An Improved Sphere-Packing Bound for Finite-Length Codes on Symmetric Memoryless Channels

Gil Wiechman   Igal Sason

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel
{igillw@tx, sason@ee}.technion.ac.il

April 19, 2007

Outline

Sphere-Packing Bounds

- Lower bounds on the decoding error probability of optimal block codes, given in terms of
  1. block length
  2. rate
  3. communication channel
- Valid under optimal ML decoding or even under list decoding.
- Based on geometrical properties of the decoding regions.
- Decay to zero exponentially with the block length.

The 1967 Sphere-Packing (SP67) Bound

Shannon, Gallager and Berlekamp [1967].

- A lower bound on the decoding error probability of block codes transmitted over discrete memoryless channels (DMCs).
- Error exponent is exact for all rates above the critical rate of the channel.
- The derivation of the bound is divided into three main steps:
  1. Lower bound on the pairwise error probability for a code of two codewords.
  2. Lower bound on maximal error prob. for fixed composition codes.
  3. Lower bound on the block error probability of arbitrary block codes.

Step 1: Lower bound on the pairwise error probability for a code of two codewords

- Consider a code which consists of two codewords $\mathbf{x}_1$ and $\mathbf{x}_2$.
- $P_i(\mathbf{y})$ - Probability of receiving $\mathbf{y}$ when $\mathbf{x}_i$ is transmitted ($i = 1, 2$).
- Let us define
  - The log-likelihood ratio $D(\mathbf{y}) \triangleq \ln\left(\frac{P_2(\mathbf{y})}{P_1(\mathbf{y})}\right)$
  - The function $\mu(s) \triangleq \ln\left(\sum_{\mathbf{y}} P_1(\mathbf{y})^{1-s} P_2(\mathbf{y})^s\right), \quad 0 < s < 1$
  - The probability distribution $Q_s(\mathbf{y}) \triangleq \frac{P_1(\mathbf{y})^{1-s} P_2(\mathbf{y})^s}{\sum_{\mathbf{y}'} P_1(\mathbf{y}')^{1-s} P_2(\mathbf{y}')^s}, \quad 0 < s < 1.$

Step 1 (Cont.)

- For $0 < s < 1$

$$\begin{aligned} \mu'(s) &= \mathbb{E}_{Q_s}\big(D(\mathbf{y})\big) \\ \mu''(s) &= \mathrm{Var}_{Q_s}\big(D(\mathbf{y})\big) \end{aligned}$$

  where $\mathbb{E}_{Q_s}$ and Var_{Q_s} designate the statistical expectation and variance, respectively, w.r.t. the probability distribution Q_s.

- The following equalities hold:

$$\begin{aligned} P_1(\mathbf{y}) &= \exp\big(\mu(s) - sD(\mathbf{y})\big)\, Q_s(\mathbf{y}) \\ P_2(\mathbf{y}) &= \exp\big(\mu(s) + (1-s)D(\mathbf{y})\big)\, Q_s(\mathbf{y}) \end{aligned}$$

- Let
  - $\mathcal{Y}_1$ and $\mathcal{Y}_2$ be disjoint decision regions for $\mathbf{x}_1$ and $\mathbf{x}_2$ and assume that they partition the observation space (i.e., $\mathcal{Y}_1 = \mathcal{Y}_2^c$ where the superscript 'c' stands for the complementary set).
  - $\mathcal{Y}_s \triangleq \left\{\mathbf{y} : |D(\mathbf{y}) - \mu'(s)| \le \sqrt{2\mu''(s)}\,\right\}$.

Step 1 (Cont.)

Lower bound on the conditional error probability given that the first codeword is transmitted

$$
\begin{aligned}
P_{e,1} &= \sum_{\mathbf{y}\in\mathcal{Y}_1^c} P_1(\mathbf{y}) \\
&\geq \sum_{\mathbf{y}\in\mathcal{Y}_1^c \bigcap \mathcal{Y}_s} P_1(\mathbf{y}) \\
&= \sum_{\mathbf{y}\in\mathcal{Y}_1^c \bigcap \mathcal{Y}_s} \exp\Big(\mu(s) - sD(\mathbf{y})\Big)\, Q_s(\mathbf{y}) \\
&\geq \exp\left(\mu(s) - s\mu'(s) - s\sqrt{2\mu''(s)}\right) \sum_{\mathbf{y}\in\mathcal{Y}_1^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y})
\end{aligned}
$$

Step 1 (Cont.)

- Similarly

$$P_{e,2} \geq \exp\left(\mu(s) + (1-s)\mu'(s) - (1-s)\,\sqrt{2\mu''(s)}\right) \sum_{\mathbf{y} \in \mathcal{Y}_2^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y})$$

- Since $\mathcal{Y}_1^c = \mathcal{Y}_2$, then

$$\sum_{\mathbf{y} \in \mathcal{Y}_1^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y}) + \sum_{\mathbf{y} \in \mathcal{Y}_2^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y}) = \sum_{\mathbf{y} \in \mathcal{Y}_s} Q_s(\mathbf{y})$$

- From the Chebychev inequality $\sum_{\mathbf{y} \in \mathcal{Y}_s} Q_s(\mathbf{y}) > \frac{1}{2}$

$$\Rightarrow \text{either} \sum_{\mathbf{y} \in \mathcal{Y}_1^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y}) > \frac{1}{4} \quad \text{or} \quad \sum_{\mathbf{y} \in \mathcal{Y}_2^c \bigcap \mathcal{Y}_s} Q_s(\mathbf{y}) > \frac{1}{4}.$$

Step 1 (Cont.)

One therefore gets that for for every $s \in (0, 1)$

$$P_{e,1} \triangleq \sum_{\mathbf{y} \in \mathcal{Y}_2} P_1(\mathbf{y}) > \frac{1}{4} \exp\Big(\mu(s) - s\mu'(s) - s\sqrt{2\mu''(s)}\Big)$$

or

$$P_{e,2} \triangleq \sum_{\mathbf{y} \in \mathcal{Y}_1} P_2(\mathbf{y}) > \frac{1}{4} \exp\Big(\mu(s) + (1-s)\mu'(s) - (1-s)\sqrt{2\mu''(s)}\Big).$$

Step 2: Fixed composition codes

- Consider a block code $\mathcal{C}$ containing M codewords of length N.
  - Denote the codewords by $\{\mathbf{x}_m\}_{m=1}^{M}$.
  - Assume that the codewords have a fixed composition.
- Transmission over a DMC with transition probabilities $P(y|x)$.
- Let
  - $\mathcal{K} = \{1, \ldots, K\}$ be the input alphabet of the DMC.
  - $\mathcal{J} = \{1, \ldots, J\}$ be the output alphabet of the DMC.

Step 2 (Cont.)

1. Consider a probability measure over N-length channel output vectors which can be factorized in the form

$$f_N(\mathbf{y}) = \prod_{n=1}^{N} f(y_n).$$

2. The size of the set $\mathcal{Y}_m$ is defined as

$$F(\mathcal{Y}_m) \triangleq \sum_{\mathbf{y}\in\mathcal{Y}_m} f_N(\mathbf{y}).$$

3. Assigning

$$\begin{aligned} P_1(\mathbf{y}) &\triangleq P(\mathbf{y}|\mathbf{x}_m), \quad P_2(\mathbf{y}) \triangleq f_N(\mathbf{y}) \\ \mathcal{Y}_1 &\triangleq \text{Decision region for } \mathbf{x}_m, \quad \mathcal{Y}_2 \triangleq \mathcal{Y}_1^c \end{aligned}$$

$\Rightarrow$ Step 1 gives a lower bound either on the conditional error probability $P_{\mathrm{e},m}$ or the size of the decision region for $\mathbf{x}_m$.

Step 2 (Cont.)

- Let q_k be the fraction of appearances of the letter k in each codeword of the fixed composition code.
- Since the channel is memoryless, and f_N is factorized as before, then

$$\mu(s, f_N) = N \sum_{k=1}^{K} q_k \mu_k(s, f)$$

where

$$\mu_k(s, f) \triangleq \ln \left(\sum_{j=1}^{J} P(j|k)^{1-s} f(j)^s \right) .$$

Step 2 (Cont.)

- Consider now the codeword with the smallest decision region
  - Since the size of the list is limited to L, then $\sum_{m=1}^{M} F(\mathcal{Y}_m) \leq L$. Hence, there is an index m for which $F(\mathcal{Y}_m) \leq \frac{L}{M}$.
  - Upper bound the conditional error probability of this codeword ($P_{\mathrm{e},m}$) by $P_{\mathrm{e,max}}$.
    Moreover, the size of its decision region is upper bounded by $\frac{L}{M}$.
- Choose $s \in (0,1)$ so that the lower bound on $\frac{L}{M}$ is not satisfied
  $\Rightarrow$ The lower bound on $P_{\mathrm{e,max}}$ for fixed composition codes holds.
- This bound depends on the composition and the choice of $f_N(\mathbf{y})$.

- For a universal bound for all fixed composition codes,
  one needs to solve the following min-max problem:
  - Optimize over the probability measure $f_N(\mathbf{y})$ to get the tightest (maximal) lower bound within this form.
  - Consider the composition which minimizes the bound.

Step 2 (Cont.)

For $0 < s < 1$, let $\mathbf{q}_s = \{q_{1,s}, \ldots, q_{K,s}\}$ satisfy the inequalities

$$\sum_j P(j|k)^{1-s} \alpha_{j,s}^{\frac{s}{1-s}} \geq \sum_j \alpha_{j,s}^{\frac{1}{1-s}} \; ; \quad \forall k$$

where $\alpha_{j,s} \triangleq \sum_{k=1}^{K} q_{k,s} P(j|k)^{1-s}$. Then, the function $f = f_s$ is given by

$$f_s(j) = \frac{\alpha_{j,s}^{\frac{1}{1-s}}}{\sum_{j'=1}^{J} \alpha_{j',s}^{\frac{1}{1-s}}}, \quad j \in \{1, \ldots, J\}.$$

Note that the input distribution $\mathbf{q}_s$ is $\underline{\text{independent of the code}}$ $\mathcal{C}$, as it only depends on the channel statistics. By setting $\rho = \frac{s}{1-s}$, this form is identical to the necessary and sufficient conditions on $\mathbf{q}$ to maximize Gallager's reliability function $E_0(\rho, \mathbf{q})$.

Step 3: Lower bound on the error probability of block codes

From a universal lower bound on the maximal error probability of fixed composition codes to such a bound for general block codes:

- For a block code of length N and alphabet size K, there are at most $\binom{N+K-1}{K-1}$ possible compositions.
- Since $\binom{N+K-1}{K-1} < N^K$, there exists a fixed composition subcode with at least $\frac{M}{N^K}$ codewords.
- The maximal error probability of any block code with M codewords of length N is lower bounded by the maximal error probability from Step 2 for a fixed composition code with $\frac{M}{N^K}$ codewords.

Step 3 (Cont.)

From the maximal error probability to the average error probability:
Expurgation $\Rightarrow$ The average error probability of a general block code is at least half the maximal error probability of the subcode containing half of the codewords with the lowest error probability.

Conclusion: The average error probability of an (N, M) block code defined over an alphabet of size K is not less than half the maximal error probability of a certain $(N, \frac{M}{2\,N^K})$ fixed composition subcode.

Combine this conclusion with the lower bound in Step 2 for fixed composition codes $\Rightarrow$ the resulting lower bound on the error probability of general block codes under list decoding is the 1967 sphere-packing (SP67) bound of Shannon, Gallager & Berlekamp.

Theorem (The 1967 Sphere-Packing Bound)

- Let $\mathcal{C}$ be a block code consisting of M codewords each of length N.
- Assume communication over a DMC, and let $P(j|k)$ designate the transition probabilities where $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, J\}$ are the channel input and output alphabets, respectively.
- Assume a list decoder where the size of the list is limited to L.
- Define

$$R \triangleq \frac{\ln\left(\frac{M}{L}\right)}{N} \quad - \quad \text{code rate in nats per channel use}$$

$$P_{\min} \quad - \quad \text{smallest non-zero transition probability of the DMC.}$$

- Then, the *average decoding error probability* is lower bounded by

$$P_e(N, M, L) \geq \exp\left\{-N\left[E_{\text{sp}}\left(R - O_1\left(\frac{\ln N}{N}\right)\right) + O_2\left(\frac{1}{\sqrt{N}}\right)\right]\right\}$$

Theorem (The 1967 Sphere-Packing Bound)

where

$$
\begin{aligned}
E_{\mathrm{sp}}(R) &\triangleq \sup_{\rho \geq 0} \left(E_0(\rho) - \rho R \right) \\
E_0(\rho) &\triangleq \max_{\mathbf{q}} E_0(\rho, \mathbf{q}) \\
E_0(\rho, \mathbf{q}) &\triangleq -\ln \left(\sum_{j=1}^{J} \left[\sum_{k=1}^{K} q_k P(j|k)^{\frac{1}{1+\rho}} \right]^{1+\rho} \right) \\
O_1\left(\frac{\ln N}{N}\right) &\triangleq \frac{\ln 8}{N} + \frac{K \ln N}{N} \\
O_2\left(\frac{1}{\sqrt{N}}\right) &\triangleq \sqrt{\frac{8}{N}} \ln\left(\frac{e}{\sqrt{P_{\min}}}\right) + \frac{\ln 8}{N}.
\end{aligned}
$$

Notes on the Classical SP67 Bound

- The original focus in the derivation of the SP67 bound was on asymptotic analysis.
- The aim was to make the derivation as simple as possible, as long as there is no loss in the asymptotic behavior.
- Problem: The SP67 bound is in general very loose for codes of short to moderate block lengths.
- Goal: Improve the tightness of the sphere-packing bound for finite-length codes, especially in light of the remarkable performance of codes defined on graphs (e.g., turbo, LDPC, RA codes etc.) even for short to moderate block lengths.

The Valembois & Fossorier (VF) Bound

A. Valembois and M. Fossorier, “Sphere-packing bounds revisited for moderate block length,” *IEEE Trans. on IT*, Vol. 50, Decemeber 2004.

- Valembois and Fossorier revisited the derivation of the SP67 bound, and found four points where the bound could be tightened for codes of short to moderate block lengths.
- These improvements also make the bound valid for memoryless channels with discrete input and continuous output.
- The resulting sphere-packing bound (referred to as the VF bound) is uniformly tighter than the SP67 bound.

The Improvements of Valembois & Fossorier

1. The considered set in Step 1 can be generalized to

$$\mathcal{Y}_s^x \triangleq \left\{\mathbf{y} : |D(\mathbf{y}) - \mu'(s)| \le x\sqrt{2\mu''(s)}\right\}, \quad x > 0$$

   where x is optimized (in the SP67 bound $x \equiv 1$).
2. Instead of upper bounding $s\sqrt{\mu''(s)}$ by $\ln\left(\frac{e}{\sqrt{P_{\min}}}\right)$, calculate the exact value.
3. To emphasize the similarity between the random-coding and the SP67 bounds, ρ was selected sub-optimally $\Rightarrow$ Select optimal ρ.
4. Instead of bounding the number of possible compositions, use the exact number.

Theorem (The Valembois & Fossorier (VF) Bound)

Under the assumptions and notation used for the SP67 bound, the *average decoding error probability* is lower bounded by

$$P_e(N, M, L) \geq \exp\left\{-N\widetilde{E}_{sp}(R, N)\right\}$$

where

$$\widetilde{E}_{sp}(R, N) \triangleq \sup_{x > \frac{\sqrt{2}}{2}} \left\{E_0(\rho_x) - \rho_x \left(R - O_1\left(\frac{\ln N}{N}, x\right)\right) + O_2\left(\frac{1}{\sqrt{N}}, x, \rho_x\right)\right\}.$$

For more details, see Thm. 7 in the paper of Valembois and Fossorier.

