On Approximation, Bounding & Exact Calculation of Block Error Probability for Random Codes

Ralf R. Müller*

*Institute for Digital Communications, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany
ralf.r.mueller@fau.de

Abstract—This paper presents a method to calculate the exact average block error probability of some random code ensembles under maximum-likelihood decoding. Deviating from Shannon’s 1959 solid angle argument, we project the problem into two dimensions and apply standard trigonometry. This enables us to also analyze Gaussian random codes in additive white Gaussian noise and binary random codes for the binary symmetric channel. We find that the Voronoi regions harden doubly-exponential in the blocklength and utilize that to propose the new median bound that outperforms Shannon’s 1959 sphere packing bound for the uniform spherical ensemble, whenever the code contains more than three codewords. Furthermore, we propose a very tight approximation to simplify computation of both exact error probability and the two bounds.

I. INTRODUCTION

In his 1959 seminal paper, Shannon [1] calculates the exact block error probability (BEP) of the uniform spherical random code ensemble in additive white Gaussian noise (AWGN). However, the exact formula turned out to be hopeless to evaluate numerically given the computing tools of that time. Still he could find accurate bounds on the exact error probability. His lower bound is based on an argument resorting to fill the entire $N$-dimensional hypersphere with hyperspheres of smaller radii and commonly referred to as the 1959 sphere packing bound (SPB). His bounds were not easy to evaluate back in the 1960s [2]. Subsequent works focussed on applying the sphere packing idea to a broader range of codes and channels, see [3] for a tutorial. Improved computing power has made the SPB a very fast and useful tool for checking code performance [4]. The ideas proposed in this work are not related to more recent bounds on average BEP based on the normal approximation [5].

The paper is organized as follows: In Section II, we propose a method to calculate average BEP and demonstrate its usefulness at the example of the uniform spherical random code ensemble in AWGN. In Section III, we propose the median bound and compare it to Shannon’s 1959 SPB. Sections IV and V apply the proposed method to Gaussian random coding in AWGN and binary random coding for the BSC, respectively. Section VI summarizes the conclusions.

II. CALCULATION OF BLOCK ERROR PROBABILITY

Consider an AWGN channel. The random codebook of size $2^{NR}$ is chosen from points uniformly distributed on an $N$-dimensional hypersphere of radius $\sqrt{NP}$. Consider Fig. 1. The transmitted codeword is denoted by $c$. It is distorted by independent AWGN of zero mean and variance 1 denoted by $z$. We decompose the noise vector into a radial component $z_r$ which is collinear to the code word $c$ and a tangential component $z_t$ which is orthogonal to the code word $c$. We denote the angle between the codeword $c$ and the received word $r$ as $\beta$ and the angle between an alternative codeword $w_i$ and the received word as $\alpha_i$. 
A. Angle to Nearest Wrong Codeword

All codewords are uniformly distributed on the hypersphere. Their joint distribution is thus invariant to any rotation around the origin. We utilize this property and rotate the hypersphere in such a way that the received word becomes collinear to the first unit vector of a Cartesian coordinate system.

Denoting \( \rho_i = \cos \alpha_i \), we can construct the squared correlation coefficient between the received word and the \( i \)th codeword out of \( N \) independent identically Gaussian distributed random variables \( g_n \), with zero mean and unit variance as

\[
\rho_i^2 = g_i^2 / \sum_{n=1}^{N} g_n^2.
\]

In the numerator, only the first Gaussian random variable shows up due to the inner product with the first unit vector of the coordinate system. The ratio in (1) is known to be distributed according to the beta distribution with shape parameters \( \frac{1}{2} \) and \( \frac{N-1}{2} \). The corresponding density is given by [6]

\[
p_{\rho} (r) = \frac{1}{B\left(\frac{1}{2}, \frac{N-1}{2}\right)} r^{\frac{1}{2}} (1-r)^{\frac{N-3}{2}}.
\]

with \( B(\cdot, \cdot) \) denoting the beta function. Substituting \( \rho = \sqrt{r} \) and utilizing the symmetry around \( \rho = 0 \), leads to the density

\[
p_{\rho} (\rho) = \frac{1}{B\left(\frac{1}{2}, \frac{N-1}{2}\right)} (1-\rho^2)^{\frac{N-3}{2}}
\]

within \([-1; +1]\).

