Gaussian Red Alert Exponents: Geometry and Code Constructions

Bobak Nazer and Stark C. Draper
University of Wisconsin, Madison
ECE Department
Email: {bobak,sdraper}@ece.wisc.edu

Abstract—In non-block “streaming” data settings it has been observed that the availability of feedback dramatically changes the attainable communication reliability as a function of delay. The codebooks upon which these results depend contain a special and extremely reliable “red alert” message, used to indicate to the decoder when its best guess of the prior transmission is erroneous. In this paper we develop analogous results for additive white Gaussian channels that operate under both average and peak power constraints. We derive achievable error exponents for the special message for all rates up to capacity and provide geometric intuition for our coding scheme and error analysis.

I. INTRODUCTION

Consider a communication system that is tasked to transmit information at rate $R$ with a vanishing error probability. On top of this requirement, the system must also be able to transmit one special message, deemed the “red alert” message, with as small an error probability as possible. Surprisingly, even when the system is transmitting regular messages at capacity, it is still possible to imbue the red alert message with a probability of error that vanishes exponentially in the blocklength.

Prior work by Borade, Nakiboğlu, and Zheng investigated this phenomena at capacity (with multiple priority levels) in the discrete memoryless case [1]. In [2], Draper and Sahai applied similar ideas to streaming data systems and derived the red alert exponent for the binary symmetric channel (BSC) at all rates. In brief, the basic idea is to communicate the standard codewords with a fixed composition codebook and devote the rest of space to sending the red alert codeword. Let $p$ denote the crossover probability of the BSC and $q$ the probability that a symbol in the codebook is a one. Then, the red alert exponent is

$$E_{\text{ALERT}}(R) = \max_{h_B(p+q) - h_B(p)} D(q||p) \quad (1)$$

where $h_B(p)$ denotes the binary entropy function and $D(q||p)$ denotes the Kullback-Leibler divergence. Geometrically, one can think of the space $\{0,1\}$ as a sphere with the red alert codeword placed at one of the poles. The remainder live below the “latitude” dictated by the probability $q$ used to generate the codebook.

In this paper, we will develop codes for sending a red alert message over additive white Gaussian noise (AWGN) channels with both peak and average power constraints. Our aim is to derive achievable error exponents for the red alert message as well as gain some intuition into the high-dimensional geometry. As it will turn out, our code construction is related, in a certain sense, to that for the BSC. In particular, our red alert message will be placed along a pole at a distance dictated by the peak power constraint and the remainder of the codewords will live on a spherical cap whose “latitude” will be set by the required rate.

We also mention the prior work on protecting special messages in [3], [4], [5]. Due to space considerations, we do not attempt a full survey of prior work in unequal error protection. We refer the interested to the reference list of [1] for more details.

II. PROBLEM STATEMENT

Definition 1 (Messages): The transmitter has a message $w$ that is chosen randomly and uniformly from $\{0, 1, 2, \ldots, M\}$. One of the messages, $w = 0$, is a red alert message that will be afforded extra error protection.

Definition 2 (Encoder): The encoder $E$ maps the message $w$ into a length-$n$ real-valued codeword $x$ for transmission over the channel, $E: \{0, 1, 2, \ldots, M\} \rightarrow \mathbb{R}^n$. Let $x(w)$ denote the codeword used for message $w$ and let $C$ denote the entire codebook, $C = \{x(0), x(1), \ldots, x(M)\}$. The codebook must satisfy both average and peak power constraints,

$$\frac{1}{M} \sum_{w=1}^{M} ||x(w)||^2 \leq P_{\text{avg}} \quad (2)$$

$$||x(w)||^2 \leq P_{\text{peak}} \quad \forall w \in \{0, 1, 2, \ldots, M\} \quad (3)$$

for some $0 < P_{\text{avg}} \leq P_{\text{peak}}$. The rate of the codebook is

$$R = \frac{1}{n} \log M \quad (4)$$

where $\log$ represents the logarithm to base $2$.

Definition 3 (Channel): The channel outputs the transmitted vector, corrupted by independent and identically distributed

1Note that we did not include the red alert codeword in the average power constraint. It can easily be included but this requires somewhat more cumbersome notation. Specifically, the standard codewords would be chosen to satisfy a power constraint $P_{\text{avg}} - \epsilon$ so that the average power constraint is met for large blocklengths. By choosing small enough $\epsilon$, we can approach the performance derived under our current formulation.
The angle between two vectors $a$ and $b$ is

$$\angle(a, b) = \cos^{-1}\left(\frac{a \cdot b}{\|a\| \cdot \|b\|}\right)$$

where $\cos^{-1}(\cdot)$ takes values between 0 and $\pi$. Let $\mathcal{V}_n(a, b, \theta)$ denote the $n$-dimensional cone with origin at $a$ with center axis running from $a$ to $b$ with half-angle $\theta$ taking values from 0 to $\pi$. Note that if the half-angle is larger than $\pi/2$, the resulting region is not convex. See Figure 1 for an illustration. We also define the spherical cap $\mathcal{U}_n(\theta)$ to be the intersection of the sphere of radius $\sqrt{nP}\avg$ and a cone of half-angle $\theta$.