Notes on the SP67 and VF bounds

- While the SP67 bound can be applied only to a DMC, the VF bound can be also applied to memoryless channels of continuous output alphabet (since it does not require that $P_{\text{min}} > 0$, and calculate instead the second derivative of μ exactly).
- The rate shift in their error exponents scales like $\frac{\ln N}{N}$ for both bounds. This is due to the need to consider fixed composition codes (i.e., Step 2 in the derivation of the SP67 and VF bounds).
- Both bounds use expurgation of half the codewords to transform a lower bound on the *maximal* error probability to a lower bound on the *average* error probability.

Discussion on Sphere-Packing Bounds

Question

Is it necessary to consider fixed composition codes first ?

Discussion on Sphere-Packing Bounds

Question

Is it necessary to consider fixed composition codes first ?

Answer

In general, yes! Since μ and its derivatives are affected by the codeword composition, it is required to fix the composition in order to find the optimal tilting measure f.

However, for symmetric memoryless channels, μ and its first two derivatives are independent of the codeword composition.

This observation yields that for symmetric memoryless channels, the sphere-packing bounding technique can be directly applied to general block codes (without necessarily a fixed composition).

Discussion on Sphere-Packing Bounds

Question

Is it necessary to consider fixed composition codes first ?

Answer

In general, yes! Since μ and its derivatives are affected by the codeword composition, it is required to fix the composition in order to find the optimal tilting measure f.

However, for symmetric memoryless channels, μ and its first two derivatives are independent of the codeword composition.

This observation yields that for symmetric memoryless channels, the sphere-packing bounding technique can be directly applied to general block codes (without necessarily a fixed composition).

Discussion on Sphere-Packing Bounds

Question

Is it necessary to consider the maximal error probability first?

Discussion on Sphere-Packing Bounds

Question

Is it necessary to consider the maximal error probability first?

Answer

By modifying the first step of the derivation to consider the average error probability over M pairs of codewords, where the index m of the selected pair is chosen uniformly at random and known at the decoder, it is possible to directly consider the *average* error probability.

This stage also requires that μ and its first two derivatives are independent of the selected pair of codewords.

An Improved Sphere-Packing (ISP) Bound

- The derivation of the ISP bound relies on
  - the observation that for symmetric memoryless channels, $\mu(s, f_s)$ and its first two derivatives are independent of the codeword composition.
  - the improvements suggested by Valembois and Fossorier for the derivation of the VF bound.

An Improved Sphere-Packing (ISP) Bound

- The derivation of the ISP bound relies on
  - the observation that for symmetric memoryless channels, $\mu(s, f_s)$ and its first two derivatives are independent of the codeword composition.
  - the improvements suggested by Valembois and Fossorier for the derivation of the VF bound.
- The ISP bound forms a tighter sphere-packing bound
  1. by considering *M* codeword pairs in the first step of the derivation
     $\Rightarrow$ direct analysis of *average* error probability,
     eliminating the need for expurgation of half the codewords
  2. by the independence of μ, μ' and μ'' from codeword composition
     $\Rightarrow$ direct analysis of general block codes,
     eliminating the need for considering fixed-composition codes.

An Improved Sphere-Packing (ISP) Bound

- The derivation of the ISP bound relies on
  - the observation that for symmetric memoryless channels, $\mu(s, f_s)$ and its first two derivatives are independent of the codeword composition.
  - the improvements suggested by Valembois and Fossorier for the derivation of the VF bound.
- Though this observation has no effect on asymptotic analysis, it affects the tightness of the bound for finite-length codes (especially, for short to moderate block lengths).

This gives the following improvement on the SP67 and VF bounds.

Theorem (Improved Sphere-Packing Bound)

- Let $\mathcal{C}$ be an arbitrary block code consisting of M codewords, each of length N.
- Assume communication over a symmetric memoryless channel specified by the transition probabilities (or densities) $P(j|k)$.
- Assume a list decoder where the size of the list is limited to L.
- Then, the *average decoding error probability* is lower bounded by

$$P_{\mathrm{e}}(N, M, L) \geq \exp\left\{-N\widetilde{E}_{\mathrm{sp}}(R, N)\right\}$$

Theorem (Improved Sphere-Packing Bound)

where $\widetilde{E}_{\text{sp}}(R,N) \triangleq \sup_{x>\frac{\sqrt{2}}{2}} \left\{ E_0(\rho_x) - \rho_x \left(R - O_1\left(\frac{1}{N}, x\right)\right) + O_2\left(\frac{1}{\sqrt{N}}, x, \rho_x\right) \right\}$

R designates the code rate in nats per channel use, and

$$\begin{aligned} O_1\left(\frac{1}{N}, x\right) &\triangleq \frac{\ln 4}{N} - \frac{\ln\left(2 - \frac{1}{x^2}\right)}{N} \\ O_2\left(\frac{1}{\sqrt{N}}, x, \rho\right) &\triangleq s(\rho)x\sqrt{\frac{8}{N}\mu_0''(s(\rho), f_{s(\rho)})} + \frac{\ln 4}{N} - \frac{\ln\left(2 - \frac{1}{x^2}\right)}{N}. \end{aligned}$$

Here, $s(\rho) \triangleq \frac{\rho}{1+\rho}$, and $\rho = \rho_x$ is determined by solving the equation

$$\begin{aligned} R - O_1\left(\frac{1}{N}, x\right) &= -\mu_0\left(s(\rho), f_{s(\rho)}\right) - \left(1 - s(\rho)\right)\mu_0'\left(s(\rho), f_{s(\rho)}\right) \\ &\quad + \left(1 - s(\rho)\right) x \sqrt{\frac{2\mu_0''\left(s(\rho), f_{s(\rho)}\right)}{N}} \end{aligned}$$

Notes on the ISP Bound

- The ISP bound differs from the VF bound in the sense that the term $\frac{\log\binom{N+K-1}{K-1}}{N}$ is removed from $O_1(\frac{\ln N}{N}, x)$.
  $\Rightarrow$ The rate shift in the error exponent of the ISP bound scales like $\frac{1}{N}$, as opposed to $\frac{\ln N}{N}$ for error exponents of the VF and SP67 bounds.
- The rate shift of $\frac{\ln 2}{N}$ and the multiplicative value of $\frac{1}{2}$ which stem from the expurgation of half the codewords are also removed.

Symmetric memoryless channels

In our work, we use the following definition for symmetric memoryless channels:

Definition (Symmetric Memoryless Channels)

A memoryless channel with input alphabet $\mathcal{K} = \{0, 1, \ldots, K-1\}$, output alphabet $\mathcal{J} \subseteq \mathbb{R}^d$ (where $K, d \in \mathbb{N}$) and transition probability (or density if $\mathcal{J}$ non-countable) $P(\cdot|\cdot)$ is said to be *symmetric* if there exists a set of unitary (bijective) mappings $\{g_k\}_{k=0}^{K-1}$ where $g_k : \mathcal{J} \to \mathcal{J}$ for all $k \in \mathcal{K}$ such that

$$\forall \mathbf{y} \in \mathcal{J}, k \in \mathcal{K} \quad P(\mathbf{y}|0) = P\big(g_k(\mathbf{y})|k\big)$$

and

$$\forall k_1, k_2 \in \mathcal{K} \quad g_{k_1}^{-1} \circ g_{k_2} = g_{(k_2 - k_1) \mathrm{mod} K} \, .$$

Symmetric memoryless channels (Cont.)

- The symmetry conditions required for the ISP bound are mild:
  - All memoryless binary-input output-symmetric (MBIOS) channels are symmetric in this sense
  - All M-ary input and symmetric output (MI-SO) channels (see [Wang et al.,IT Jan 07]) are symmetric in this sense.
- The ISP bound is valid in particular for M-ary PSK modulated signaling, coherently detected over the AWGN channel.

The 1959 sphere-packing (SP59) bound

C. E. Shannon, “Probability of error for optimal codes in a Gaussian channel,” *Bell System Technical Journal*, vol. 38, May 1959.

Consider a block code $\mathcal{C}$ of length N and rate R nats per channel use

- Transmission takes place over an AWGN channel.
- Assume that the codewords are mapped to equal-energy signals.
  $\Rightarrow$ Each Voronoi cell is a polyhedric cone whose tip lies on the origin.

The 1959 sphere-packing (SP59) bound

To measure decoding region size, we use the solid angle of a cone (solid angle of a cone $\triangleq$ area of unit sphere cut out by the cone).

- Denote the solid angle of a circular cone with half-angle θ whose point lies on the origin by $\Omega_N(\theta)$.
- The area of the N-dimensional unit sphere is given by $\Omega_N(\pi)$.
- The decoding regions form a partition of $\mathbb{R}^N$
  $\Rightarrow$ The sum of their solid angles is $\Omega_N(\pi)$.

The derivation of the SP59 bound relies on two main observations:

Main Observations

- Among cones of a given solid angle, the lowest error probability is given by the circular cone centered around the code signal.
- It is best to share the total solid angle equally among the $\exp(NR)$ Voronoi cells.

Corollary

The decoding regions of the optimal code are circular cones whose half angle θ satisfies $\frac{\Omega_N(\theta)}{\Omega_N(\pi)} = exp(-NR)$.

Theorem (The 1959 sphere-packing bound)

Assume a block code of length N and rate R nats per channel use is transmitted over an AWGN channel with noise spectral density $\frac{N_0}{2}$. Then, under ML decoding, the error probability is lower bounded by

$$P_e(\mathrm{ML}) > P_{\mathrm{SPB}}(N, \theta, A)\ , \quad A \triangleq \sqrt{\frac{2E_s}{N_0}}$$

where E_s is the average energy per symbol, $\theta \in [0, \pi]$ satisfies the inequality $exp(-NR) \leq \frac{\Omega_N(\theta)}{\Omega_N(\pi)}$,

$$P_{\mathrm{SPB}}(N, \theta, A) \triangleq \frac{(N-1)e^{-\frac{NA^2}{2}}}{\sqrt{2\pi}} \int_{\theta}^{\frac{\pi}{2}} (\sin\phi)^{N-2} f_N(\sqrt{N}A\cos\phi)\, d\phi + Q(\sqrt{N}A).$$

and

$$f_N(x) \triangleq \frac{1}{2^{\frac{N-1}{2}}\Gamma(\frac{N+1}{2})} \int_0^{\infty} z^{N-1} \exp(-\frac{z^2}{2} + zx)\, dz\ , \quad \forall\, x \in \mathbb{R},\ N \in \mathbb{N}.$$

Asymptotically tight approximations for the SP59 bound

- For block lengths of $N = 1000$ and more, the SP59 bound becomes difficult to calculate due to numerical problems.
- Shannon presented some asymptotic formulas which give a good estimation of the bound for large block lengths.
- The calculation of these approximations is done in the log-domain, thus eliminating numerical difficulties for large block lengths.

Theorem (Shannon's approximation for the SP59 bound)

Defining

$$\begin{aligned}
G(\theta) &\triangleq \frac{A\cos\theta + \sqrt{A^2\cos^2\theta + 4}}{2} \\
E_L(\theta) &\triangleq \frac{A^2 - AG(\theta)\cos\theta - 2\ln(G(\theta)\sin\theta)}{2} \\
\alpha(\theta) &\triangleq \left(\sqrt{\pi\left(1 + G^2(\theta)\right)}\sin\theta(AG(\theta)\sin^2\theta - \cos\theta)\right)^{-1}
\end{aligned}$$

if $\theta > \cot^{-1}(A)$, then

$$P_{\text{SPB}}(N, \theta, A) \approx \frac{\alpha(\theta)e^{-NE_L(\theta)}}{\sqrt{N}}$$

Recursive algorithm for calculating the SP59 bound

A. Valembois and M. Fossorier, "Sphere-packing bounds revisited for moderate block length," *IEEE Trans. on IT*, Vol. 50, December 2004.

Theorem

The set of functions $\{f_N\}$ can be expressed in the alternative form

$$f_N(x) = P_N(x) + Q_N(x) \exp(\frac{x^2}{2}) \int_{-\infty}^{x} \exp(-\frac{t^2}{2})\, dt\ , \qquad \begin{array}{l} x \in \mathbb{R}, \\ N \in \mathbb{N} \end{array}$$

where P_N and Q_N are polynomials defined by the recursive equation

$$\begin{aligned} P_N(x) &= \frac{2N-5+x^2}{N-1}\, P_{N-2}(x) - \frac{N-4}{N-1}\, P_{N-4}(x)\ , \\ Q_N(x) &= \frac{2N-5+x^2}{N-1}\, Q_{N-2}(x) - \frac{N-4}{N-1}\, Q_{N-4}(x), \qquad N \geq 5 \end{aligned}$$

with initial conditions given in the paper of Valembois and Fossorier.

Recursive algorithm for calculating the SP59 bound

$$P_N(x) = \frac{2N-5+x^2}{N-1}\, P_{N-2}(x) - \frac{N-4}{N-1}\, P_{N-4}(x)\,,$$
$$Q_N(x) = \frac{2N-5+x^2}{N-1}\, Q_{N-2}(x) - \frac{N-4}{N-1}\, Q_{N-4}(x), \quad N \geq 5$$

- The coefficients of the higher powers of x decay to zero like $\frac{1}{N!}$.

Recursive algorithm for calculating the SP59 bound

$$P_N(x) = \frac{2N-5+x^2}{N-1}\, P_{N-2}(x) - \frac{N-4}{N-1}\, P_{N-4}(x)\,,$$
$$Q_N(x) = \frac{2N-5+x^2}{N-1}\, Q_{N-2}(x) - \frac{N-4}{N-1}\, Q_{N-4}(x), \quad N \geq 5$$

- The coefficients of the higher powers of x decay to zero like $\frac{1}{N!}$.

$$P_{\text{SPB}}(N,\theta,A) \triangleq \frac{(N-1)e^{-\frac{NA^2}{2}}}{\sqrt{2\pi}} \int_{\theta}^{\frac{\pi}{2}} (\sin\phi)^{N-2}\, f_N(\sqrt{N}A\cos\phi)\, d\phi + Q(\sqrt{N}A).$$

- f_N is evaluated at $O(\sqrt{N})$.

Recursive algorithm for calculating the SP59 bound

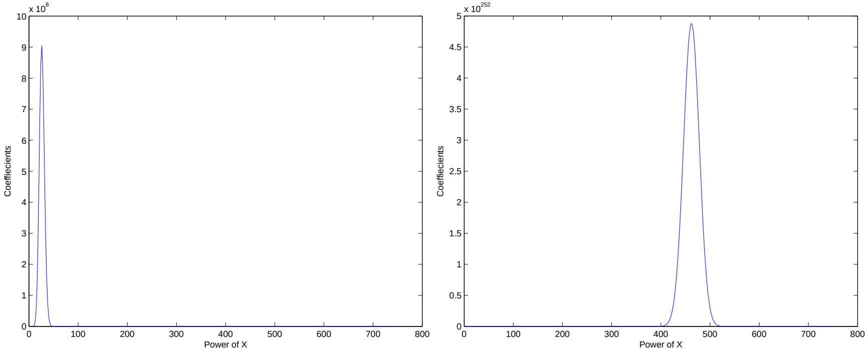
$$P_N(x) = \frac{2N-5+x^2}{N-1}\, P_{N-2}(x) - \frac{N-4}{N-1}\, P_{N-4}(x)\,,$$
$$Q_N(x) = \frac{2N-5+x^2}{N-1}\, Q_{N-2}(x) - \frac{N-4}{N-1}\, Q_{N-4}(x), \quad N \geq 5$$

- The coefficients of the higher powers of x decay to zero like $\frac{1}{N!}$.

$$P_{\text{SPB}}(N,\theta,A) \triangleq \frac{(N-1)e^{-\frac{NA^2}{2}}}{\sqrt{2\pi}} \int_{\theta}^{\frac{\pi}{2}} (\sin\phi)^{N-2}\, f_N(\sqrt{N}A\cos\phi)\, d\phi + Q(\sqrt{N}A).$$

- f_N is evaluated at $O(\sqrt{N})$.
- For block lengths of more than several hundreds the high coefficients cause underflows.
- This causes considerable inaccuracy in the calculations.

Example: Coefficients of $P_N(x)$ and $\widetilde{P}_N(x) \triangleq P_N(\sqrt{N}x)$ for $N = 750$



Motivation

- Existing algorithms for calculating the SP59 bound perform the calculation in the probability domain.
  $\Rightarrow$ The calculation is numerically unstable and causes over and underflows when the block length exceeds several hundreds.
- Asymptotic approximations performs the calculation in the logarithmic domain, avoiding numerical problems.
- For moderate block lengths, the approximations do not provide an accurate assessment of the bound.

Question

Is it possible to avoid the numerical difficulties associated with the calculation of the SP59 bound; e.g. by working in the log-domain?

Log-Domain Calculation of the SP59 Bound

- We present an algorithm which performs the *exact* calculation of the bound entirely in the logarithmic domain.
- Eliminates virtually all the numerical problems in the calculation of the SP59 bound for moderate to large block lengths.

Theorem (Log domain calculation of the SP59 bound)

The term $P_{\mathrm{SPB}}(N,\theta,A)$ can be rewritten as

$$\begin{aligned} P_{\mathrm{SPB}}(N,\theta,A) \quad = \quad & \int_{\theta}^{\frac{\pi}{2}} \exp\Bigg[\ln(N-1) - \frac{NA^2}{2} - \frac{1}{2}\ln(2\pi) + (N-2)\ln\sin\phi \\ & + \max{}^{*}\Big(d(N,0,\sqrt{N}A\cos\phi),\ldots,d(N,N-1,\sqrt{N}A\cos\phi)\Big)\Bigg]d\phi \\ & + Q(\sqrt{N}A)\,, \qquad N\in\mathbb{N},\ \theta\in[0,\frac{\pi}{2}],\ A\in\mathbb{R}^{+} \end{aligned}$$

Theorem (Log domain calculation of the SP59 bound)

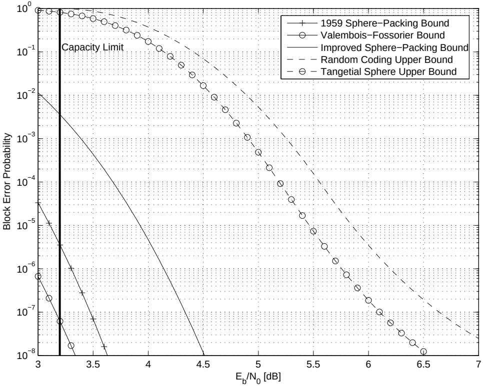
where

$$\begin{aligned} d(N,j,x) \triangleq\ & \frac{x^2}{2} + \ln\Gamma\left(\frac{N}{2}\right) - \ln\Gamma\left(\frac{j}{2}+1\right) - \ln\Gamma(N-j) + \\ & +(N-1-j)\ln\left(\sqrt{2}\,x\right) - \frac{\ln 2}{2} \\ & +\ln\left[1+(-1)^j\,\tilde{\gamma}\left(\frac{x^2}{2},\frac{j+1}{2}\right)\right], \qquad \begin{array}{c} N\in\mathbb{N},\ x\in\mathbb{R} \\ j=0,1\ldots,N-1 \end{array} \end{aligned}$$

- $\Gamma(\cdot)$ and $\tilde{\gamma}(\cdot,\cdot)$ designate the complete and incomplete (regularized) Gamma functions, respectively.
- $\max{}^*(x_1,\ldots,x_m) \triangleq \ln\left(\sum_{i=1}^m e^{x_i}\right)$ is the function commonly used in the log-domain implementation of the BCJR algorithm.

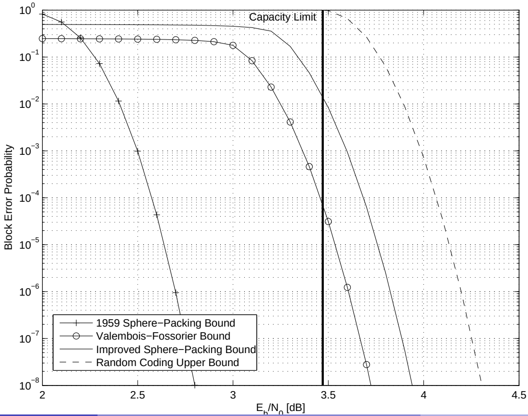
Numerical Results

- Transmission over a QPSK modulated AWGN channel
- $N = 300$ bits (150 channel symbols), $R = 1.8 \frac{\text{bits}}{\text{channel use}}$.
- The ISP bound gives an improvement of 0.7 and 1 dB over the SP59 and VF bounds, respectively. The SP59 bound refers to the 1959 sphere-packing bound of Shannon for the Gaussian channel.



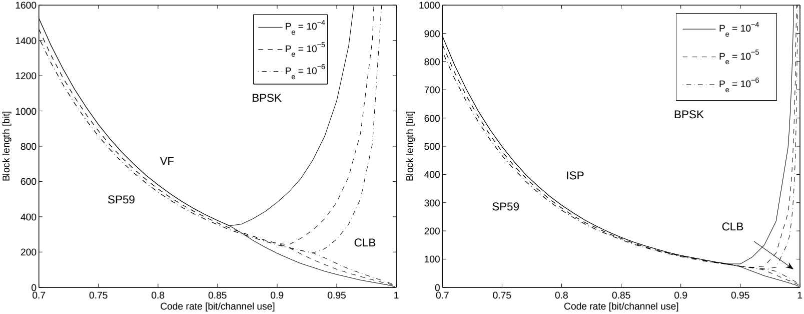
Numerical Results

- Transmission over a 8-PSK modulated AWGN channel
- $N = 5580$ bits (1680 channel symbols), $R = 2.2 \frac{\text{bits}}{\text{channel use}}$.
- ISP bound gives an improvement of 0.2 dB over the VF bound.
- 0.4 dB gap between ISP lower bound and random-coding upper bound.



Comparison of Bounds

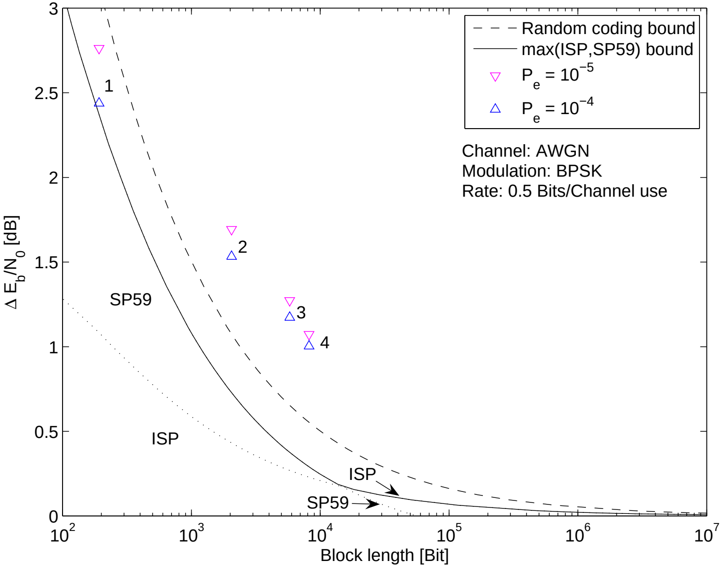
- Two-dimensional regions of N and R where VF bound (left plot), and the ISP bound (right plot) outperform the SP59 bound and the capacity-limit bound.
- The plots refer to BPSK modulation and focus on high code rates.



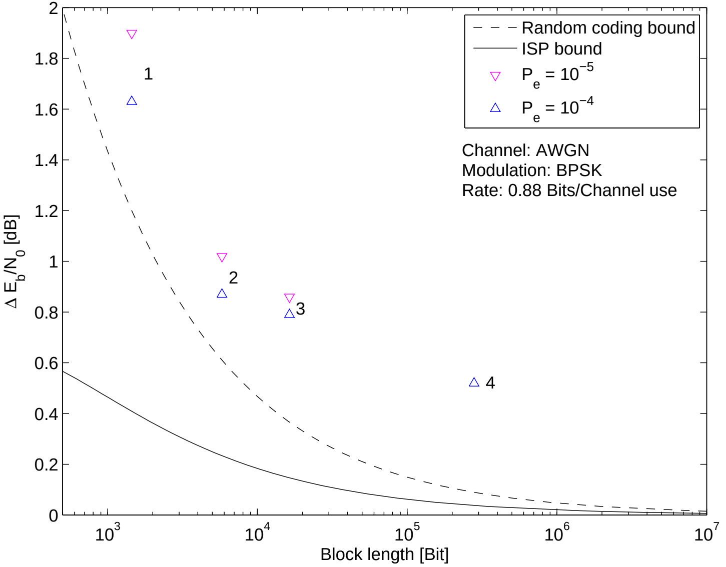
Minimal Block Length as a Function of Performance

- Fixing:
  1. the communication channel model
  2. the code rate
  3. the block error probability
- Sphere-packing bounds $\Rightarrow$ lower bounds on the minimal block length required to achieve the desired performance on the given channel using an arbitrary block code and decoding algorithm.
- Random coding bound of Gallager $\Rightarrow$ upper bound on the block length required for ML decoded random codes to achieve the desired performance on the given communication channel.

Comparison of upper and lower bounds on the block the length with the performance of iteratively decoded codes.



Comparison of upper and lower bounds on the block the length with the performance of iteratively decoded codes.



Summary

- We introduce an improved sphere-packing (ISP) bound for finite-length codes whose transmission takes place over symmetric memoryless channels.
- The ISP bound is uniformly tighter than the SP67 and VF bounds, especially for codes of short to moderate block lengths.
- Applications of the ISP bound are exemplified.

Summary (Cont.)

- An algorithm which performs the calculation of the SP59 bound in the logarithmic domain is introduced.
- The new algorithm is numerically stable and enables to perform the calculation for arbitrary block lengths.
- The ISP bound provides an interesting alternative to the sphere-packing bound of Shannon for the Gaussian channel, especially for high code rates.
- The sphere-packing bounds are employed as lower bounds on minimal block length required to achieve a desired performance on a given channel model.

Full Paper Version

Further results and proofs are provided in the full paper version:

- G. Wiechman and I. Sason, "An Improved Sphere-Packing Bound for Finite-Length Codes on Symmetric Channels," submitted to *IEEE Trans. on Information Theory*, March 2007.
- Available at http://arxiv.org/abs/cs.IT/0608042.