Given the received word \( r \), all angles \( \alpha_i \) are jointly statistically independent due to the spherical code construction, i.e. for any disjoint sets of angles \( A_1 \) and \( A_2 \), \( A_1 \mapsto r \mapsto A_2 \) is a Markov chain. This can easily be verified recognizing that no information about an angle \( \alpha_i \) is revealed, if we know \( r \) and \( \alpha_i \) with \( i \neq j \). It is still uniformly distributed on the hypersphere. Furthermore, \( p_{\rho|\tau} (\cdot) \) due to rotational invariance. Since there are \( 2^{NR} - 1 \) alternative codewords, the cumulative distribution function of \( \rho = \max_i \rho_i \) becomes

\[
P_{\rho|r} (\rho) = \left[ \int_{0}^{\rho} p_{\rho} (\rho) d\rho \right]^{2^{NR}-1}.
\]

Defining the lower regularized incomplete beta function

\[
B(a,b,x) = \frac{1}{B(a,b)} \int_{0}^{x} \xi^{a-1} (1-\xi)^{b-1} d\xi
\]

which is readily available in Matlab as \texttt{betainc(x,a,b)}, and substituting \( x = \rho^2 \), we get

\[
P_{\rho|r} (\rho) = \left[ \frac{1}{2} + \frac{\text{sign}(\rho)}{2} B\left(\frac{1}{2}, \frac{N-1}{2}, \rho^2\right) \right]^{2^{NR}-1}.
\]

The numerical evaluation of (6) is not straightforward if \( NR \) exceeds values around 40. The bracketed expression is a cumulative distribution function which for a wide range of \( \rho \) is very close to 1. Raised to a huge power like \( 2^{NR} \), tiny deviations from unity have high impact. This issue can be fixed as follows: Note from its definition (5) that the incomplete beta function obeys the following intrarelationship

\[
B(a,b,x) = 1 - B(b, a, 1-x).
\]

So if \( B(a,b,x) \) is close to one and causes numerical trouble, \( B(b,a,1-x) \) is close to zero and can be calculated very accurately in floating point arithmetic. Furthermore, we have the series expansion

\[
(1-x)^a = e^a \ln(1-x) = \sum_{i=1}^{\infty} e^{-\frac{a}{i}}
\]

converging for \( x \in (0; 1) \). In practice, we will hardly need more than a single factor, as we only need this expansion, if \( x \) is so tiny that direct evaluation of the left hand side causes numerical issues. Utilizing (7) to (8), we find

\[
P_{\rho|r} (\rho) = \prod_{i=1}^{\infty} e^{(1-x^{2^i})} \left[ \frac{\text{sign}_{\rho_i}}{2} B\left(\frac{1}{2}, \frac{N-1}{2}, \rho^2\right) \right]^{2^{NR}-1}
\]

with \( 1_{x<0} \) equaling 1 and 0 for \( x < 0 \) and \( x \geq 0 \), respectively. If (6) runs into numerical trouble, (9) will not, and vice versa.

There is still one issue to be resolved. If \( NR \) reaches 1024 or beyond, \( 2^{NR} \) is numerically represented as \( \texttt{Inf} \) while the incomplete beta function might be numerically 0. This would make their product a numerical exemption, i.e. \( \texttt{NaN} \). This issue is overcome by the upper bound (proven in the extended journal version [7])

\[
B\left(\frac{a}{2}, \frac{1}{2}, x\right) \leq \frac{2^{a} B\left(\frac{a}{2}, \frac{1}{2}, 1-x\right)}{B\left(\frac{1}{2}, \frac{N-1}{2}\right)}
\]

with equality for \( x = 0 \), that is very tight for large \( N \) and \( x \) not too close to unity. The resulting product on the right hand side of

\[
2^{NR} B\left(\frac{N-1}{2}, \frac{1}{2}, 1-\rho^2\right) \leq 2^{NR} \left[ B\left(\frac{a}{2}, \frac{1}{2}, 1-x\right) \right]^\frac{N-1}{2}
\]

does not risk to run into such numerical exemptions. Note that, for large \( NR \), the sign and the truth functions in (9), as well as higher order terms of the series expansion, give vanishing contributions. Thus,

\[
P_{\rho|r} (\rho) \approx \exp \left[ 2^{NR} \left(1 - \rho^2\right)^{\frac{N-1}{2}} - \texttt{sign}_{\rho_i} B\left(\frac{1}{2}, \frac{N-1}{2}\right) \right]
\]

For \( NR \) around 1000, this approximation leads to relative errors below \( 10^{-3} \) for the full range of practically interesting block error probabilities. For even larger blocklength and rate, the approximation becomes even more accurate. Note also from (12) that the distribution of \( \rho \) is doubly-exponential in the blocklength.
B. Angle to Correct Codeword

To derive the BEP, we also need the distribution of the angle between received word and codeword. Consider the right triangle formed by the received word \( r \), the tangential noise component \( z_t \), and the sum of codeword and radial noise component \( c + z_r \) in Fig. 1.

Denoting
\[
\chi = \| z_r \|,
\]
we recognize that \( \chi^2 \) follows a chi-square distribution with \( N - 1 \) degrees of freedom and probability density
\[
P_{\chi^2}(x) = \frac{x^{N-2} e^{-x/2}}{2^{(N-1)/2} \Gamma(N/2)}.
\]

The radial component of the noise is Gaussian distributed with zero mean and unit variance and will be denoted by \( z_r \). It is collinear to the codeword. The sum of the two is denoted as
\[
s = \sqrt{N P + z_r},
\]
and Gaussian distributed with mean \( \sqrt{N P} \) and unit variance.

The angle to the true codeword can be expressed as
\[
\cos \beta = \frac{s}{\sqrt{s^2 + \chi^2}}
\]
by standard trigonometric considerations, cf. Fig. 1. Note that due to the joint independence of all codewords and noise, the angles \( \alpha_i \) and \( \beta \) are statistically independent.

C. Exact Block Error Probability

A maximum likelihood decoder assigns any received word to that codeword which is closest in angle to the received word. A decoding error occurs, if \( \alpha_i = \min_i |\alpha_i| < |\beta| \). Since the cosine function is strictly decreasing within \([0, \pi]\), the error probability is given by
\[
P_e = P_r(\cos \alpha_{\min} > \cos \beta | r)
\]
\[
= P_r \left( \varrho > \frac{s}{\sqrt{s^2 + \chi^2}} | r \right)
\]
\[
= 1 - \int_0^\infty P_{\varrho|r} \left( \frac{s}{\sqrt{s^2 + x}} \right) \frac{e^{(-\chi^2)/(2\pi x)}}{\sqrt{2\pi x}} dx p_{\chi^2}(x) dx
\]
with \( P_{\varrho|r}(\cdot) \) specified in (6), (9), or (12).

The numerical evaluation of this integral is straightforward. The integration over \( s \) and \( x \) can be performed by Gauss-Hermite and Gauss-Laguerre quadrature, respectively, to speed up the computation time.

Following [1], the distribution of the ratio
\[
t = \frac{s}{\chi} \sqrt{N - 1}
\]
is identified as a noncentral \( t \)-distribution with \( N - 1 \) degrees of freedom and noncentrality parameter \( \sqrt{NP} \). In newer versions of Matlab, it is available as \( \text{nctpdf}(t, N-1, \sqrt{NP}) \). Its density will be denoted by \( p_t(t, N - 1, \sqrt{NP}) \), in the sequel. This way, the double integral in (19) can be simplified to a single one
\[
P_e = 1 - \int_r p_{\varrho|r} \left( \frac{t}{\sqrt{t^2 + N - 1}} \right) p_t(t) dt.
\]

This was the method of choice in the late 1950s when integrals were evaluated by tables, but it is not clear, whether it is preferable for evaluation on a modern computer. The noncentral \( t \)-distribution cannot be expressed by polynomials combined with exponentials and/or Gaussian functions. Thus, neither Gauss-Hermite nor Gauss-Laguerre quadrature are straightforward to apply. The combination of sticking to the noncentral \( t \)-distribution together with the numerical issues of (6) for large \( N R \) may explain, why previous literature considered the exact calculation of the BEP as numerically intractable for the most interesting ranges of blocklength.

III. LOWER BOUNDS ON BLOCK ERROR PROBABILITY

While the BEP can be calculated exactly, as shown in the previous section, some applications prefer calculation speed over accuracy. In the sequel, we present a novel tight lower bound on BEP and discuss its connection to the well-known 1959 SPB.

A. Median Bound

A fast and accurate approximation for the exact error probability can be obtained as follows: Note from (12) that \( P_{\varrho|r}(\cdot) \) scales doubly exponential in the blocklength.

This behavior can be utilized to derive very tight lower bounds on error probability: If we replace the exact shape of \( P_{\varrho|r}(\cdot) \) by a unit step function at its median value \( m_\varrho \), we reduce the error probability. This is, as we only shift probability for close wrong codewords to more distant wrong codewords. Clearly, the close wrong codewords can be reached more easily by the noise than the distant ones.

Due to (21), the jump occurs at
\[
m_\varrho = \frac{t}{\sqrt{t^2 + N - 1}}.
\]

Solving for \( t \), the error probability is lower bounded by
\[
P_e > P_t \left( m_\varrho \sqrt{N - 1 \over 1 - m_\varrho^2}, N - 1, \sqrt{NP} \right)
\]
with \( P_t(t, N - 1, \sqrt{NP}) \) denoting the cumulative distribution function of the noncentral \( t \)-distribution. It is available in newer versions of Matlab as \( \text{nctcdf}(t, N-1, \sqrt{NP}) \).

The median value \( m_\varrho \) is easily found by setting the cumulative distribution function (6) to \( \frac{1}{2} \). Since the median \( m_\varrho \) is clearly positive, we obtain with (7)
\[
m_\varrho^2 = 1 - B^{-1} \left( \frac{N - 1}{2} \frac{22}{2} - 2^{1 - (2^{N-1}-1)} \right)
\]
\[
> 1 - B^{-1} \left( \frac{N - 1}{2} \frac{2\ln 2}{2^{N-1} - 1} \right)
\]
where $B^{-1}(a, b, x)$ denotes the inverse with respect to composition of $B(a, b, x)$. It is available in Matlab as betaincinv\((x, a, b)\). The inequality follows from the first order Taylor series of the exponential function
\[
2^{-x} > 1 - x \ln 2 \quad \text{for} \quad x = (2^{NR} - 1)^{-1}.
\]
Inequality (25) is useful as (24) can numerically only be evaluated for moderate values of $NR$.

### B. Sphere Packing Bound

Another tight lower bound utilizing the double exponential scaling of $P_{\text{BEP}}(\cdot)$ is Shannon’s 1959 SPB. In [1], he derives it by arguing that if all Voronoi regions would be equal in size and circular, error probability would be improved. Thus, the total solid angle should be $2^{NR}$ times the solid angle of one circular Voronoi region. In the view of this article, his argument corresponds to the following line of thought: The edge of the circular Voronoi region is such that the probability of being within it is $2^{-NR}$. Specifying the edge of the circular Voronoi region by the cosine of the angle between edge and center, we get with (3)
\[
2^{-NR} = \frac{1}{\rho_{\text{SP}}} p_\rho(\rho) d\rho = \frac{1}{2} \frac{1}{2} B \left( \frac{1}{2}, \frac{N - 1}{2}, \rho_{\text{SP}}^2 \right)
\]
and with (7)
\[
\rho_{\text{SP}}^2 = 1 - B^{-1} \left( \frac{N - 1}{2}, \frac{2}{2NR} \right).
\]
Substituting $\rho_{\text{SP}}$ for $m_{ij}$ in (23) gives the 1959 SPB.

Note that for any $x \geq 2$, $2^x - 1 > 2^x \ln 2$. Thus, whenever the code contains at least 4 codewords, the SPB is less tight than the median bound (25). Unless $NR$ is small, the number of codewords is much larger than 1. In that regime, the sphere packing assumption, i.e. the assumption that the Voronoi regions are hyperspheres, is equivalent to having $1/\ln 2 \approx 1.443$ times more codewords available at the same given minimum distance.

### C. Numerical Results

In the sequel, we compare the exact calculation of the error probability to the median bound and the 1959 SPB. Fig. 2 compares the three over a wide range of blocklengths for a rate only 0.2% below channel capacity. The median bound is observed to lie pretty much in the middle between the exact result and the 1959 SPB. Note that the exact calculation becomes numerically troublesome, if the product of rate and blocklength exceeds 1000. For larger blocklengths, the approximation (12) is used. It can be observed that this approximation is sufficiently tight for practical use even for much lower blocklengths.

For rates further away from capacity, the behavior is very similar, but hard to depict, as the error probabilities span a very wide range in logarithmic scale. Note that the tightness of the approximation primarily depends on the product $NR$.

For larger rates, the approximation becomes tight at even smaller blocklengths.

In order to utilize approximation (12) for the two bounds, one needs to solve (12) for $\rho = \cos \alpha_{\text{min}}$. This cannot be done in closed form, but very efficiently by fixed point iteration: Solve (12) for the $\rho$ in the numerator (on the right hand side) assuming $\rho$ in the denominator being fixed. Then, start the fixed point iteration for some given value of the $\rho$ that was assumed fixed. It is shown in the extended journal version of this manuscript [7] that $N < 2NR$ is a sufficient condition for the fixed point iterations to converge.

### IV. Gaussian Random Code Ensemble

Consider now the case that the codewords are not on the hypersphere, but their components are independent and identically distributed (iid) Gaussian with zero mean and variance $P$.

We decompose the vector to an alternative codeword $w_i$ into a radial and a tangential component relative to the received word $r$
\[
w_i = w_{r,i} + w_{t,i}.
\]
Note that $||w_{t,i}||^2 / P$ is chi-squared distributed with $N - 1$ degrees of freedom. Its probability density is given by (14). The radial component $w_{r,i}$ is zero mean Gaussian with variance $P$. The two components are statistically independent.

Define the normalized squared distance to the received word as
\[
d_i^2 = \frac{||w_{t,i}||^2 + (||r|| - w_{r,i})^2}{P}.
\]
Conditioned on the received word $r$, it follows a noncentral chi-square distribution with $N$ degrees of freedom, non-
centrality parameter $||r||^2/P$, and cumulative distribution function [6]

$$P_{d^2|r}(d) = 1 - Q_{\frac{d}{\sqrt{P}}} \left( \frac{||r||}{\sqrt{P}}, \sqrt{d} \right) \quad (31)$$

with $Q_{\alpha}(a,b)$ denoting the generalized Marcum Q-function. The distribution (31) is available in Matlab as `ncx2cdf(d,N,norm(r)^2/P)`.

We are interested in the distribution of the minimum distance to the received word:

$$P_{d} \left( x < P \min_i d_i \right) r = P_{x^2/P < d^2} \left( \frac{x}{\sqrt{P}} \right) \chi_{2N-1}$$

$$= Q_{\frac{||r||}{\sqrt{P}}} \left( \frac{x}{\sqrt{P}} \right) \chi_{2N-1} \quad (32)$$

The squared distance of the received word to the correct codeword is $||z||^2 = \chi^2 + z^2$ with (13). Furthermore, $||r||^2 = \chi^2 + (z_i + ||c||)^2$. Collecting these results, we get the error probability as

$$P_e = 1 - P_{||r||^2 < \min_i d_i^2} \quad (34)$$

$$= 1 - \frac{E_{||c||,z_i,\chi} Q_{\frac{\sqrt{\chi^2+||c||^2+z_i^2}}{\sqrt{\chi^2+z_i^2}}} \chi_{2N-1}} {\chi_{2N-1}} \quad (35)$$

The random variables $||c||, z_i$, and $\chi$ are statistically independent and their distributions are known. The expectation in (35), however, can only be evaluated numerically. Depending on the number of codewords $2^{NR}$, it can, however, be calculated directly or utilizing (8).

Gaussian random coding is compared against spherical random coding in terms of BEP vs. blocklength in Fig. 3. For blocklength below 17, Gaussian random coding performs superior. This effect is the more pronounced the larger the code rate. For low dimensions, spherical random codes suffer from not utilizing the radial component for data transmission. For large dimensions, this rate loss is negligible.

V. Binary Symmetric Channel

Consider now a BSC with iid. binary codewords having equal probability. In order to streamline the presentation with respect to the previous chapters, let the alphabet of the code be $\{+1,-1\}$. Since all codewords and the received word lie on an hypersphere in $N$ dimensions, we can argue in the same way as for spherical random coding on the AWGN channel: No error occurs, as long as the smallest angle between the received word and a wrong codeword is larger than the angle between the received word and the true codeword.

The correlation coefficient between a random wrong codeword and the received word scaled by the blocklength $\alpha = N \cos \alpha_i$ is called antipodal overlap in the sequel. In order to simplify subsequent notation, we also introduce the Boolean overlap

$$l_i = \frac{N + \alpha_i}{2} \quad (36)$$

which corresponds to codewords in $\{0,1\}^N$. Since the two overlaps are strictly increasing functions of each other, maximizing the one also maximizes the other.

The Boolean overlap follows the binomial distribution [8]

$$P_l(l) = 2^{-N} \sum_{i=0}^{l} \binom{N}{i} = B \left( \frac{[N-l], [l]+1, \frac{1}{2} \right)$$

for $l \in [0; N]$. Defining $l = \max_i l_i$, we get

$$P_{\ell\mid r}(\ell) = B \left( \frac{[N-l], [\ell]+1, \frac{1}{2} \right) 2^{2N-1} \quad (37)$$

For large values of $NR$, direct evaluation of (37) can run into numerical trouble. However, one can circumvent that in the same way as for spherical Gaussian codes via (7). In analogy to (9), this leads to

$$P_{\ell\mid r}(\ell) = \prod_{i=1}^{\infty} e^{(1-2^NR)b(|\ell|+1,[N-\ell],\frac{1}{2})} \quad (38)$$

In practice, very few factors (typically only a single one) of the infinite product is required to obtain sufficient accuracy.

The statistics of the antipodal overlap between the received word and the correct codeword can be found in a similar way. We can model the BSC as a multiplicative noise $z \in \{-1,+1\}^N$ on top of the correct codeword $c$. For crossover probability $f$, we have $Pr(z_i = -1) = f$. The Boolean overlap is given by

$$\tau = \frac{\sum_{i=1}^{N} 1 + z_i}{2} \quad (39)$$

It is binomially distributed with parameters $N$ and $f$. 

![Fig. 3: BEP vs. blocklength for Gaussian random coding (black line) for $P = 100$ and $R = 3$. The colored lines refer to spherical random coding (colors as in Fig. 2).](image-url)
We do not get the exact error probability applying (18), but the following upper and lower bounds \( P_l < P_e < P_u \).

\[
P_l = 1 - \sum_{i=0}^{N} \binom{N}{i} f^{N-i}(1-f)^i P_{\ell|x}(i)
\]

\[
P_u = 1 - \sum_{i=1}^{N} \binom{N}{i} f^{N-i}(1-f)^i P_{\ell|x}(i-1).
\]

For discrete random variables, there is a non-zero probability that the two Boolean overlaps equal each other. The lower bound and upper bound assume that, in case of equality, the decisions are always correct (too optimistic) and always erroneous (too pessimistic), respectively. Clearly, the pessimistic attitude is typically much closer to reality.

For large blocklengths, the binomial coefficients in (40) and (41) become very large, while the powers involving the crossover probabilities may become very small. The resulting numerical trouble is overcome by calculating their logarithms. This can be overcome as follows: We replace the factor \( (1-\xi)^{\ell} \) in (5) by its Taylor series at \( \xi = x \). Restricting the Taylor series to first order and taking only the first factor in (38), we obtain an approximation that is sufficiently accurate for most practical cases. It reads

\[
P_{\ell|x}(\ell) \approx \exp \left[ \frac{2^{N(1-R)}}{B(N-\ell, \ell+1)} \left( \frac{N-\ell-1}{\ell+2} - \frac{N-\ell}{\ell+1} \right) \right]
\]

assuming \( \ell \) integer within \([0; N]\). Like (12), it is doubly exponential in the blocklength.

The trade-off between rate and blocklength is depicted in Fig. 4. Apparently, approximation (42) is very tight for the full range of blocklengths. Furthermore, the upper bound is tighter than the random coding union bound of [5]. This is hardly visible in Fig. 4, but is apparent in Fig. 5, which is a magnification of Fig. 4. Comparisons to other bounds can easily be made with the help of [5, Figures 1 & 2].

VI. CONCLUSIONS

Shannon’s formula for average BEP can be evaluated exactly by fast numerical integration methods for rate-blocklength products up to about 1000. Almost the same holds for the BSC where a very tight upper bound can be found. For larger rate-blocklength products, the average BEP can be approximated with relative errors on the order of \( 10^{-3} \) or less. The same holds for Shannons’s 1959 SPB and the tighter median bound. If the blocklength is sufficiently short and the rate sufficiently high, the uniform spherical code ensemble falls behind the iid. Gaussian code ensemble.

ACKNOWLEDGMENT

The author would like to thank S. Asaad and A. Bereyhi for proofreading of the manuscript, M. Bossert, G. Caire, A. Guillén i Fàbregas, G. Kramer, and G. Liva for stimulating discussions, as well as M. Coşkun for help with the figures on the BSC.

REFERENCES

[1] C. E. Shannon, “Probability of error for optimal codes in a Gaussian channel,” The Bell System Technical Journal, vol. 38, no. 3, pp. 611–656, May 1959.

[2] D. Slepian, “Bounds on communication,” The Bell System Technical Journal, vol. 42, pp. 681–707, 1963.

[3] I. Sason and S. Shamai, Performance Analysis of Linear Codes under Maximum-Likelihood Decoding: A Tutorial. Now Foundations and Trends, 2006.

[4] G. Liva and F. Steiner, “pretty-good-codes.org: Online library of good channel codes,” http://pretty-good-codes.org, Feb. 2020.

[5] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” IEEE Transactions on Information Theory, vol. 56, no. 5, pp. 2307–2359, May 2010.

[6] G. Casella and R. L. Berger, Statistical Inference, 2nd ed. Duxbury Pacific Grove, CA, USA, 2002.

[7] R. Müller, “On approximation, bounding & exact calculation of block error probability for random code ensembles,” Submitted to IEEE Transactions on Communications, Mar. 2020, https://arXiv:2003.06807.

[8] G. P. Wadsworth and J. G. Bryan, Introduction to Probability and Random Variables. New York: McGraw-Hill, 1960.