Variations on the Gallager Bounds, Connections, and Applications

Igal Sason
Department of Electrical Engineering, Technion
Haifa 32000, Israel
Sason@ee.technion.ac.il

Joint work with Shlomo Shamai
The error performance of coded communication systems rarely admits exact expressions ⇒ Tight analytical bounds emerge as a useful tool for assessing performance.

For long enough block codes, the union bound is useless at rates above the cutoff rate \( R_0 \) of the channel ⇒ A need for improved upper bounds on the decoding error probability, which are not subject to the cutoff rate limitation for long enough codes.
- The motivation for introducing and applying tight bounds on the decoding error probability has increased with the discovery of the turbo-like codes.

- The optimal ML decoding is prohibitively complex for long enough codes. However, the derivation of bounds on the ML decoding is of interest, as it provides an ultimate indication on the system performance.
The fine structure of efficient codes is usually not available!
⇒ Efficient ML bounds rely only on basic features, such as the distance spectrum of the codes.

It is desirable that efficient bounding techniques encompass both specific codes as well as ensembles of structured codes.
The Gallager (1965) Bound

Fixed Codes – Maximum Likelihood (ML) decoding:

\[ P_{e|m} \leq \sum_y p_N(y|x^m) \left( \sum_{m' \neq m} \left( \frac{p_N(y|x^{m'})}{p_N(y|x^m)} \right)^\lambda \right)^\rho \]

\[ \lambda, \rho \geq 0 \]

- \( P_{e|m} \) – block error probability conditioned on the transmitted code word \( x^m \) (\( m = 1, 2, \ldots, M \)).
- \( x^m \) – the transmitted length \(-N\) code word.
- \( y \) – the observation vector (\( N \) components).
- \( p_N(y|x) \) – the channel’s transition probability measure (for a block of length \( N \)).
The Gallager (1965) Bound (Cont.)

- Usually **impractical** to evaluate in terms of basic features of particular codes, *but* for example, orthogonal codes and the special case where $\rho = 1, \lambda = 1/2$, which yields the Bhattacharyya-Union bound.
Gallager ’65 Bound: Random Codes

- Select each codeword independently by the distribution \( q_N(x) \).
- By applying the Jensen inequality: \( E[x^\rho] \leq (E[x])^\rho \) \( (0 \leq \rho \leq 1) \), and setting \( \lambda = \frac{1}{1+\rho} \), the Gallager random coding bound results in

\[
P_e \leq (M - 1)^\rho \sum_y \left( \sum_x q_N(x)p_N(y|x)^{1/1+\rho} \right)^{1+\rho}, \ 0 \leq \rho \leq 1
\]

where

- \( P_e \) – the average ML decoding error probability
- \( M \) – the number of codewords.
The Gallager 1965 Bound: Random Codes (Cont.)

- Memoryless channel: \( p_N(y|x) = \prod_{\ell=1}^{N} p(y_\ell|x_\ell) \).
- Memoryless input-distribution: \( q_N(x) = \prod_{\ell=1}^{N} q(x_\ell) \).

The 1965 Gallager random coding bound reads

\[
P_e \leq e^{-N \cdot E(R,q)}
\]

\[
E(R,q) = \ln \frac{M}{N} - \text{code rate}
\]

\[
E_0(\rho, q) = -\ln \left( \sum_y \left( \sum_x q(x) p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right)
\]

\[
E(R, q) = \max_{0 \leq \rho \leq 1} \left( E_0(\rho, q) - \rho R \right).
\]
Generalization of the Duman & Salehi Variation (1998)

- Suggested by Sason & Shamai as a generalization (second version) of the Duman and Salehi bound which was originally derived for the AWGN channel (Duman’s Ph.D. dissertation 1998).

- Let $\psi_{m,N}^N(y)$ be a measure (may depend on $x^m$).
The Duman & Salehi Variation (Cont.)

\[ P_{e|m} \leq \sum_{y} \psi_{N}^{m}(y)\psi_{N}^{m}(y)^{-1}p_{N}(y|x^{m}) \left( \sum_{m' \neq m} \left( \frac{p_{N}(y|x^{m'})}{p_{N}(y|x^{m})} \right)^{\lambda} \right)^{\rho} \]

\[ \lambda, \rho \geq 0 \]

\[ = \sum_{y} \psi_{N}^{m}(y) \left( \psi_{N}^{m}(y)^{-1}p_{N}(y|x^{m})^{1/\rho} \left( \sum_{m' \neq m} \left( \frac{p_{N}(y|x^{m'})}{p_{N}(y|x^{m})} \right)^{\lambda} \right) \right)^{\rho} \]

\[ \leq \text{Jensen} \left( \sum_{m' \neq m} \sum_{y} p_{N}(y|x^{m})^{1/\rho} \psi_{N}^{m}(y)^{(1-1/\rho)} \left( \frac{p_{N}(y|x^{m'})}{p_{N}(y|x^{m})} \right)^{\lambda} \right)^{\rho} \]

