mathbf{s}_i \to \mathbf{s}_j\} = \int_{\Omega_j} f(\boldsymbol{x})d\boldsymbol{x} \tag{97}$$

where Ω_j is the decision region for $\mathbf{s}_j$ while the reference frame is centered on $\mathbf{s}_i$, and the AWGN density $f(\boldsymbol{x})$ is as in (2). The second derivative in the noise power σ_0^2 is

$$\Pr\{\mathbf{s}_i \to \mathbf{s}_j\}''_{\sigma_0^2} = \int_{\Omega_j} f(\boldsymbol{x})''_{\sigma_0^2} \, d\boldsymbol{x} \tag{98}$$
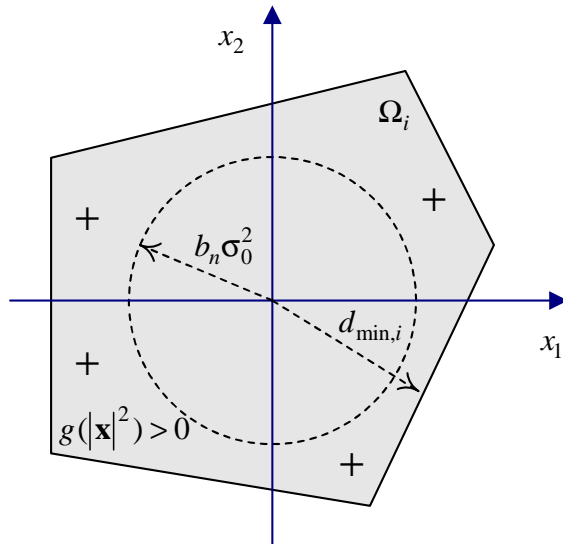
Fig. 5. Two-dimensional illustration of the problem geometry for the case $d_{\min,i}^2 \geq b_n \sigma_0^2$. The decision region Ω_i is shaded. $g(|\boldsymbol{x}|^2)$ has a sign as indicated by "+" and "-".

where $f(\boldsymbol{x})''_{\sigma_0^2}$ can be expressed as

$$f(\boldsymbol{x})''_{\sigma_0^2} = \frac{\sigma_0^8}{4} \left(\frac{\sigma_0^2}{2\pi} \right)^{n/2} \exp \left\{ -\frac{|\boldsymbol{x}|^2}{2\sigma_0^2} \right\} g(|\boldsymbol{x}|^2) \tag{99}$$

and

$$g(x) = (x - b_n \sigma_0^2)(x - a_n \sigma_0^2) \tag{100}$$

where a_n, b_n are as in Theorem 11. Clearly, $g(x) \geq 0$ if $x \geq b_n \sigma_0^2$ or $x \leq a_n \sigma_0^2$, so that $f(\boldsymbol{x})''_{\sigma_0^2} \geq 0 \; \forall \boldsymbol{x} \in \Omega_j$ if $d_{\min,i}^2 \geq b_n \sigma_0^2$ or $a_n \sigma_0^2 \geq (d_{ij} + d_{\max,j})^2$ and the result follows. Figures 5 and 6 illustrate these two cases.

REFERENCES

[1] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004

[2] A. Ben-Tal, A. Nemirovski, Lectrures on Modern Convex Optimization, MPS-SIAM Series on Optimization, Philadelphia, 2001.

[3] J.M. Wozencraft, I.M. Jacobs, Principles of Communication Engineering, Wiley, 1965.

[4] J.R. Barry, E.A. Lee, D.G. Messerschmitt, Digital Coomunications (3rd Ed.), Kluwer, Boston, 2004.

[5] S. Benedetto, E. Biglieri, Principles of Digital Transmission With Wireless Applications. New York: Kluwer Academic, 1999.

