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

Gil Wiechman Igal Sason

Department of Electrical Engineering
Technion, Haifa, Israel

2008 Information Theory and Applications (ITA) Workshop
UCSD, San-Diego
January 28, 2008.

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
- Based on geometrical properties of the decoding regions.
- Decay to zero exponentially with the block length.

The 1967 sphere-packing (SP67) bound (Shannon, Gallager & Berlekamp)

- Applies to codes transmitted over discrete memoryless channels (DMCs).
- Valid under optimal ML decoding or even under list decoding.
- Error exponent is exact between the critical rate and the channel capacity.

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.

In order to consider possible improvements of the SP67 bound for finite-length codes, we first outline the original derivation of this bound.

Derivation of the 1967 Sphere-Packing Bound

Step 1: Lower bound on the error prob. 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 ($i = 1, 2$).
- $\mathcal{Y}_i$ - Decoding region of $\mathbf{x}_i$ ($\mathcal{Y}_2 = \mathcal{Y}_1^c$).
- By considering typical output sequences w.r.t. a certain prob. distribution, it was shown that for all $s \in (0, 1)$

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

or

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

where $\mu(s) \triangleq \ln\left(\sum_{\mathbf{y}} P_1(\mathbf{y})^{1-s} P_2(\mathbf{y})^s\right), \quad 0 < s < 1.$

Step 2: Fixed-composition block codes

- Consider a fixed-composition block code containing M codewords of length N, transmitted over a DMC, and decoded by a list decoder of size L.
- An arbitrary memoryless probability measure F_N over channel output vectors of length N is introduced. It is used to measure the size of the decoding regions.
- It is easy to show that there is a codeword $\mathbf{x}_m$ so that the size of its decoding region does not exceed $\frac{L}{M}$.

Step 2: Fixed-composition block codes

- Use in Step 1 with the setting: $P_1(\mathbf{y}) = P_N(\mathbf{y}|\mathbf{x}_m)$, $P_2(\mathbf{y}) = F_N(\mathbf{y})$. $\Rightarrow$ Step 1 provides a lower bound on the conditional error prob. of this codeword for all values of s for which the inequality related to the size of the decoding region is violated (we do not know the exact size of this set, but it is upper bounded by $\frac{L}{M}$).
- Since we do not know this specific codeword, we replace its conditional error probability by an upper bound which is the maximal error probability (over all codewords).

To achieve the tightest universal lower bound:

1. Find F_N which *maximizes* the lower bound.
2. Find the composition which *minimizes* the lower bound.

Step 3: Lower bound on the average error prob. of general block codes

Proposition: 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.

Proof outline:

- 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.
- 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.

Combine this conclusion with the lower bound in Step 2 for fixed composition codes, and the SP67 bound follows.

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

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

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

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

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

Theorem (The 1967 Sphere-Packing Bound)

where

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

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.

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

$$P_e(N,M,L) \geq \exp\Big\{-NE_{\mathrm{VF}}(R,N)\Big\}$$

where

$$E_{\mathrm{VF}}(R,N) \triangleq \inf_{x>\frac{\sqrt{2}}{2}} \left\{E_0(\rho_x) - \rho_x\left(R - O_1\left(\frac{\ln N}{N},x\right)\right) + O_2\left(\frac{1}{\sqrt{N}},x,\rho_x\right)\right\}.$$

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 $P_{\min} > 0$, 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 of 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.

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 the *average* error probability,
     eliminating the need for expurgation of half of 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