$$\mathcal{U}_n(\theta) \triangleq \mathcal{S}_n(0, \sqrt{nP}\avg) \cap \mathcal{V}_n(0, -1, \theta).$$

We now provide a detailed description of our random codebook construction. As mentioned earlier, all of the standard codewords will live in a spherical cap $\mathcal{U}_n(\theta)$ formed by the intersection of the sphere of radius $\sqrt{nP}\avg$ and the cone of half-angle $\theta$. Our codebook construction for $C$ consists of the following steps:

1. Choose $\epsilon > 0$ and $R < C - \epsilon$.
2. The red alert codeword is placed either at the origin, $x(0) = 0$, or at the boundary of the peak power constraint, $x(0) = \sqrt{\text{peak}}\cdot 1$.
3. Choose $\theta > 0$ such that the fraction of the surface area of the sphere $\mathcal{S}_n(0, \sqrt{nP}\avg)$ covered by the spherical cap $\mathcal{U}_n(\theta)$ is equal to $2^{-n(C-\epsilon)-\epsilon} + 1$.
4. Generate a codebook $\mathcal{C}_{\text{sphere}}$ with $2^{n(C-\epsilon)}$ codewords, each drawn independently and uniformly over the surface of the sphere $\mathcal{S}_n(0, \sqrt{nP}\avg)$.
5) The standard messages $x(1), \ldots, x(2^{nR})$ of $\mathcal{C}$ are comprised of the codewords of $\mathcal{C}_{\text{sphere}}$ that fall within the spherical cap $\mathcal{U}_n(\theta)$. If the intersection contains more than $2^{nR}$ codewords, we choose $2^{nR}$ independently and uniformly and throw out the rest. If it contains less, we declare that the entire codebook is in error.

We now proceed to bound the probability that the spherical cap contains less than $2^{nR}$ codewords.

**Lemma 1:** For a random instance of $\mathcal{C}_{\text{sphere}}$, the probability that $\mathcal{C}_{\text{sphere}} \cap \mathcal{U}_n(\theta)$ (where $\theta$ is selected as in the construction above) contains less than $2^{nR}$ codewords is upper bounded as follows:

$$\mathbb{P}( |\mathcal{C}_{\text{sphere}} \cap \mathcal{U}_n(\theta)| < 2^{nR} ) \leq \exp \left( -2^{nR-2} \right).$$

Due to space constraints, we have omitted the proof (which follows by a Chernoff bound).

Assuming the red alert codeword is not used, we now show that the expected average probability of error for the codewords to derive a bound on the maximum error probability for the conical codebook construction is small. Note that the expectation is taken over the randomness in the codebook construction.

**Remark 1:** Usually, it is sufficient to throw out half the codewords to derive a bound on the maximum error probability. However, in our setting, this would correspond to throwing out $2^{n(C-\epsilon)}$ potential codewords over the surface of the sphere. This is much larger than the size of our $2^{nR}$ conical codebook and could easily wipe out the entire set.

**Lemma 2:** For any $\epsilon > 0$ and $n$ large enough, the expected probability of error for the standard codewords in the codebook $\mathcal{C}$ is upper bounded by $\epsilon$,

$$\mathbb{E}_{\mathcal{C}}[p_{\text{avg}}] = \mathbb{E}_{\mathcal{C}} \left[ \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \mathbb{P}(\hat{w} \neq w) \right] \leq \epsilon .$$

**Proof:** Let $\hat{x}(w)$ represent a codeword from $\mathcal{C}_{\text{sphere}}$ for $w = 1, 2, \ldots, 2^{n(C-\epsilon)}$. Using standard typicality arguments, it can be shown that for maximum likelihood decoding, the probability of decoding to $\hat{x}(w_2)$ when $\hat{x}(w_1), w_1 \neq w_2$, was transmitted is at most $2^{-n(C-\frac{\epsilon}{2})}$. We now use this result to upper bound the probability of error between two codewords that lie in the spherical cap $\mathcal{U}_n(\theta)$. Without loss of generality, assume that $\hat{x}(w_1)$ lies in $\mathcal{U}_n(\theta)$. Let $\mathcal{K}(w_1, w_2)$ denote the event that $w_2$ is decoded when $w_1$ is transmitted. We can decompose the error probability as follows:

$$\mathbb{P}(\mathcal{K}(w_1, w_2)) = \mathbb{P}(\mathcal{K}(w_1, w_2) | x(w_2) \in \mathcal{U}_n(\theta)) \mathbb{P}(x(w_2) \in \mathcal{U}_n(\theta))$$

$$+ \mathbb{P}(\mathcal{K}(w_1, w_2) | x(w_2) \notin \mathcal{U}_n(\theta)) \mathbb{P}(x(w_2) \notin \mathcal{U}_n(\theta)).$$

Rearranging terms, we can upper bound the conditional error probability for when both codewords lie in the spherical cap,

$$\mathbb{P}(\mathcal{K}(w_1, w_2) | x(w_2) \in \mathcal{U}_n(\theta)) \leq \frac{\mathbb{P}(\mathcal{K}(w_1, w_2))}{\mathbb{P}(x(w_2) \in \mathcal{U}_n(\theta))}$$

$$= \frac{\mathbb{P}(\mathcal{K}(w_1, w_2))}{2^{-n(C-\epsilon)}} \leq 2^{-n(C-\frac{\epsilon}{2})} + 1 \leq 2^{-n(C-\frac{\epsilon}{2})}.$$  \hspace{1cm} (23)

Now, consider the (randomly chosen) set of $2^{nR}$ codewords of $\mathcal{C}_{\text{sphere}}$ that lie in the spherical cap that make up the codewords $x(1), \ldots, x(2^{nR})$ of $\mathcal{C}$. First, we need to guarantee that the spherical cap contains enough codewords. By Lemma 1, the probability that the cap contains fewer than $2^{nR}$ codewords is upper bounded by $\exp(-2^{nR-2})$. From the argument above, for maximum likelihood decoding (ignoring the red alert codeword), the probability that we decode $x(w_2)$ when $x(w_1), w_1 \neq w_2$, was sent is at most $2^{-n(R+\frac{\epsilon}{2})}$. Therefore, via the union bound, the probability that $x(w_1)$ is incorrectly decoded is upper bounded as follows:

$$\mathbb{P}(\hat{w} \neq w_1) = \sum_{w_2 \neq w_1} \mathbb{P}(\mathcal{K}(w_1, w_2))$$

$$\leq 2^{nR} 2^{-n(R+\frac{\epsilon}{2})} = 2^{-n\epsilon/2}$$  \hspace{1cm} (26)

Since this bound holds over all codewords in $\mathcal{C}$, the total expected error probability can be made arbitrarily small for $n$ large enough.

The following lemma establishes a relationship between the half-angle $\theta$ and the fraction of the surface area of the sphere occupied by the spherical cap.

**Lemma 3 (Shannon):** The fraction of surface area of an $n$-dimensional sphere occupied by a spherical cap of half-angle $\theta$ satisfies

$$\frac{\text{Vol}(\mathcal{U}_n(\theta))}{\text{Vol}(S_n(0, \sqrt{nR_{\text{avg}}}))} = \frac{\sin^n \theta}{\sqrt{2\pi n \sin \theta \cos \theta}} \left( 1 + O \left( \frac{1}{n} \right) \right).$$

See the math leading up to equation (28) in [6] for a proof.

**Corollary 1:** Assume that $\theta$ is chosen such that the fraction of surface area of the sphere $S_n(0, \sqrt{nR_{\text{avg}}})$ occupied by the spherical cap $\mathcal{U}_n(\theta)$ satisfies

$$\frac{\text{Vol}(\mathcal{U}_n(\theta))}{\text{Vol}(S_n(0, \sqrt{nR_{\text{avg}}}))} = 2^{-n\lambda}.$$  \hspace{1cm} (27)

for some fixed $\lambda > 0$. Then for any $\epsilon > 0$ and $n$ large enough, we have that

$$\sin \theta \leq (1 + \epsilon) 2^{-\lambda}.$$  \hspace{1cm} (28)

**Proof:** Using Lemma 3, it follows that

$$\frac{\sin^n \theta}{\sqrt{2\pi n \sin \theta \cos \theta}} \left( 1 + O \left( \frac{1}{n} \right) \right) = 2^{-n\lambda}.$$  \hspace{1cm} (29)
Rearranging terms we get
\[ \sin^n \theta = \frac{2^{-n\lambda} \sqrt{2\pi n \sin \theta \cos \theta}}{(1 + O(\frac{1}{n}))} < 2^{-n\lambda} \sqrt{2\pi n} . \] (31)
Taking the \( n \)th root yields the desired result for \( n \) large enough.

IV. ERROR ANALYSIS

We will demonstrate that the decoding region for the standard codewords can be restricted to the intersection of a cone with a shell determined by the average power constraint and the noise variance. Then, we allocate the remainder of \( \mathbb{R}^n \) to the red alert decoding region and bound the resulting probability of a missed detection. Our upper bound splits the red alert missed detection event into two parts. First, the observation must fall inside the cone centered at the red alert decoding region and bound the resulting probability of a missed detection. Second, the magnitude of the noise must be large enough to reach the decoding region of the standard codewords.

A. Standard Decoding Region

The following lemma formalizes the notion that the squared \( \ell_2 \) norm of an i.i.d. Gaussian vector concentrates sharply around its variance.

**Lemma 4:** Let \( z \) be a length-\( n \) vector with i.i.d. zero-mean Gaussian entries with variance \( N \). Then, for any \( \beta > 0 \),
\[ \mathbb{P}(\|z\|^2 \geq nN(1 + \beta)) \leq \exp \left( -\frac{n}{2} (\beta - \ln(1 + \beta)) \right) \] (32)
and, for any \( 0 < \beta < 1 \),
\[ \mathbb{P}(\|z\|^2 \leq nN(1 - \beta)) \leq \exp \left( -\frac{n}{2} (\beta - \ln(1 - \beta)) \right) . \] (33)
See, for instance, Proposition 2.2 in [7] for a proof.

Therefore, for large \( n \), the decoding region for \( C_{\text{sphere}} \) can be restricted to a thin spherical shell.

The next lemma well-known that an i.i.d. Gaussian vector is approximately orthogonal with respect to any fixed vector.

**Lemma 5:** Let \( z \) be a length-\( n \) vector with i.i.d. zero-mean Gaussian entries with variance \( N \) and let \( a \) be a length-\( n \) vector with \( \|a\|^2 = na \) for some fixed \( \alpha > 0 \). Then, for any \( \epsilon > 0 \) and \( n \) large enough
\[ \mathbb{P}(\|a^T z\| \geq \epsilon \|a\|^2) < \epsilon . \] (34)

**Proof:** First, we write the probability that \( |a^T z| \) is greater than \( t \) in terms of the Q-function,
\[ \mathbb{P}(\|a^T z\| \geq t) = 2Q \left( \frac{t}{\sqrt{N\|a\|}} \right) \] (35)
\[ < \frac{\|a\|}{t} \sqrt{\frac{2N}{\pi}} \exp \left( -\frac{t^2}{2N\|a\|^2} \right) . \] (36)
Substituting \( t = \epsilon \|a\|^2 \) yields,
\[ \mathbb{P}(\|a^T z\| \geq \epsilon \|a\|) < \sqrt{\frac{2N}{\pi \epsilon^2 n\alpha}} \exp \left( -\frac{\epsilon^2 n\alpha}{2N} \right) . \] (37)
which can be driven arbitrarily close to zero for \( n \) large enough.

**Lemma 6:** Let \( z \) be a length-\( n \) vector with i.i.d. zero-mean Gaussian entries with variance \( N \) and let \( a \) be a length-\( n \) vector with \( \|a\|^2 = na \) for some fixed \( \alpha > 0 \). For any \( \nu > 0 \) and \( n \) large enough, the probability that the angle between \( a \) and \( a + z \) exceeds \( \cos^{-1} \left( \frac{\alpha}{\alpha + N} \right) + \nu \) is upper bounded by \( \nu \).
\[ \mathbb{P}(\angle(a, a + z) \geq \cos^{-1} \left( \frac{\alpha}{\alpha + N} \right) + \nu) < \nu . \] (38)

**Proof:** The angle between \( a \) and \( a + z \) is
\[ \angle(a, a + z) = \cos^{-1} \left( \frac{a^T(a + z)}{\|a\|\|z\|} \right) . \] (39)
From Lemma 5, for any \( \epsilon > 0 \) and \( n \) large enough, the probability that \( a^T z > \epsilon \|a\|^2 \) is at most \( \epsilon \). Therefore, \( a^T(a + z) \geq (1 + \epsilon)na \) with probability at most \( \epsilon \). Combining Lemmas 5 and 4, we can also show that the probability that \( \|a + z\| < (1 - \epsilon)\sqrt{\alpha/\alpha + N} \) is at most \( \epsilon \) for \( n \) large enough. Thus, the probability that
\[ \frac{a^T(a + z)}{\|a\|\|z\|} < \frac{(1 + \epsilon)na}{\sqrt{n\epsilon(1 - \epsilon)\sqrt{\alpha/\alpha + N}}} \] (40)
\[ = \frac{1 + \epsilon}{1 - \epsilon} \sqrt{\frac{\alpha}{\alpha + N}} \] (41)
is at most \( 2\epsilon \). Choosing \( \epsilon \) small enough yields the desired result.

We now have all the tools we need to show that the decoding region of our “conical” codebook can be constrained to a spherical cap of the noise-inflated shell. In the next lemma, we will establish that there exists at least one codebook of rate \( R \) with average probability of message error \( p_{\text{MSG}} \) at most \( \epsilon \) when the decoding region is taken to be this spherical cap.

\[ \theta + \phi + \epsilon \]
\[ \sqrt{n P_{\text{seg}}} \]
\[ \sqrt{n P_{\text{seg}}} + N - \epsilon \]
Lemma 7: Choose $R < C$ and $\epsilon > 0$. Let $\theta = \sin^{-1}(2^{R-C}+\epsilon)$ and $\phi = \sin^{-1}(2^{-C})$. For $n$ large enough, there exists a codebook with $2^{nR}$ codewords, all in the spherical cap $U_n(\theta)$, and average probability of error $P_{\text{MSG}} < \epsilon$ whose decoding region for the standard messages $L_{\text{MSG}}$ is restricted to the intersection of the cone with half-angle $\theta + \phi + \epsilon$ and the spherical shell from radius $\sqrt{n(P_{\text{avg}} + N - \epsilon)}$ to radius $\sqrt{n(P_{\text{avg}} + N + \epsilon)}$.

$L_{\text{MSG}} \subseteq \mathcal{V}(0, -1, \theta + \phi + \epsilon) \cap \cdots \cap T_n(0, \sqrt{n(P_{\text{avg}} + N - \epsilon)}, \sqrt{n(P_{\text{avg}} + N + \epsilon)}) \quad (42)$

**Proof:** We will use the ensemble of random codebooks developed in Section III. We characterize three error events:

1) First, we bound the average error probability of a random codebook.

2) Second, we bound the probability that the received signal lies outside the cone $\mathcal{V}(0, -1, \theta + \phi + \epsilon)$. This happens only if the magnitude of the noise exceeds $L$ and has magnitude exceeding $\epsilon$. The probability at most $\epsilon/3$ for $n$ large enough if we use maximum-likelihood decoding for the standard messages over $\mathbb{R}^n$, ignoring the red alert codeword. We now show that the error probability only increases by a small amount if we restrict the decoding regions to the spherical cap.

The received signal $y = x + z$ is most likely to end up outside the cone if we assume the transmitted codewords lies on the edge of the spherical cap $U(\theta)$. In this case, $y$ lies outside the cone if the angle between $x$ and $y$ exceeds $\phi + \epsilon$.

From Lemma 6, we have that for any $\nu > 0$ and $n$ large enough,

$$\mathbb{P}\left(\angle(x, y) \geq \cos^{-1}\left(\frac{P_{\text{avg}}}{P_{\text{avg}} + N}\right) + \nu\right) < \nu \quad . \quad (43)$$

Using the trigonometric identity $\sin(\cos^{-1}(a)) = \sqrt{1 - a^2}$ and the capacity formula $2^C = \sqrt{P_{\text{avg}} N}$, we can show

$$\sin\left(\cos^{-1}\left(\frac{P_{\text{avg}}}{P_{\text{avg}} + N}\right)\right) = \sqrt{1 - \frac{P_{\text{avg}}}{P_{\text{avg}} + N}} = \frac{N}{P_{\text{avg}} + N} = 2^{-C} \quad , \quad (45)$$

so that $\cos^{-1}\left(\frac{P_{\text{avg}}}{P_{\text{avg}} + N}\right) = \sin^{-1}(2^{-C}) = \phi$. Now, set $\nu = \epsilon/3$ to get that the received vector lies outside the cone of half-angle $\theta + \phi + \epsilon/3$ with probability at most $\epsilon/3$.

Finally, using Lemma 4, it can be shown that the received vector lies in the spherical shell with probability at most $\epsilon/3$. Using the union bound, we get that the average error probability over the decoding region constrained to the intersection of the cone and the shell is at most $\epsilon$. It then follows that there must exist at least one good codebook with average error probability at most $\epsilon$ which completes the proof.

**B. Red Alert Decoding Region**

Now that we know the decoding region can be confined to a conical section of the shell, we can bound the probability for the red alert codeword. We break our analysis into two cases based on whether the half-angle of the standard decoding region is less than or greater than $\pi/2$. Recall from Lemma 7 that the half-angle (neglecting $\epsilon$ terms) is $\sin^{-1}(2^{R-C}) + \sin^{-1}(2^{-C})$. We now verify that this half-angle is greater than $\pi/2$ if $R \geq \max (\frac{1}{2} \log_2 \left(\frac{P_{\text{avg}}}{n}\right), 0)$. We will need the trigonometric identity $\sin^{-1}(x) + \sin^{-1}(y) = \sin^{-1}(x\sqrt{1 - y^2} + y\sqrt{1 - x^2})$. First, consider the case where $\frac{1}{2} \log_2 \left(\frac{P_{\text{avg}}}{n}\right)$ is positive.

$$\sin^{-1}(2^{R-C}) + \sin^{-1}(2^{-C}) \quad (46)$$
$$= \sin^{-1}\left(2^{R-C} \sqrt{1 - 2^{-2C} + 2^{-C} \sqrt{1 - 2^{2R-2C}}\right) \quad (47)$$
$$\geq \sin^{-1}\left(\frac{P_{\text{avg}}}{P_{\text{avg}} + N} + \frac{N}{P_{\text{avg}} + N}\right) = \sin^{-1}(1) = \pi/2 \quad (48)$$

Now, if $\frac{1}{2} \log_2 \left(\frac{P_{\text{avg}}}{n}\right)$ is negative, we get that the capacity is less than one, $C \leq 1$. From here, we immediately get that $\sin^{-1}(2^{-C}) \geq \pi/4$ and $\sin^{-1}(2^{R-C}) \geq \pi/4$, which implies the angle is at least $\pi/2$.

1) Low Rate Case: We will decompose the error into two events as illustrated in Figure 3. An error occurs if the noise vector carries the transmitted red alert codeword into the standard decoding region. This happens only if the magnitude of the noise exceeds $L$ and the noise lies in the cone of half-angle $\psi$. These quantities can be calculated using simple trigonometry.

![Diagram](image-url)

**Fig. 3.** Standard codeword decoding region as seen from the red alert codeword for the case $\theta + \phi + \epsilon \leq \pi/2$. Note that the standard decoding lies between distances $\sqrt{n(P_{\text{avg}} + N - \epsilon)}$ and $\sqrt{n(P_{\text{avg}} + N + \epsilon)}$ from the origin. An error occurs if the noise pushes the received vector from the red alert codeword (square) to the standard decoding region (shaded region). We bound this event by noting that an error will only occur if the noise falls in the cone of half-angle $\psi$ and has magnitude exceeding $L$.

First, the minimum distance $L$ from the red alert codeword
to the standard decoding region satisfies
\[
L^2 = \left( \sqrt{n P_{\text{peak}}} + \sqrt{n (P_{\text{avg}} + N - \epsilon)} \cos(\theta + \phi + \epsilon) \right)^2 + \left( \sqrt{n (P_{\text{avg}} + N - \epsilon)} \sin(\theta + \phi + \epsilon) \right)^2
\]
\[
= n \left( P_{\text{avg}} + P_{\text{peak}} + N - \epsilon \right) + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N - \epsilon)} \cos(\theta + \phi + \epsilon) .
\]
(49)

Again, by choosing \( \epsilon \) small as we would like, the following lower bound on \( L^2 \) holds for any \( \nu > 0 \) and \( n \) large enough:
\[
L^2 \geq n \left( P_{\text{avg}} + P_{\text{peak}} + N + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa} \right) .
\]
(50)

Next, we choose \( \psi \) so that the cone centered at the red alert codeword with half-angle \( \psi \) contains the standard decoding region. The minimum angle required satisfies
\[
\sin \psi \leq \frac{\sqrt{n(N + P_{\text{avg}} + \epsilon) \sin \kappa}}{L}
\]
(52)

Again, by choosing \( \epsilon \) small enough and \( n \) large enough, the following upper bound on \( \sin \phi \) holds for any \( \nu > 0 \):
\[
\sin \psi \leq \frac{1 + \nu \sqrt{N + P_{\text{avg}} \sin \kappa}}{\sqrt{P_{\text{avg}} + P_{\text{peak}} + N + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa}}} .
\]
(53)

Using the bounds above, we now derive the red alert exponent for low rates.

Lemma 8: If the rate \( R \) is less than max \( \left( \frac{1}{2} \log_2 \left( \frac{P_{\text{avg}}}{\mu_{\text{TAN}}} \right) , 0 \right) \), a red alert exponent of
\[
\beta = \frac{P_{\text{avg}} + P_{\text{peak}} + 2 \cos \kappa \sqrt{(P_{\text{avg}} + N) P_{\text{peak}}}}{N}
\]
(54)

\[
E_{\text{ALERT}}(R) = \frac{\beta}{2} - C \ln 2 - \ln(\sin \kappa)
\]
(55)

is achievable.

Proof: There is an error if the noise lies in the cone of half-angle \( \phi \) and has magnitude larger than \( L \). Using Lemma 4, we can show that for any \( \nu > 0 \) and \( n \) large enough, the probability that the magnitude of the noise vector \( \mathbf{z} \) exceeds \( L \) is upper bounded by
\[
\mathbb{P}(\|\mathbf{z}\| \geq L) \leq \exp \left( -\frac{n}{2} \left( \beta - \ln(1 + \beta) \right) - \nu \right) .
\]
(56)

The probability that the received vector falls into the cone of half-angle \( \psi \) is given by the fraction of surface area of a sphere carved out by the cone. Using Lemma 3, this can be calculated as
\[
\mathbb{P} \left( \mathbf{x}(0) + \mathbf{z} \in V_n \left( \sqrt{n P_{\text{avg}} \mathbf{1}} , -1 , \psi \right) \right) = \frac{\sin^n \psi}{\sqrt{2 \pi n} \sin \psi \cos \psi} \left( 1 + O \left( \frac{1}{n} \right) \right) .
\]
(57)

Pulling terms into the exponent we get
\[
\exp \left( -n \left( -\ln(\sin \psi) \right) \right) \left( \cdots + \frac{1}{n} \ln \left( \sqrt{2 \pi n} \sin \psi \cos \psi \right) + O \left( 1/n \right) \right)
\]
(59)

and for any \( \nu > 0 \) and \( n \) large enough we get that the probability is upper bounded by \( \exp \left( -n \left( -\ln(\sin \psi) - \nu \right) \right) \). Since the noise is an i.i.d. Gaussian vector, its magnitude and direction are independent. Therefore, if the red alert code word is transmitted, the probability that the received vector lies in the decoding region \( L_{\text{MSG}} \) for the standard messages is upper bounded by
\[
\mathbb{P} \left( \mathbf{x}(0) + \mathbf{z} \in L_{\text{MSG}} \right) \leq \mathbb{P}(\|\mathbf{z}\| \geq L) \mathbb{P} \left( \left( \sqrt{n P_{\text{avg}} 1} , -1 , \psi \right) \right) \leq \exp \left( -n \left( \frac{P_{\text{avg}} + P_{\text{peak}} + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa}}{2N} \right) \right)
\]
(60)

Since \( \nu \) can be chosen as small as desired (for \( n \) large enough), the proof is complete.

2) High Rate Case: We now examine the case where \( R \geq \max \left( \frac{1}{2} \log_2 \left( \frac{P_{\text{avg}}}{\mu_{\text{TAN}}} \right) , 0 \right) \). We can still use a cone centered at the red alert codeword to capture the standard decoding region. At first, if we take the cone to intersect the frontier of the standard decoding region, the half-angle \( \phi \) continues to increase with rate and the distance \( L \) continues to decrease. However, the cone will eventually become tangent to the sphere of radius \( \sqrt{n (P_{\text{avg}} + N + \epsilon)} \) and afterwards the half-angle to the frontier will decrease. After this threshold, we can no longer use the frontier to find the angle of our cone as this will cut through the standard decoding region. Therefore, once the enclosing cone is tangent to the sphere, we keep the half-angle fixed and only allow the distance to decrease with rate. See Figure 4 for an illustration.

Some simple trigonometry reveals that when the enclosing cone is tangent to the sphere,
\[
-\sqrt{n (P_{\text{avg}} + N + \epsilon)} = \sqrt{n P_{\text{peak}} \cos \mu}
\]
(61)

\[
\mu_{\text{TAN}} = \pi - \cos^{-1} \left( \frac{P_{\text{avg}} + N}{P_{\text{peak}}} \right) .
\]
(62)

Note that this implicitly assumes that we can place the red alert codeword outside the sphere which is only possible if the peak power exceeds the average power plus the noise variance, \( P_{\text{peak}} > P_{\text{avg}} + N \).

Assuming that \( \kappa \leq \mu_{\text{TAN}} \), we are free to take the cone to frontier of the decoding region as in the low rate case. The distance and angle bounds are the same as in the low rate case.
The probability that the noise lies within the cone of half-angle $\theta$ received vector from the red alert codeword (square) to the standard decoding region (shaded region). The half-angle $\theta$ of the enclosing cone increases with $\theta + \phi + \epsilon$ until it becomes tangent to the sphere and then stays fixed. The distance to the red alert codeword $L$ continues to decrease until the standard decoding region becomes a full sphere.

**Lemma 9:** If $P_{\text{peak}} > P_{\text{avg}} + N$ and $\kappa \leq \mu_{\text{TAN}}$, the following red alert exponent is achievable:

$$\beta = \frac{P_{\text{avg}} + P_{\text{peak}} + 2 \cos \kappa \sqrt{(P_{\text{avg}} + N) P_{\text{peak}}}}{N} \quad (63)$$

$$\beta^+ = \max(0, \beta) \quad (64)$$

$$E_{\text{ALERT}}(R) = \frac{\beta^+}{2} - \frac{1}{2} \ln(1 + \beta^+) - \ln(\sin \kappa) \cdots \quad (65)$$

$$\cdots + \frac{1}{2} \ln \left( 1 + \frac{P_{\text{peak}} + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa}}{N + P_{\text{avg}}} \right) \quad (66)$$

**Proof:** We would like to bound the probability that the noise lies in the cone of half-angle $\phi$ and has magnitude larger than $L$. If $L^2$ exceeds the noise variance, we get a positive error exponent due to the magnitude. Using Lemma 4, we get that

$$\mathbb{P}(\|z\| \geq L) \leq \exp \left( -n \left( \frac{\beta^+}{2} - \frac{1}{2} \ln(1 + \beta^+) - \nu \right) \right). \quad (67)$$

The probability that the noise lies within the cone of half-angle $\phi$ is upper bounded just as in the proof of Lemma 8,

$$\mathbb{P}\left( z(0) + z \in V_n \left( \sqrt{n} P_{\text{avg}} \mathbf{1}, -1, \psi \right) \right) \leq \exp \left( -n \left( \frac{1}{2} \ln \left( 1 + \frac{P_{\text{peak}} + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa}}{N + P_{\text{avg}}} \right) \cdots \ln(\sin \kappa) - \ln(1 + \nu) \right) \right). \quad (68)$$

Combining these two terms yields the desired result.

Once the entire sphere is enclosed, we can hold the half-angle of the cone fixed at $\psi = \frac{\pi}{2} - \cos^{-1}(\sqrt{P_{\text{peak}} / (P_{\text{avg}} + N + \epsilon)})$. The distance $L$ will continue to decrease with the frontier of the decoding region. As before, for any $\nu > 0$ and $n$ large enough,

$$L^2 \geq n \left( P_{\text{avg}} + P_{\text{peak}} + N + 2 \sqrt{P_{\text{peak}} (P_{\text{avg}} + N) \cos \kappa - \nu} \right)$$

**Lemma 10:** If $P_{\text{peak}} > P_{\text{avg}} + N$ and $\kappa \geq \mu_{\text{TAN}}$, the following red alert exponent is achievable:

$$\beta = \frac{P_{\text{avg}} + P_{\text{peak}} + 2 \cos \kappa \sqrt{(P_{\text{avg}} + N) P_{\text{peak}}}}{N} \quad (69)$$

$$\beta^+ = \max(0, \beta) \quad (70)$$

$$E_{\text{ALERT}}(R) = \frac{\beta^+}{2} - \frac{1}{2} \ln(1 + \beta^+) + \frac{1}{2} \ln \left( \frac{P_{\text{peak}}}{P_{\text{avg}} + N} \right)$$

**Proof:** The probability of the magnitude exceeding $L$ is bounded just as in Lemma 9. For the angle, the probability is also bounded as in Lemma 9, except that the angle is set to $\mu_{\text{TAN}}$ instead of $\kappa$. We get that

$$\cos(\mu_{\text{TAN}}) = \sqrt{\frac{P_{\text{avg}} + N}{P_{\text{peak}}}} \quad (70)$$

$$\sin(\mu_{\text{TAN}}) = \sqrt{1 - \frac{P_{\text{avg}} + N}{P_{\text{peak}}}} \quad (71)$$

Substituting these into the error exponent expression for the angle, we get

$$\frac{1}{2} \ln \left( \frac{P_{\text{peak}} - P_{\text{avg}} - N}{P_{\text{avg}} + N} \right) - \frac{1}{2} \ln \left( \frac{P_{\text{peak}} - P_{\text{avg}} + N}{P_{\text{peak}}} \right)$$

$$= \frac{1}{2} \ln \left( \frac{P_{\text{peak}}}{P_{\text{avg}} + N} \right) \quad (72)$$

Combining terms for the magnitude and angle completes the proof.

3) Low SNR Case: Consider a scenario where the noise variance far exceeds both the average and peak power constraints. In this setting, it is often beneficial to place the red alert codeword at the origin rather than at the boundary of the peak power constraint. This increases the distance between the red alert codeword and the standard decoding region as illustrated in Figure 5. Notice that if the half-angle $\kappa$ that defines the standard decoding region is less than $\pi/2$, it is always better to place the red alert codeword as far away as possible. Therefore, we can assume that if red alert codeword is at the origin, $\kappa > \pi/2$. As a result, the probability that the noise lands in the direction of the standard decoding region is at least $1/2$ and there is no error exponent associated with the angle of the noise. The only contribution to the error decay is the distance between the red alert codeword and the standard decoding region which is at least $\sqrt{n}(P_{\text{avg}} + N - \epsilon)$.

**Lemma 11:** The following red alert exponent is achievable

$$E_{\text{ALERT}}(R) = \frac{P_{\text{avg}}}{2N} - C \ln 2 \quad (73)$$
This way, the noise magnitude must be larger than the noise variance, it is best to place the red alert codeword at the origin.

Fig. 5. When both the average and peak power constraints are small compared to the noise variance, it is best to place the red alert codeword at the origin. If we place the red alert codeword at the boundary of the peak power constraint, the distance to the standard decoding region may be much smaller. Since the probability that the noise lands in the direction of the standard decoding region is greater than $1/2$, there is no error exponent from the angular component.

**Proof:** We just need to bound the probability that the magnitude of the noise $z$ exceeds $\sqrt{n(P_{\text{avg}} + N - \epsilon)}$. For any $\epsilon > 0$ and $n$ large enough, this probability is upper bounded by

$$\exp\left(-\frac{n}{2} P_{\text{avg}} - \ln\left(\frac{N + P_{\text{avg}}}{N}\right) - \epsilon\right) \quad (74)$$

$$= \exp\left(-n \left(\frac{P_{\text{avg}}}{2N} - C \ln 2 - \epsilon\right)\right). \quad (75)$$

via Lemma 4.

4) **Red Alert Exponent:** Combining these three cases, we arrive at our main result.

**Theorem 1:** Let $\kappa = \sin^{-1}\left(2^{R - C}\right) + \sin^{-1}\left(2^{C}\right)$ and let $\mu_{\text{TAN}} = \pi - \cos^{-1}\left(\sqrt{\frac{P_{\text{peak}}}{P_{\text{avg}} + N}}\right)$. Define

$$\beta = \frac{P_{\text{avg}} + P_{\text{peak}} + 2 \cos \kappa \sqrt{(P_{\text{avg}} + N)P_{\text{peak}}}}{N}$$

$$\beta^+ = \max(0, \beta)$$

$$E_1(R) = \frac{\beta}{2} - C \ln 2 - \ln(\sin \kappa)$$

$$E_2(R) = \frac{\beta^+}{2} - \frac{1}{2} \ln(1 + \beta^+) - \ln(\sin \kappa) \cdots$$

$$\cdots + \frac{1}{2} \ln\left(1 + \sqrt{\frac{P_{\text{peak}} + 2 \sqrt{P_{\text{peak}}(P_{\text{avg}} + N)} \cos \kappa}{N + P_{\text{avg}}} + \frac{P_{\text{peak}}}{P_{\text{avg}} + N}\right) \quad (78)$$

$$E_3(R) = \frac{\beta^+}{2} - \frac{1}{2} \ln(1 + \beta^+) + \frac{1}{2} \ln\left(\frac{P_{\text{peak}}}{P_{\text{avg}} + N}\right) \quad (79)$$

$$E_{\text{LOW}}(R) = \frac{P_{\text{avg}}}{2N} - C \ln 2. \quad (80)$$

The following red alert exponent is achievable:

$$E_{\text{ALERT}}(R) = \begin{cases} E_1(R) & \kappa < \pi/2 \\ \max\left(E_2(R), E_{\text{LOW}}(R)\right) & P_{\text{peak}} > P_{\text{avg}} + N, \kappa \leq \mu_{\text{TAN}} \\ \max\left(E_3(R), E_{\text{LOW}}(R)\right) & P_{\text{peak}} > P_{\text{avg}} + N, \kappa > \mu_{\text{TAN}} \end{cases}$$

$$\quad \text{otherwise.} \quad (81)$$

**Proof:** $E_1(R)$ is due to Lemma 8 which applies if $\kappa < \pi/2$ or, alternatively, $R \leq \max\left(0, \frac{1}{2} \log_2\left(\frac{P_{\text{avg}}}{N}\right)\right)$. Lemmas 9 and 10 apply when $P_{\text{peak}} > P_{\text{avg}} + N$ and establish $E_2(R)$ and $E_3(R)$, respectively, as a special case. Finally, $E_{\text{LOW}}(R)$ from Lemma 11 can be applied in any scenario (although it is always less than $E_1(R)$).

![Fig. 6. Red alert error exponent with an average power constraint of 5, 3, and 0 dB. The peak power constraint is a factor of 10 higher than the average constraint.](http://www.math.lsa.umich.edu/barvinok/total710.pdf)

**REFERENCES**