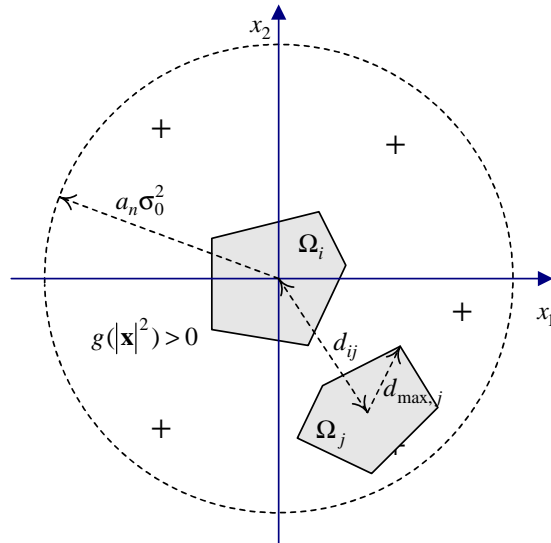Fig. 6. Two-dimensional illustration of the problem geometry for the case $a_n \sigma_0^2 \geq (d_{ij} + d_{\max,j})^2$.

[6] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization, IEEE Trans. Signal Processing, v.51, N.9, pp. 2381-2401, Sep. 2003.

[7] D. P. Palomar and Y. Jiang, MIMO Transceiver Design via Majorization Theory, Found. Trends Commun. Inf. Theory, vol. 3, no. 4-5, pp. 331-551, 2006.

[8] V. Kostina, S. Loyka, On Optimum Power Allocation for the V-BLAST, IEEE Transactions on Communications, v. 56, N. 6, pp. 999-1012, June 2008.

[9] Y.-P. Lin, S.-M. Phoong, BER Minimized OFDM Systems With Channel Independent Precoders, IEEE Trans. Signal Processing, v.51, N.9, pp. 2369-2380, Sep. 2003.

[10] C.C. Yeh, J.R. Barry, Adaptive Minimum Bit-Error Rate Equalization for Binary Signaling, IEEE Trans. Communications, v.48, N.7, pp. 1226-1235, Jul. 2000.

[11] I. Sason, S. Shamai, Performance Analysis of Linear Codes Under Maximum-Likelihood Decoding: A Tutorial, Found. Trends Commun. Inf. Theory, vol. 3, no. 1/2, pp. 1-222, 2006.

[12] M. Azizoglu, Convexity Properties in Binary Detection Problems, IEEE Trans. Inform. Theory, v. 42, N. 4, pp. 1316-1321, July 1996.

[13] S. Loyka. V. Kostina, F. Gagnon, Symbol Error Rates of Maximum-Likelihood Detector: Convex/Concave Behavior and Applications, IEEE International Symposium on Information Theory (ISIT'07), June 2007, Nice, France.

[14] S. Loyka, V. Kostina, F. Gagnon, Error Rates of the Maximum-Likelihood Detector for Arbitrary Constellations: Convex/Concave Behavior and Applications, IEEE Transactions on Information Theory, v. 56, N. 4, pp. 1948-1960, Apr. 2010.

[15] A. Conti et al., Log-Concavity Property of the Error Probability with Application to Local Bounds for Wireless Communications, IEEE Trans. Information Theory, v.55, N.6, pp. 2766-2775, June 2009.

[16] S. Loyka, V. Kostina, F. Gagnon, Bit Error Rate is Convex at High SNR, 2010 International Zurich Seminar on

Communications (IZS-10), ETH Zurich, Switzerland, Mar. 3-5, 2010.

[17] S. Loyka, F. Gagnon, V. Kostina, Error Rates of Capacity-Achieving Codes Are Convex, IEEE International Symposium on Information Theory, Austin, USA, June 2010.

[18] D. Middleton, Statistical-Physical Models of Electromagnetic Interference, IEEE Trans. EMC, v. 19, N. 3, pp. 106-127, Aug. 1977.

[19] D. Middleton, Non-Gaussian Noise Models in Signal Processing for Telecommunications: New Methods and Results for Class A and Class B Noise Models, IEEE Trans. Information Theory, v. 45, N. 4, pp. 1129-1149, May 1999.

[20] A. M. Vershik, Some characteristic properties of Gaussian stochastic processes, Theory Prob. Appl. (USSR), vol. 9, pp. 353-356, 1964.

[21] I. F. Blake and J. B. Thomas, On a class of processes arising in linear estimation theory, IEEE Trans. Inform. Theory, vol. IT-14, pp. 12-16, Jan. 1968.

[22] K. Yao, A Representation Theorem and Its Applications to Spherically-Invariant Random Processes, IEEE Trans. Info. Theory, v.IT-19, N. 5, pp. 600-608, Sep. 1973.

[23] K. Yao, Shperically Invariant Random Processes: Theory and Applications, in V.K. Bhargava et al, Communications, Information and Network Security, Kluwer, Boston, 2003.

[24] J. Goldman, Detection in the Presence of Spherically Symmetric Random Vectors, IEEE Trans. Info. Theory, v. IT-22, N. 1, pp. 52-59, Jan. 1976.

[25] G.L. Wise, N.C. Gallagher, On Spherically Invariant Random Processes, IEEE Trans. Info. Theory, v. IT-24, N. 1, pp. 118-120, Jan. 1978.

[26] E. Conte, M. Di Bisceglie, M. Lops, Canonical Detection in Spherically Invariant Noise, IEEE Trans. Communications, v. 43, N. 2/3/4, pp. 347-353, Feb./Mar./Apr. 1995.

[27] E. Conte, M. Di Bisceglie, M. Lops, Optimum Detection of Fading Signals in Impulsive Noise, IEEE Trans. Communications, v. 43, N. 2/3/4, pp. 869-876, Feb./Mar./Apr. 1995.

[28] E.S. Sousa and J.A. Silvester, "Optimum Transmission Tanges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network," IEEE Journal on Selected Areas in Communications, vol.8, no.5, pp.762-771, Jun 1990.

[29] J. Ilow and D. Hatzinakos, "Analytic Alpha-Stable Noise Modeling in a Poisson Field of Interferers or Scatterers," IEEE Transactions on Signal Processing, vol.46, no.6, pp.1601-1611, Jun 1998.

[30] V. Mordachev and S. Loyka, "On Node Density - Outage Probability Tradeoff in Wireless Networks," IEEE Journal on Selected Areas in Communications, vol.27, no.7, pp.1120-1131, Sept. 2009.

[31] M. Haenggi, R.K. Ganti, Interference in Large Wireless Networks, Foundations and Trends in Networking, v. 3, N. 2, pp. 127-248, 2009.

[32] J. Lassing et al, Computation of the Exact Bit-Error Rate of Coherent M-ary PSK with gray Code Bit Mapping, IEEE Trans. Communications, v. 51, N. 11, pp. 1758-1760, Nov. 2003.

[33] C.E. Shannon, Communication in the Presence of Noise, the Proceedings of the IRE, vol. 37, N.1, pp. 10-21, Jan. 1949.

[34] V. Kostina, Optimization and Performance Analysis of the V-BLAST Algorithm, MS Thesis, University of Ottawa, Canada, 2006.

[35] T.W. Anderson, An Introduction to Multivariate Statistical Analysis (3rd Ed.), Wiley-Interscience, 2003.

[36] S. Verdu, The Exponential Distribution in Information Theory, Problems of Information Transmission, v. 32, N. 1, pp. 86-95, 1996.

[37] J. K. Lindsey, Multivariate Elliptically Contoured Distributions for Repeated Measurements, Biometrics, Vol. 55, No. 4, pp. 1277-1280, Dec. 1999.

[38] S. Kotz et al, The Laplace distribution and generalizations: a revisit with applications to communications, economics, and finance, Birkhauser, Boston, 2001.

[39] E. Gomez-Sanchez-Manzano et al, Multivariate Exponential Power Distributions as Mixtures of Normal Distributions with Bayesian Applications, Communications in Statistics - Theory and Methods, v. 37, N. 6, pp. 972 - 985, 2008.

[40] N.C. Sagias, G.K. Karagiannidis, Gaussian Class Multivariate Weibull Distributions: Theory and Applications in Fading Channels, IEEE Trans. Information Theory, v. 51, N.10, pp. 3608-3619, Oct. 2005.

[41] M. Sekine, Y. Mao, Weibull Radar Clutter, IEE, London, 1990.

[42] H. Rinne, The Weibull Distribution: a Handbook, Chapman and Hall/CRC, Boca Raton, 2008.