$$P_{\mathrm{e}}(N,M,L) \geq \exp\Big\{-NE_{\mathrm{ISP}}(R,N)\Big\}$$

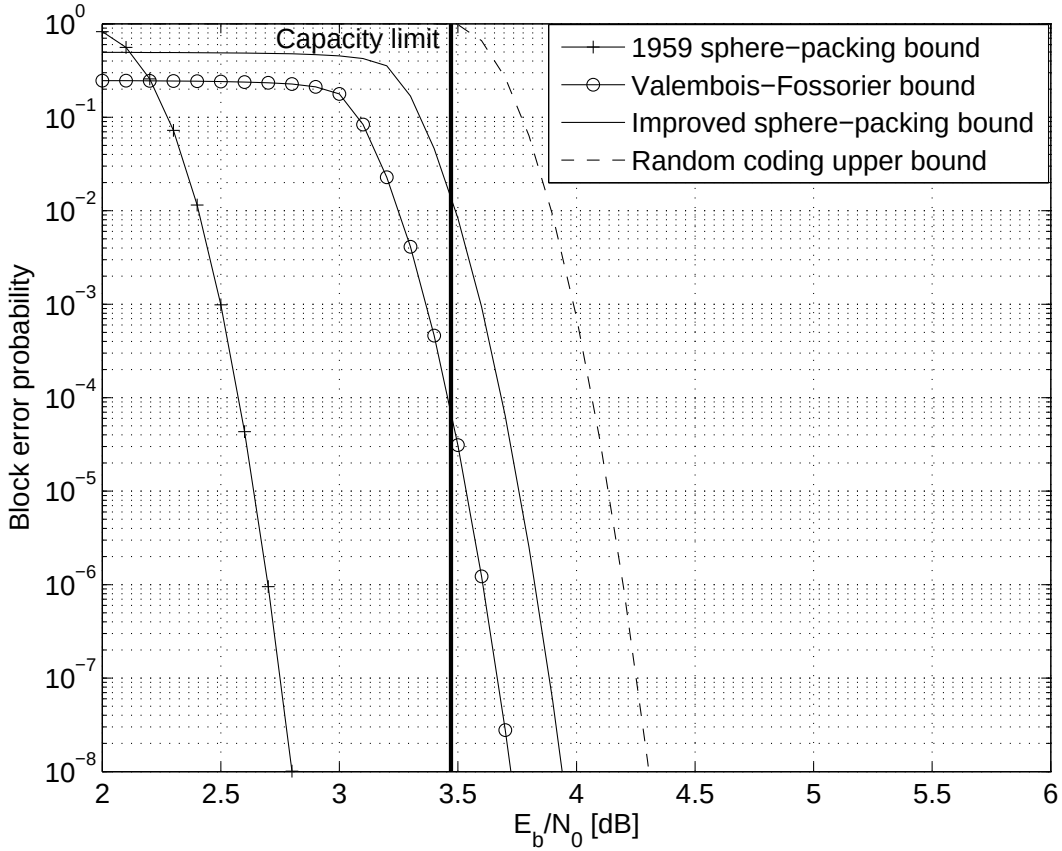
where

$$E_{ISP}(R,N) \triangleq \inf_{x>\frac{\sqrt{2}}{2}} \left\{E_0(\rho_x) - \rho_x\left(R - O_1\left(\frac{1}{N},x\right)\right) + O_2\left(\frac{1}{\sqrt{N}},x,\rho_x\right)\right\}$$

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.

Numerical Results

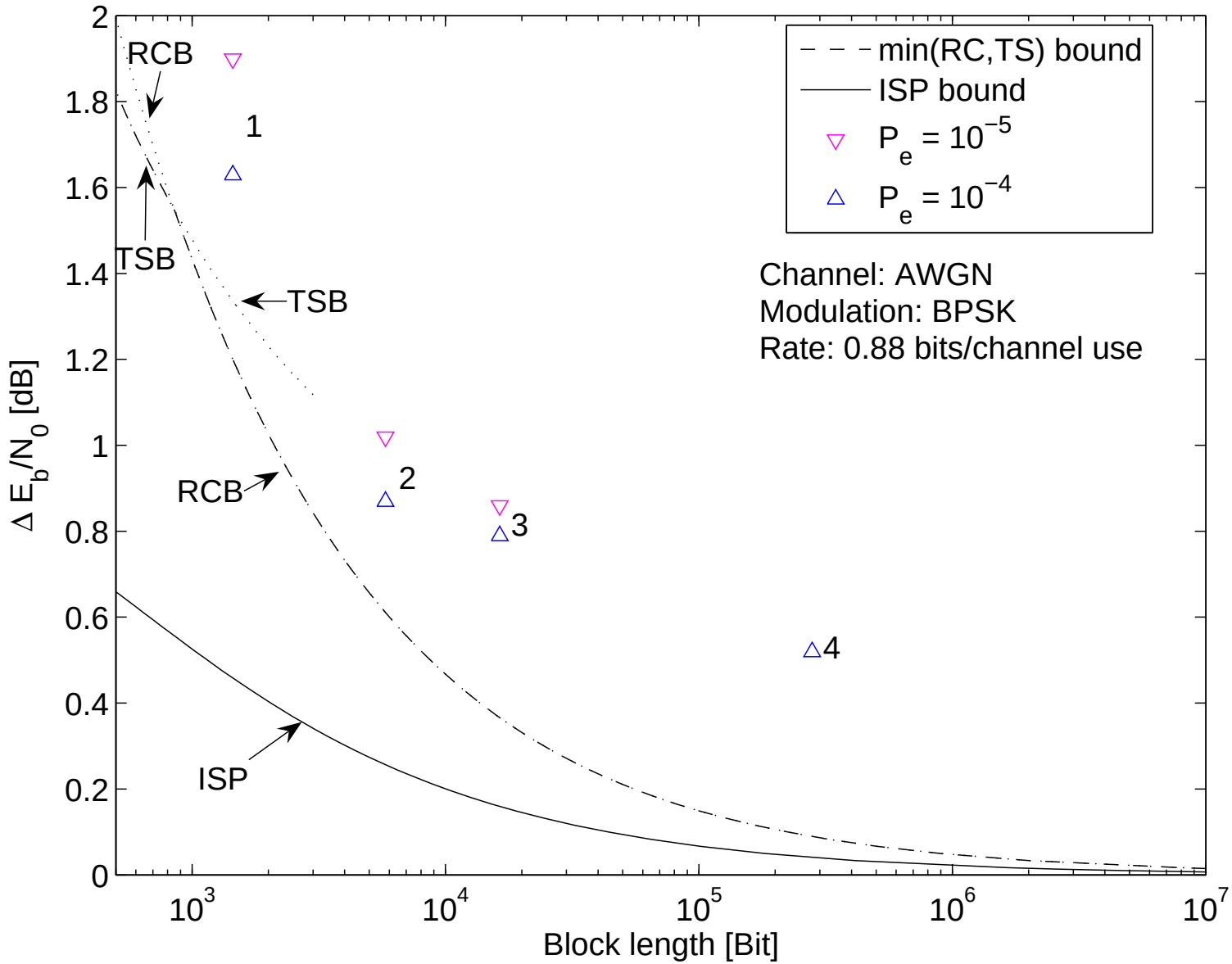


- Transmission over a 8-PSK modulated AWGN channel
- $N = 5580$ bits (1680 channel symbols), $R = 2.2 \frac{\text{bits}}{\text{channel use}}$.
- ISP bound gives an improvement of 0.2 dB over the VF bound.
- 0.4 dB gap between ISP lower bound and random-coding upper bound.

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.
- Upper bounds on the error prob. of random codes $\Rightarrow$ upper bounds 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.



Summary

- We introduce an improved sphere-packing (ISP) bound for finite-length codes whose transmission takes place over symmetric memoryless channels.

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.

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

- 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.
- The ISP bound provides an interesting alternative to the sphere-packing bound of Shannon for the Gaussian channel, especially for high code rates.

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.
- 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.

Journal Paper

G. Wiechman and I. Sason, “An improved sphere-packing bound for finite-length codes on symmetric memoryless channels,” submitted to *IEEE Trans. on Information Theory*, March 2007. [Online]. Available: `http://www.arxiv.org/abs/cs.IT/0608042`.
Latest version available at:
`http://www.ee.technion.ac.il/people/sason/ISP.pdf`.

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

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

and

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

Symmetry conditions

- 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 coherently detected M-ary PSK modulated signals, over fully-interleaved fading channels, when the decoder has full knowledge of the fading samples.