\[ 0 \leq \rho \leq 1, \ \lambda \geq 0. \]
Let
\[ \psi_m^N(y) = \frac{G^m_N(y) p_N(y | x^m)}{\sum_y G^m_N(y) p_N(y | x^m)} \]

\[ \Rightarrow \]

\[ P_{e|m} \leq \left( \sum_y G^m_N(y) p_N(y | x^m) \right)^{1-\rho} \]

\[ \left\{ \sum_{m' \neq m} \sum_y p_N(y | x^m) G^m_N(y)^{1-\frac{1}{\rho}} \left( \frac{p_N(y | x^{m'})}{p_N(y | x^m)} \right)^{\lambda} \right\}^\rho \]

\[ 0 \leq \rho \leq 1 \]

\[ \lambda \geq 0 . \]
**Advantage:** relatively *easy* to evaluate (in parallel to standard union bounds) for particular block codes in terms of basic code features as the distance-spectrum, *relaxing* the need of the fine details of the code structure.
Applications of the Generalization of the Duman & Salehi Bound

- Performance Evaluation of Codes on the AWGN Channel.
- Performance Evaluation on Fully interleaved and Block-Fading Channels.
- Performance Evaluation for Mismatched Decoding.
- Finite state channels.
- General alphabets
  
  and many more applications....
Application 1: AWGN Channel

We compare here between various improved bounds on the ML decoding error probability for the binary-input AWGN channel, referring to the following ensemble of *Uniformly Interleaved Turbo Codes*:
Application 1: AWGN Channel (cont.)
Application 2: A Memoryless (Fully Interleaved) Fading Channel

\[ y = ax + n \]

\( n \) - AWGN (normalized).
\( x \) - binary \( \{0,1\} \rightarrow \pm \sqrt{\text{SNR}} \).

\( a \) — fading amplitude coefficient (normalized Rayleigh, Rician, etc.)
  i.i.d. (fully interleaved channel) and available ideally to the receiver.

\( y \) — observation.
Application 2: Fading Channel (Cont.)

Duman & Salehi generalized version [Sason & Shamai 2001]

\[ P_e \leq \sum_{d=d_{\text{min}}}^{N} P_e(d) \]

⇒ code separation to equi-distance \(d\) (w.r.t. codeword) subcodes + Union.

\[
P_e(d) \leq (S_d)^{\rho} \left\{ \left( \int_{-\infty}^{\infty} dy \int_{0}^{\infty} da \psi(y, a)^{1-\frac{1}{\rho}} p(y, a|0)^{\frac{1}{\rho}} \right)^{1-\frac{d}{N}} \right\}^{N} \]

\[
\left( \int_{-\infty}^{\infty} dy \int_{0}^{\infty} da \psi(y, a)^{1-\frac{1}{\rho}} p(y, a|0)^{\frac{1-\lambda\rho}{\rho}} p(y, a|1)^{\lambda} \right)^{\rho d_{N}}
\]

\[ 0 \leq \rho \leq 1, \quad \lambda \geq 0. \]
A Fully interleaved Rayleigh fading channel + MRC diversity of order $L=6$ (with perfect CSI). Uniformly Interleaved Turbo Codes, Interleaver length=1000, rate 1/3

1. Duman and Salehi generalized bound [Sason & Shamai 2002].
3. Union bound & Comparison to 10 iterations of a practical Log-MAP algorithm.
Variations on the Gallager Bounds and Connections

- The generalized version of the Duman & Salehi bound is a rather strong upper bounding technique which fits for specific codes (giving results in terms of basic code parameters as the distance spectrum) as well as random codes.

- It is useful for matched and mismatched metrics, a variety of channels as well as general input and output alphabets.

- It is directly applicable to bit error probability calculations.

- Replace the distance spectrum \( \{ S_d \} (d = d_{\text{min}}, \ldots, N) \) by \( \sum_{w=1}^{K} \frac{w}{K} A_{w,d} \).

  \( A_{w,d} \) denotes the number of codewords of weight \( d \) and information weight \( w \).
Many useful bounds are *special cases*:

- Divsalar ’99, Viterbi & Viterbi ITW’98,
- Viterbi & Viterbi ISIT’98,
- Duman & Salehi version I ’98,
- Chernoff versions of:
  - Hughes ’94, Herzberg & Poltyrev ’94,
  - Dolinar & Ekroot & Polara ’94,
  - Engdahl & Zingangirov ’99,
  - Generalized Viterbi & Viterbi bound and the generalized Duman & Salehi (version I) for fully interleaved fading channels with perfect CSI [Sason, Shamai and Divsalar 2003].
The generalized version of the Duman and Salehi bound [Sason and Shamai, 2002] provides the link between the Gallager ’65 and the Gallager ’63 bounds, and it facilitates a geometric interpretation (Fano ’61 style) of the decision region tilting. Many upper bounds which were derived independently are actually some particular cases of this generalized bound.
Figure 1: A diagram which shows the interconnections among the various upper bounds on the error probability with the ML decoding algorithm.
For more information and applications of Gallager-type bounds and their variations, we refer the audience to our paper:

Further Reading (Cont.)

A tutorial paper in the subject is currently in preparation: