Stopping Criteria for Turbo Decoding and Turbo Codes for Burst Channels

Telex Magloire Ngatched

Centre for Radio Access Technologies
University Of Natal
Durban, South-Africa

Overview

- Introduction
- Turbo Encoder and Decoder
- Stopping Criteria for Turbo Decoding
- Comparison of Coding Systems
- Turbo Codes for Burst channels
- Conclusions

Introduction

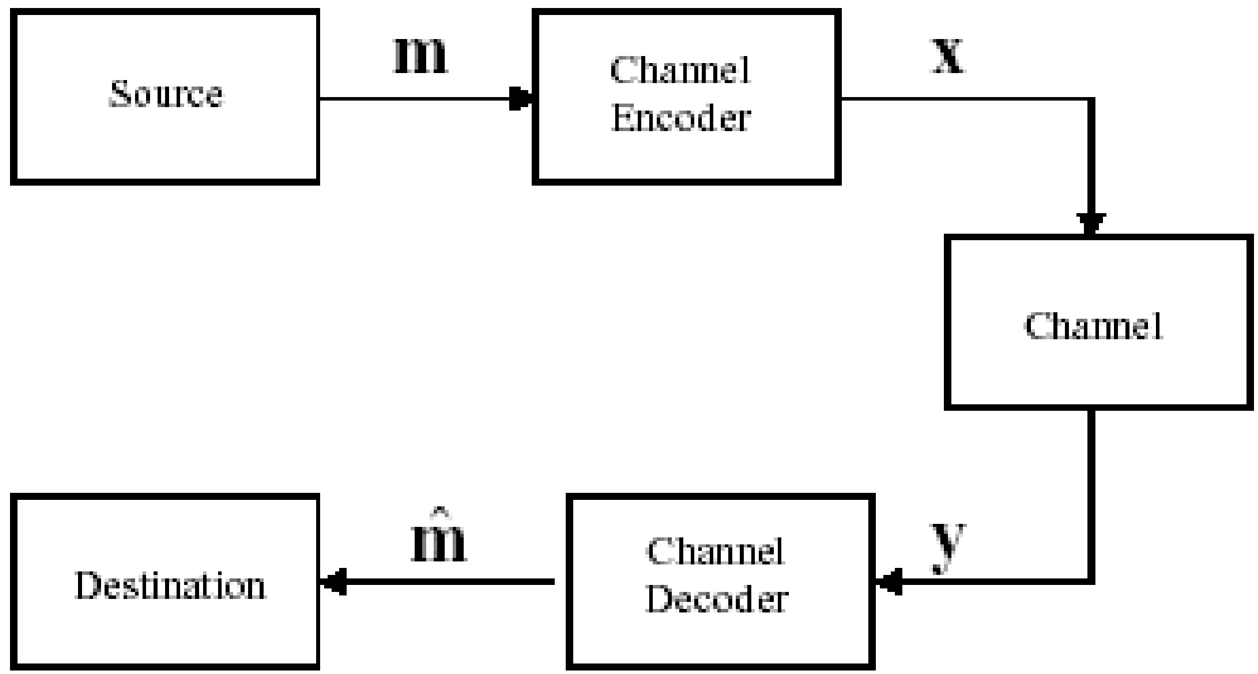


Fig. 1: Generalized Communication System

Turbo Encoder

- Turbo codes, introduced in June 1993, represent the most recent successful attempt in achieving Shannon's theoretical limit.

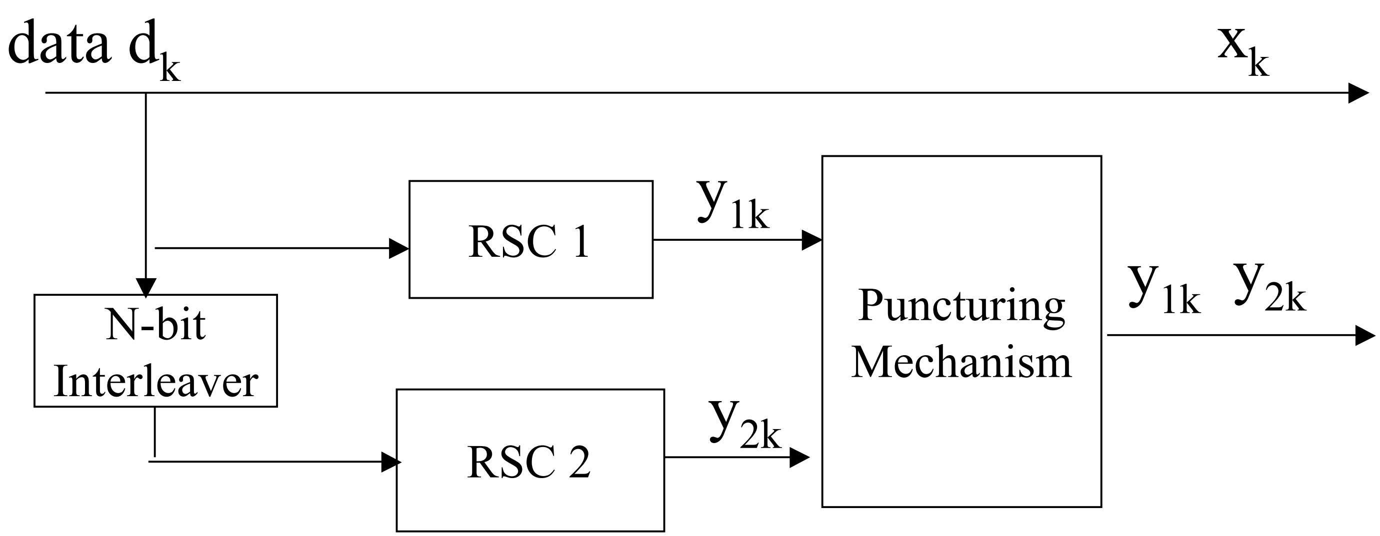


Fig. 2: A turbo encoder

Turbo Decoder

- Two component decoders are linked by interleavers in a structure similar to that of the encoder.

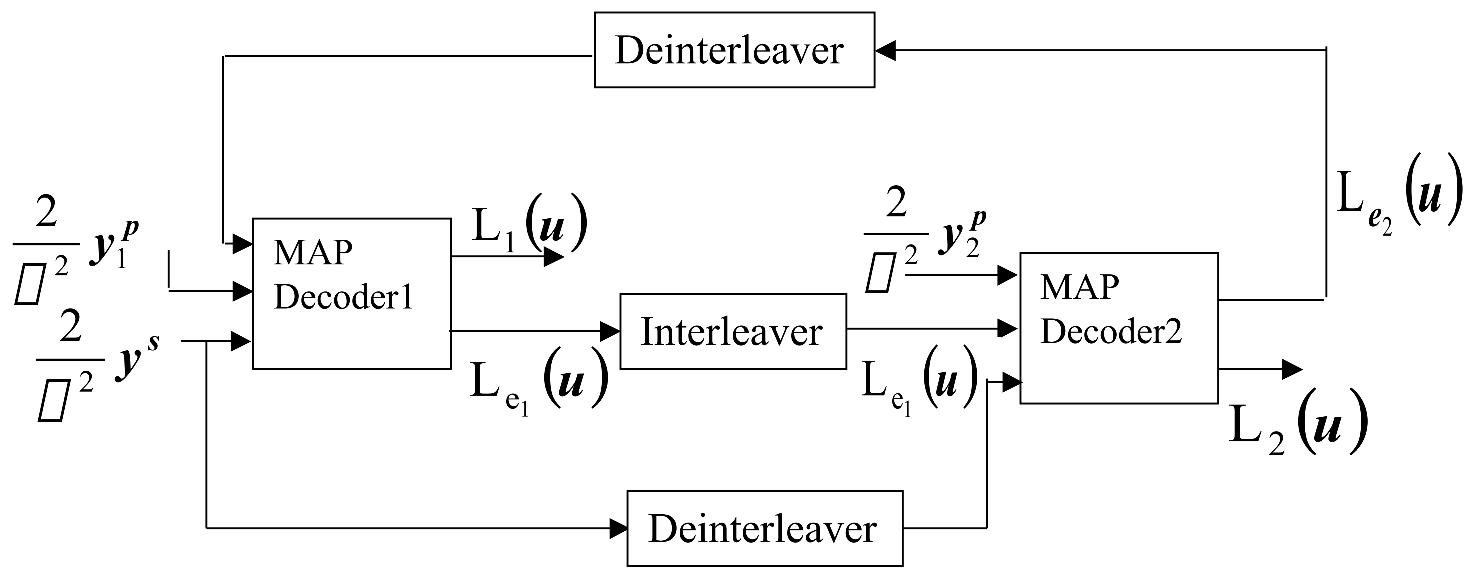


Fig. 3: Turbo Decoder

Turbo Decoder

- Each decoder takes three types of soft inputs
  - The received noisy information sequence.
  - The received noisy parity sequence transmitted from the associated component encoder.
  - The a priori information, which is the extrinsic information provided by the other component decoder from the previous step of decoding process.

Turbo Decoder

- The soft outputs generated by each constituent decoder also consist of three components:
  - A weighted version of the received information sequence
  - The a priori value, i.e. the previous extrinsic information
  - A newly generated extrinsic information, which is then provided as a priori for the next step of decoding.

$$\mathrm{L}_1^{(i)}(\hat{u}_k) = \mathrm{L}_{\mathrm{e}_2}^{(i-1)}(\hat{u}_k) + \frac{2}{\sigma^2} y_k^s + \mathrm{L}_{\mathrm{e}_1}^{(i)}(\hat{u}_k) \qquad (1)$$

$$\mathrm{L}_2^{(i)}(\hat{u}_k) = \mathrm{L}_{\mathrm{e}_1}^{(i)}(\hat{u}_k) + \frac{2}{\sigma^2} y_k^s + \mathrm{L}_{\mathrm{e}_2}^{(i)}(\hat{u}_k) \qquad (2)$$

Turbo Decoder

- The turbo decoder operates iteratively with ever-updating extrinsic information to be exchanged between the two decoder until a reliable hard decision can be made.
- Often, a fixed number, say M, is chosen and each frame is decoded for M iterations.
- Usually, M is set with the worst corrupted frames in mind.
- Most frames, however, need fewer iterations to converge
- It is therefore important to terminate the iterations for each individual frame immediately after the bits are correctly estimated

Stopping Criteria for Turbo Decoding

- Several schemes have been proposed to control the termination:
  - Cross Entropy (CE)
  - Sign Change ratio (SCR)
  - Hard Decision-Aided (HDA)
  - Sign Difference Ratio (SDR)
  - Improved Hard Decision-Aided (IHDA) (Ngatched scheme)

Stopping Criteria for Turbo Decoding

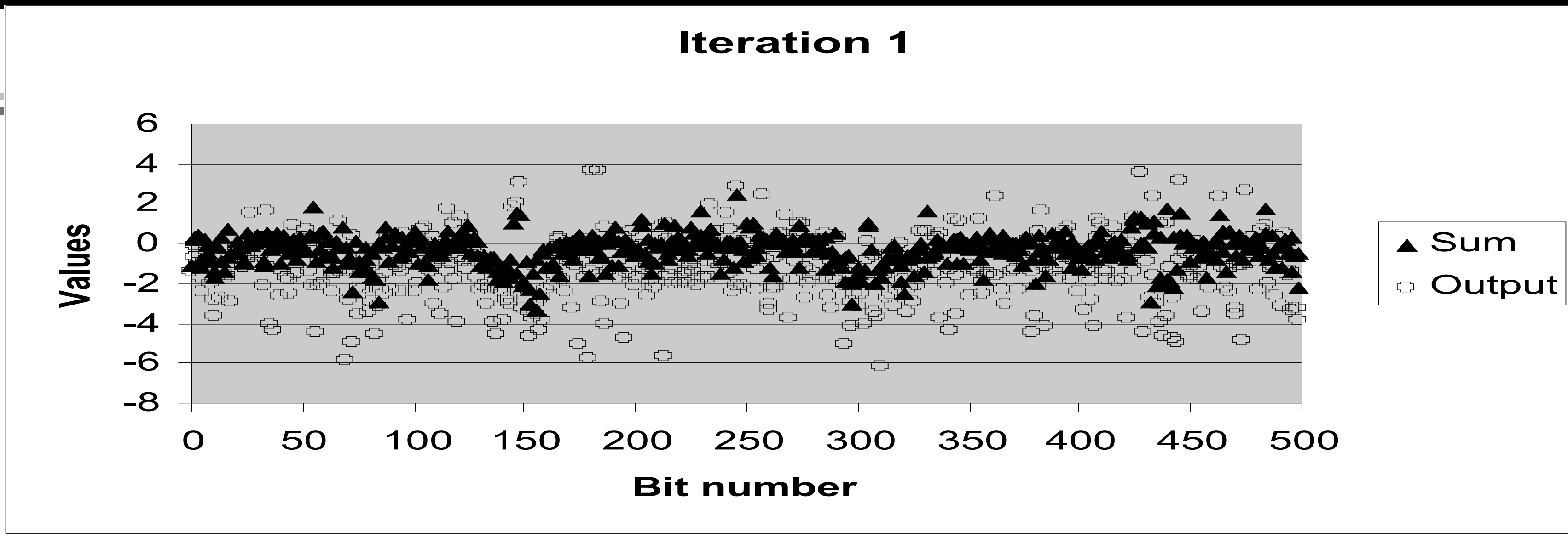
- Cross Entropy (CE)
  - Computes $T(i) \approx \sum_k \frac{\left|\Delta \mathrm{L}_{e_2}^{(i)}(\hat{u}_k)\right|^2}{\exp\left(\left|\mathrm{L}_1^{(i)}(\hat{u}_k)\right|\right)}$ (3)
  - Terminates when $T(i)$ drops to $\left(10^{-2} \text{ to } 10^{-4}\right) T(1)$
- Sign Change Ratio (SCR)
  - Computes $C(i)$ the number of sign changes of the extrinsic information from the second decoder between two consecutive iterations $(i-1)$ and i.
  - Terminates when $C(i) \le qN$, $0.005 \le q \le 0.03$; N is the frame size.

Stopping Criteria for Turbo Decoding

- Hard-Decision-Aided (HDA)
  - Terminates if the hard decision of the information bits based on $L_2^{(i-1)}(\hat{u}_k)$ at iteration $(i-1)$ agrees with the hard decision based on $L_2^{(i)}(\hat{u}_k)$ at iteration i for the entire block.
- Sign Difference Ratio (SDR)
  - Terminates at iteration i if the number of sign difference between $L_{ej}^{(i)}(\hat{u}_k)$, D_{ji}, satisfies $D_{ji} \le p \times N$ $0.001 \le p \le 0.01$, N is the frame size.

Stopping Criteria for Turbo Decoding

- The influence of each term on the *a-posteriori* LLR depends on whether the frame is “good” (easy to decode) or “bad” (hard to decode).
- For a “bad” frame, the *a-posteriori* LLR is greatly influenced by the channel soft output.
- For a “good” frame, the *a-posteriori* LLR is essentially determined by the *extrinsic information* as the decoding converges.
- These observations, together with equations (1) and (2) led us to the following stopping criterion.



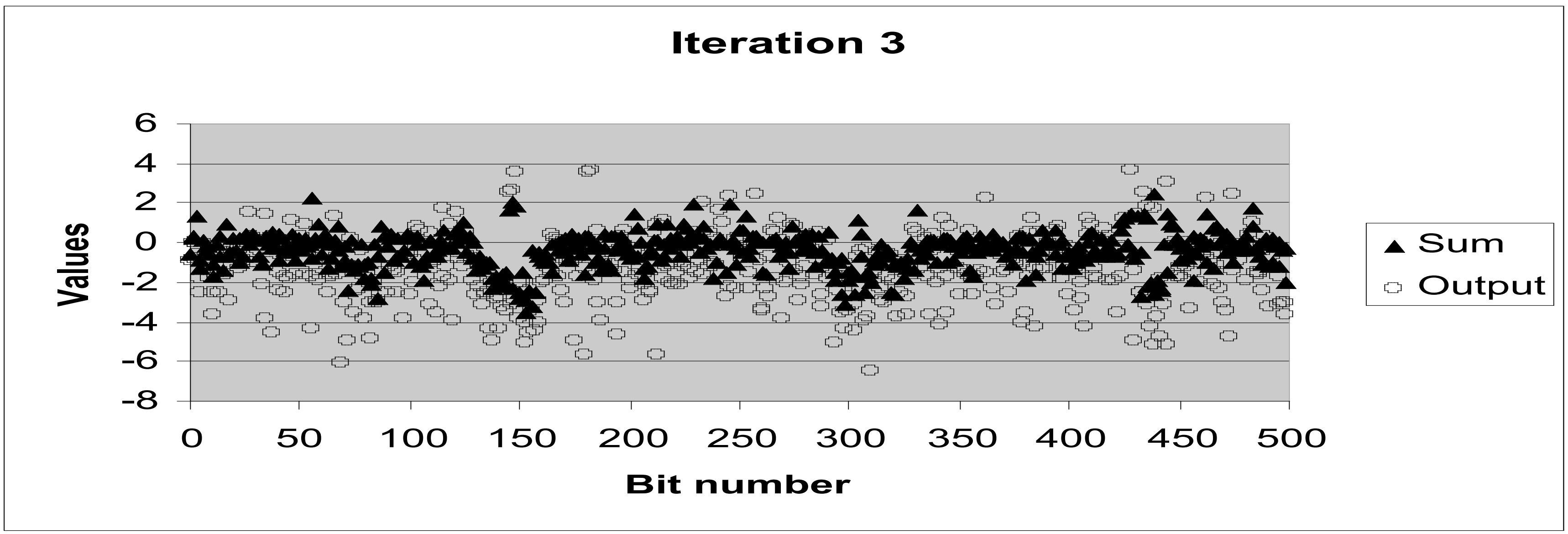
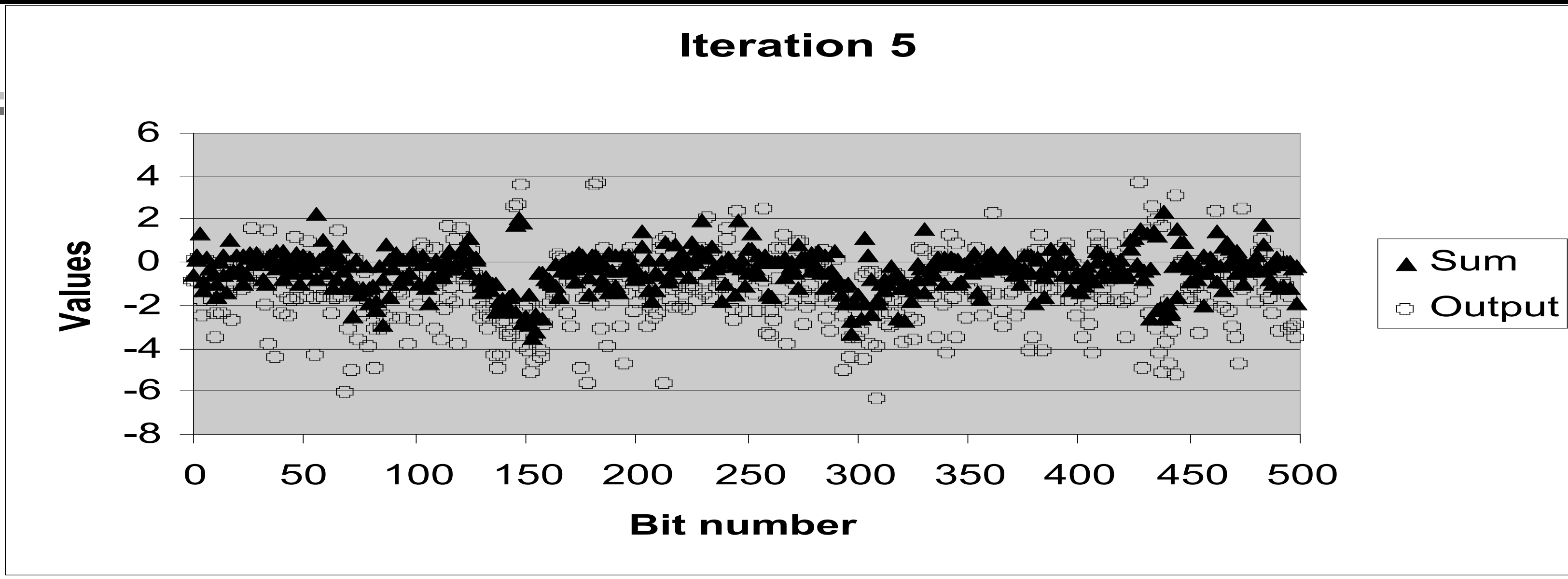


Fig. 4.a: Outputs from the decoder for a transmitted “bad” stream of -1



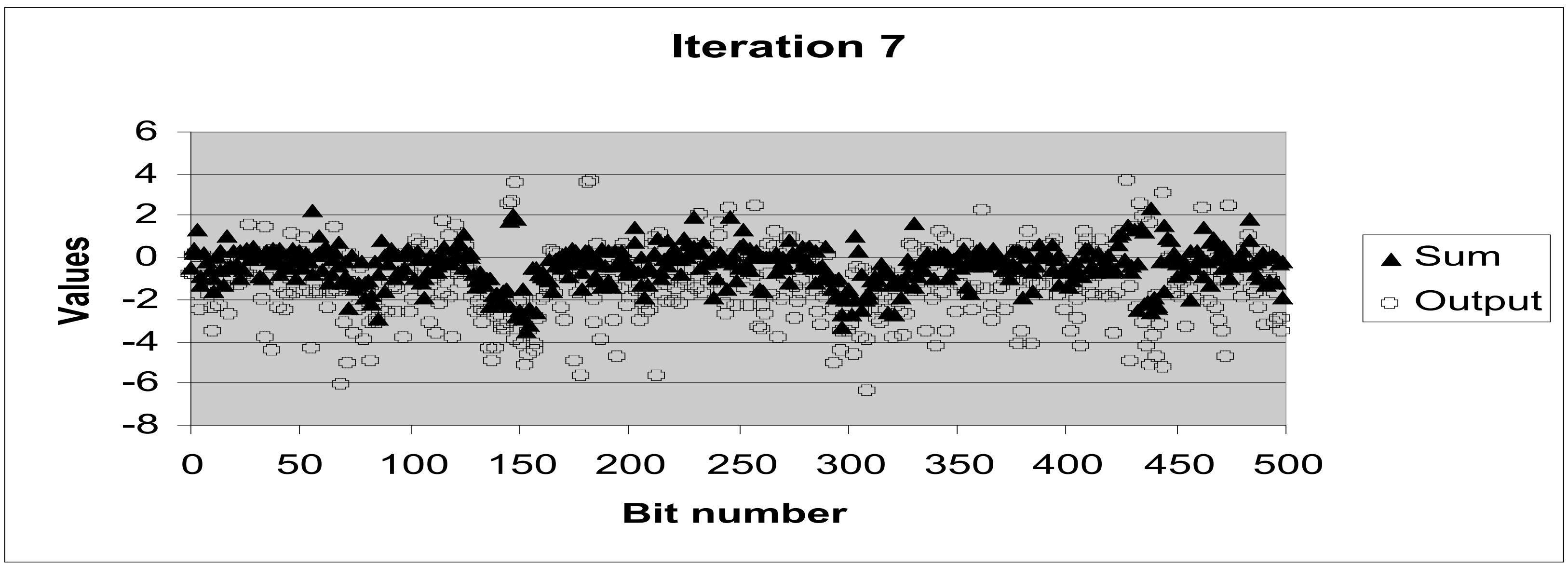


Fig. 4.b: Outputs from the decoder for a transmitted “bad” stream of -1

[redacted]

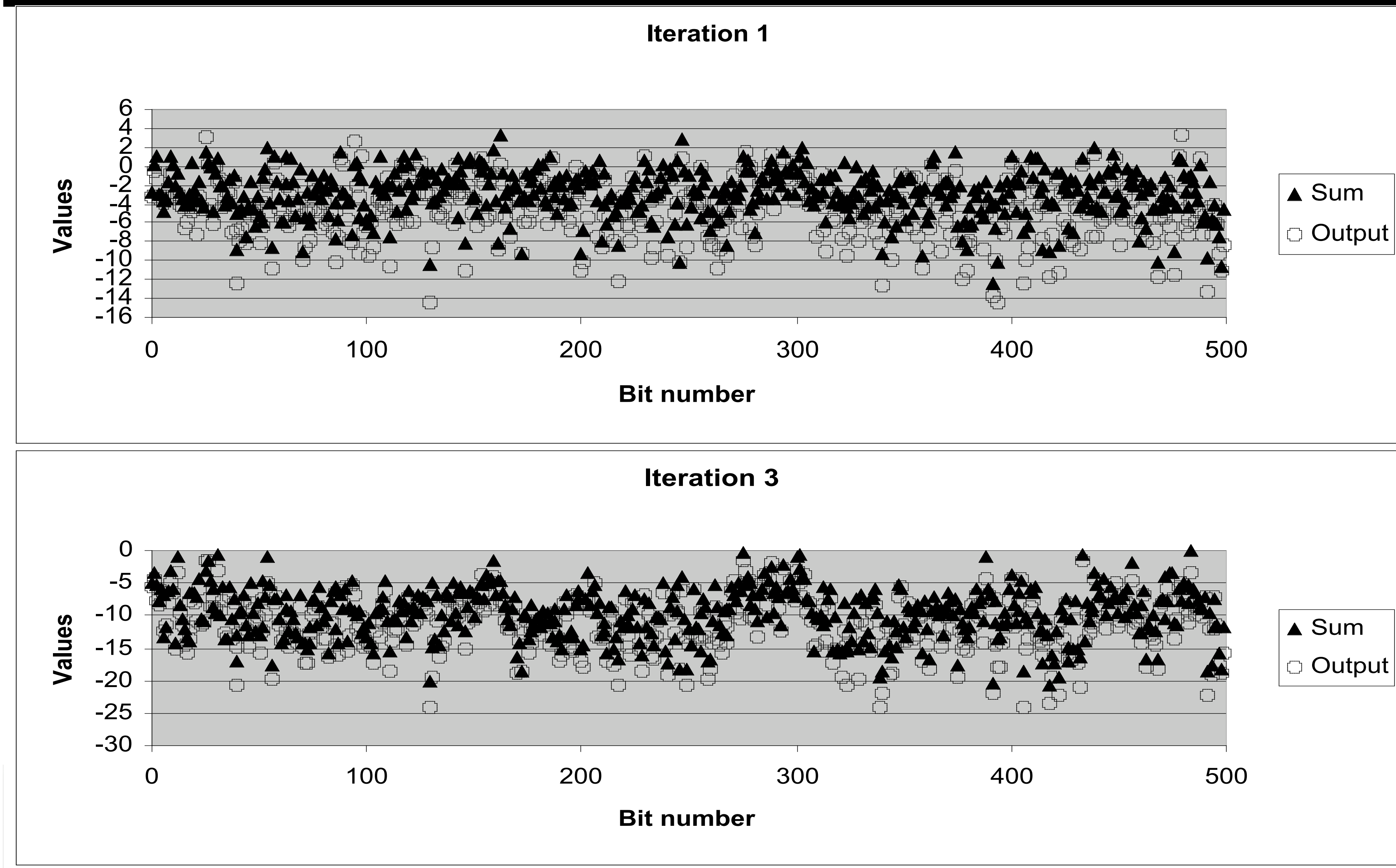


Fig. 5.a: Outputs from the decoder for a transmitted “good” stream of -1

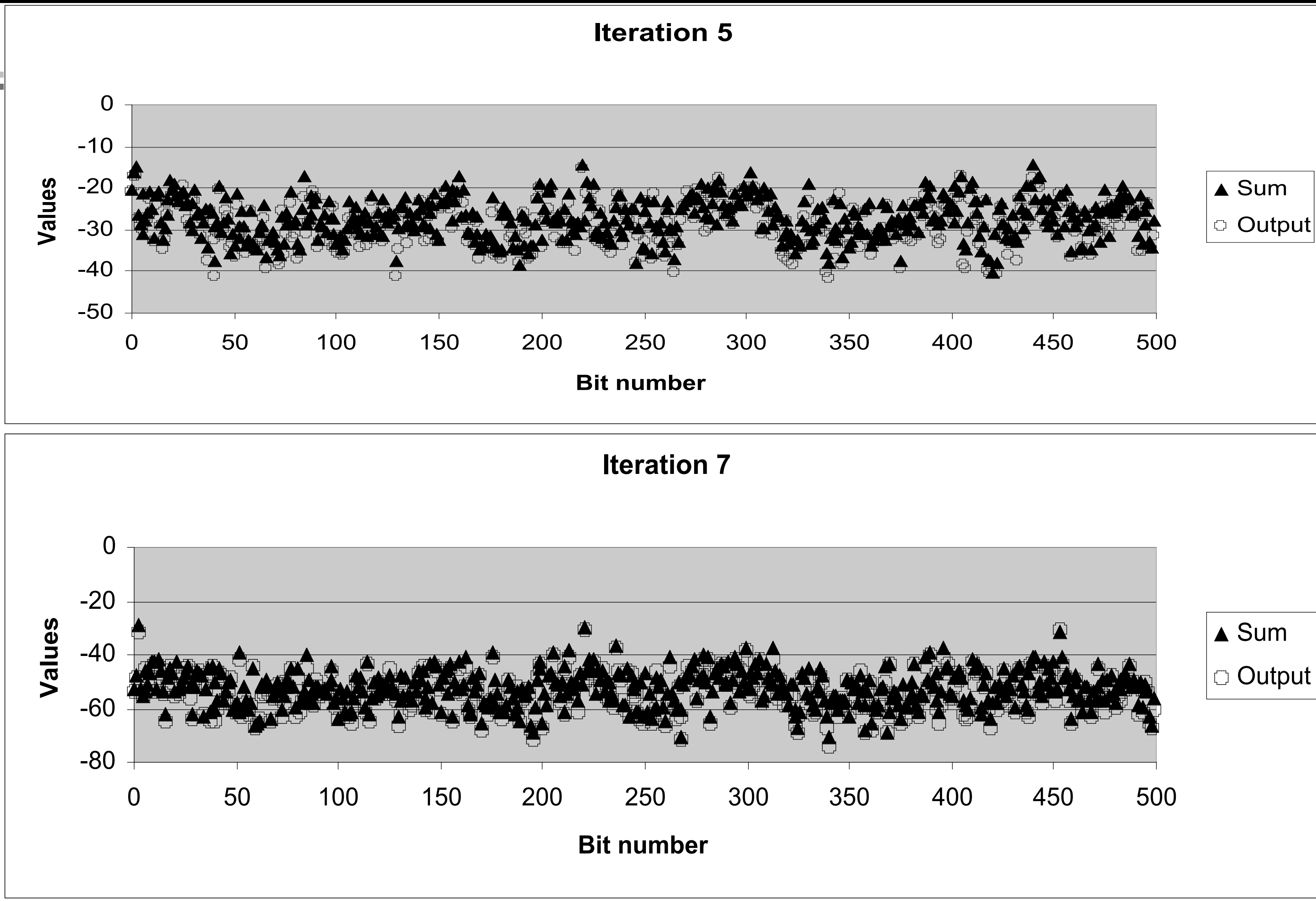


Fig. 5.b: Outputs from the decoder for a transmitted “good” stream of -1

Stopping Criteria for Turbo Decoding

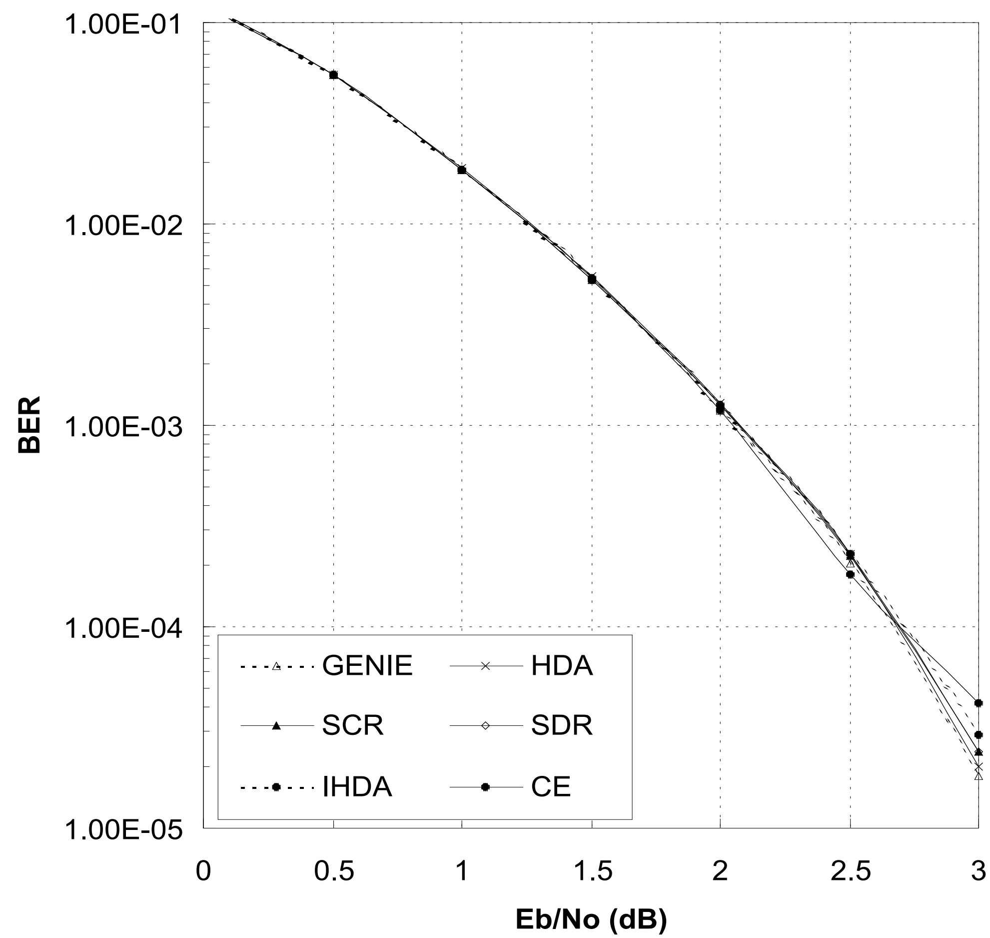
- Improved Hard-Decision-Aided (IHDA)
  - Terminates at iteration i if the hard decision of the information bit based on $\left(\frac{2}{\sigma^2} y_k^s + L_{e_1}^{(i)}(\hat{u}_k)\right)$ agrees with the hard decision of the information bit based on $L_2^{(i)}(\hat{u}_k)$ for the entire block.

Comparison of Stopping Criteria for Turbo Decoding

- Simulation Model
  - Code of rate 1/3, rate one-half RSC component encoders of memory 3 and octal generator (13, 15).
  - Frame size 128, AWGN channel.
  - MAP decoding algorithm with a maximum of 8 iterations.
  - Five terminating schemes are studied:
    - CE ($T(i)=10^{-3}T(1)$), SCR ($q=10^{-3}$), HAD, SDR ($p=10^{-4}$) and IHDA.
  - The "GENIE" case is shown as the limit of all possible schemes.

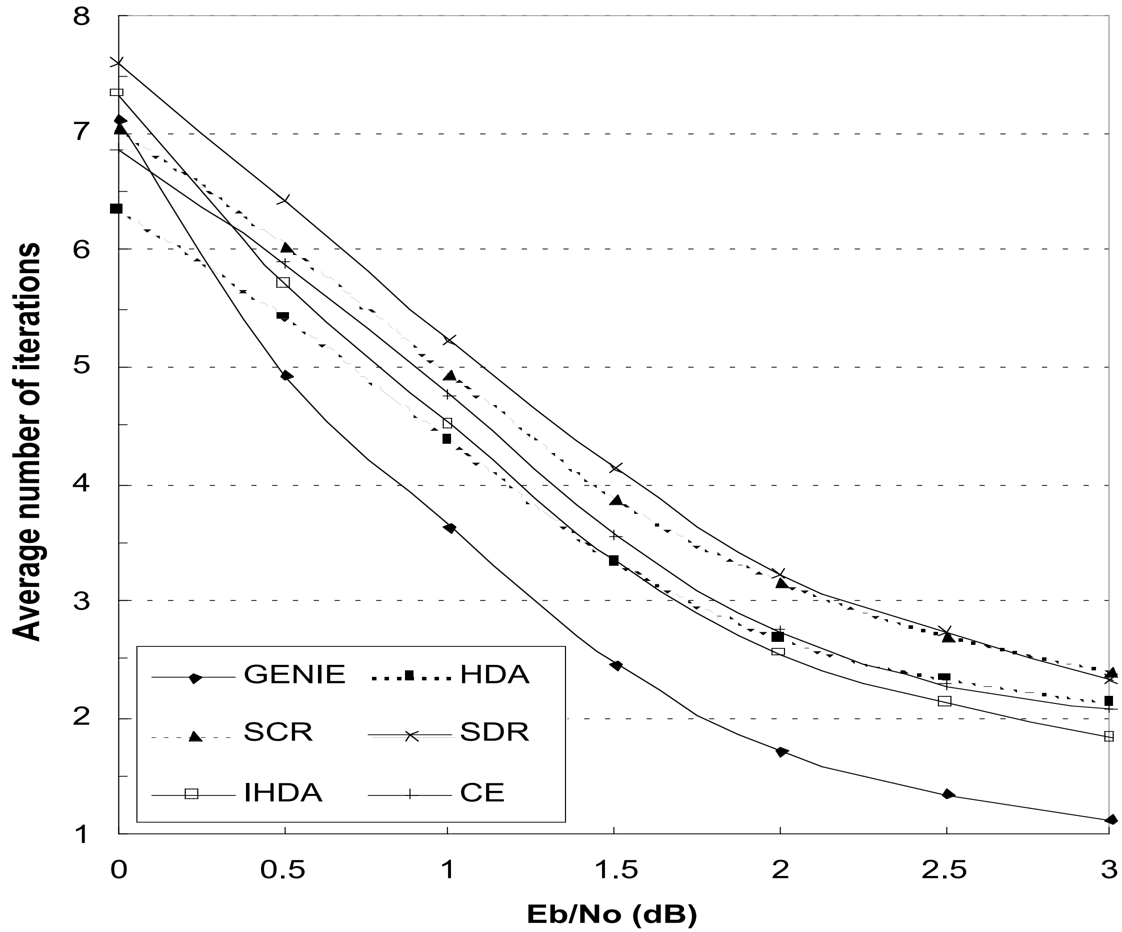
Results: BER Performance

Graph 1: Simulated BER performance for six stopping schemes.



Results: Average number of iterations

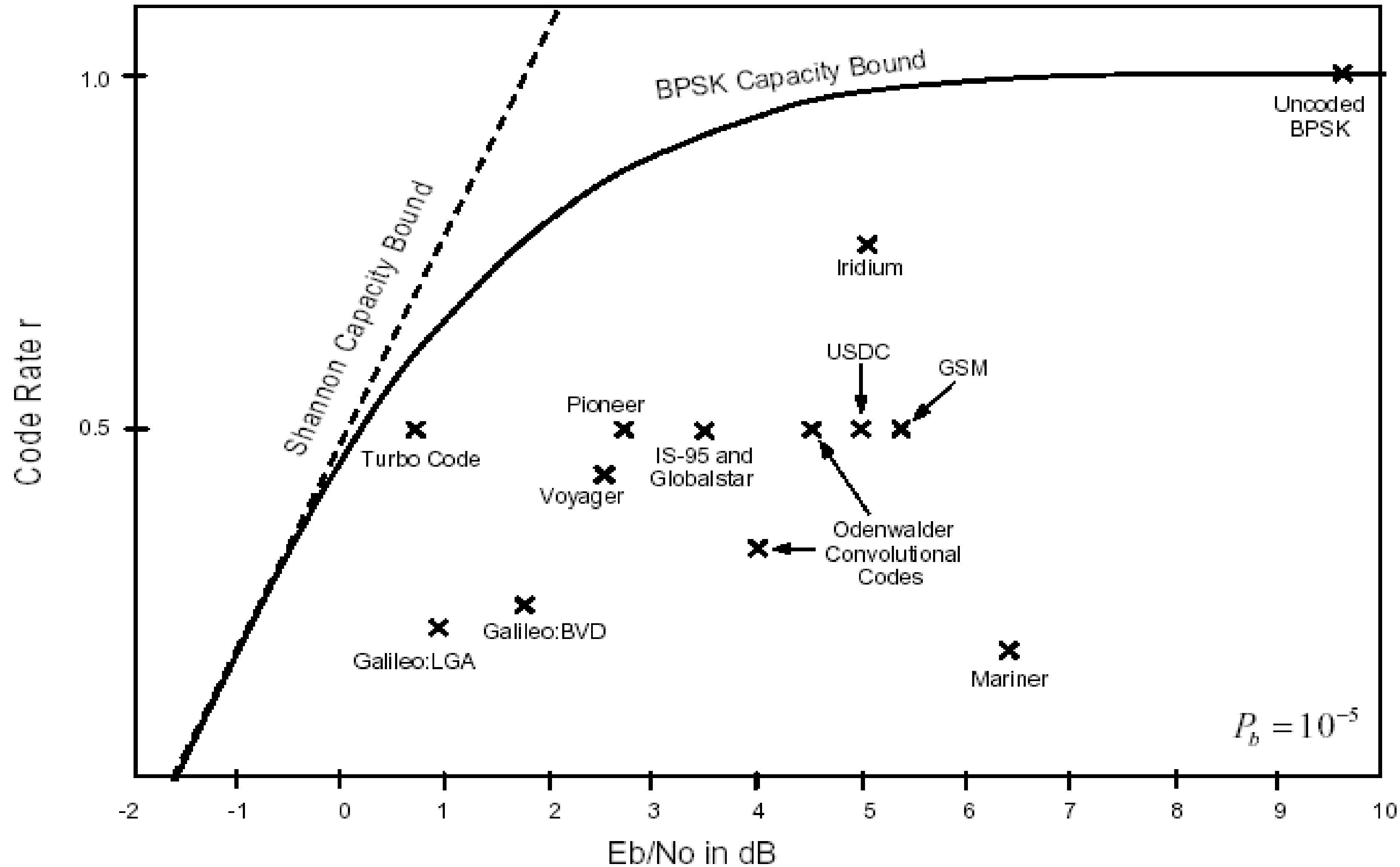
- Graph 2: Simulated Average number of iteration for the six schemes



Comparison of Stopping Criteria for Turbo Decoding

- All six schemes exhibit similar BER performance. The HDA, however, presents a slight degradation at high SNR.
- The IHDA saves more iterations for small interleaver sizes.
- The CE, SCR and HDA require extra data storage. The SCR and HDA however require less computation than the CE.
- Both the IHDA and the SDR have the advantage of reduce storage requirement.
- The IHDA has the additional advantage that its performance is independent of the choice of any parameter.

Comparison of Coding Systems



Turbo Codes for Burst Channels

- Studies of the performance of error correcting codes are most often concerned with situations where the channel is assumed to be memoryless, since this allows for a theoretical analysis.
- For a channel with memory, the Gilbert-Elliott (GE) channel is one of the simplest and practical models.
- We model a slowly varying Rayleigh fading channel with autocorrelation function $R(\tau) = J_0(2\pi f_m|\tau|)$ by the Gilbert-Elliott channel model.
- We then use this model to analytically evaluate the performance of Turbo-coded system.

The Gilbert-Elliott Channel Model

- The GE channel is a discrete-time stationary model with two states: one bad state which generally has high error probabilities and the other state is a good state which generally has low error probabilities.

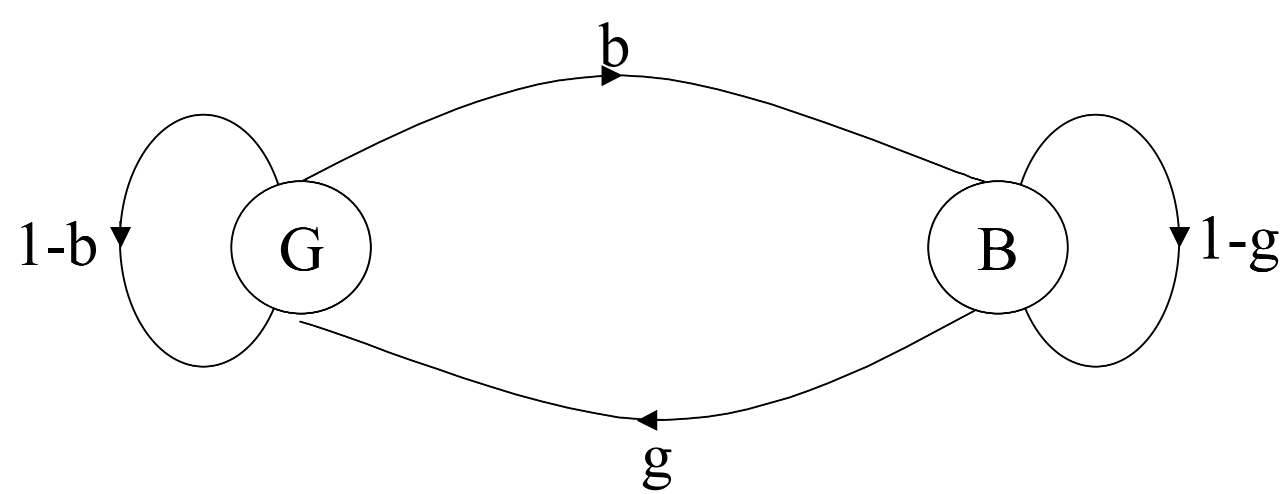


Fig. 6: The Gilbert-Elliott channel model.

The Gilbert-Elliott Channel Model

- The dynamics of the channel are modeled as a first-order Markov chain.
- In either state, the channel exhibits the properties of a binary symmetric channel.
- Important statistics
  - Steady state probabilities.

$$\pi_{\mathrm{G}} = \frac{g}{b+g} \qquad \pi_{\mathrm{B}} = \frac{b}{b+g}$$

  - Average time units in each state

$$\overline{\mathrm{T}}(\mathrm{G}) = \mathrm{E}[\mathrm{T}(\mathrm{G})] = \frac{1}{\mathrm{b}} \qquad \overline{\mathrm{T}}(\mathrm{B}) = \mathrm{E}[\mathrm{T}(\mathrm{B})] = \frac{1}{\mathrm{g}}$$

Matching the GE channel to the Rayleigh fading channel

- We let the average number of time unit the channel spends in the good (bad) state to be equal to the expected non-fade (fade) duration, normalized by the symbol time interval.
- In doing this, we obtain the following transition probabilities:

$$b = f_D T_S \sqrt{2\square\square}$$

$$g = \frac{f_D T_S \sqrt{2\square\square}}{e^{\square} \square 1}$$

Matching the GE channel to the Rayleigh fading channel

- The error probabilities in the respective states in the GE channel are taken to be the conditional error probabilities of the Rayleigh fading channel, conditioned on being in the respective state. For BPSK, the simplified expressions are:

$$P_e(G) = \frac{1}{2}\left\{ \mathrm{erfc}\left(\sqrt{\square_T}\right) \square \frac{\exp(\square)}{\sqrt{1+\frac{1}{\square}}}\,\mathrm{erfc}\left(\sqrt{\square_T + \square}\right) \right\}$$

$$P_e(B) = \frac{1 \square \mathrm{erfc}\left(\sqrt{\square_T}\right) \exp(-\square) \square \frac{1}{\sqrt{1+\frac{1}{\square}}}\left[1 \square \mathrm{erfc}\left(\sqrt{\square_T + \square}\right)\right]}{2\left[1 - \exp(-\square)\right]}$$

Performance Analysis of Turbo-Coded System on a Gilbert-Elliott Channel

- There are two main tools for the performance evaluation of turbo codes: Monte Carlo simulation and standard union bound.
- Monte Carlo simulation generates reliable probability of error estimates as low as 10^{-6} as is useful for rather low SNR.
- The union bound provides an upper bound on the performance of turbo codes with maximum likelihood decoding averaged over all possible interleavers.
- The expression for the average bit error probability is given as

$$P_{bit} = \sum_{d=d_{min}}^{3N} \sum_{i} \sum_{d_1} \sum_{d_2} \frac{i}{N} \binom{N}{i} P(d_1/i) P(d_2/i) P_2(d)$$

Performance Analysis of Turbo-Coded system on a Gilbert-Elliott Channel

- $P(d_1/i)$ and $P(d_2/i)$ are the distribution of the parity sequences and are given as

$$P(d_p/i) = \frac{t(N,i,d_p)}{\sum_{d_p} t(N,i,d_p)} = \frac{t(N,i,d_p)}{\binom{N}{i}}$$

- $P_2(d)$ is the pairwise-error probability and depends on the channel.
- We derive and expression for $P_2(d)$ for the Gilbert-Elliott channel.

Pairwise-error Probability for the Gilbert-Elliott Channel Model

- If the channel state is known exactly to the decoder we assume that amongst the d bits in which the incorrect path and the correct path differ, there are d_B in the bad state and $d_G = d\text{-}d_B$ in the good state.
- Amongst the d_B bits, there are e_B bits in error and amongst the d_G bits, e_G are in error.
- Let $CM^{(1)}$ and $CM^{(0)}$ be the metric of the incorrect path and the correct path respectively.

$$CM^{(1)} = e_B log\left(1 - P_e(B)\right) + \left(d_B - e_B\right) logP_e(B) + e_G log\left(1 - P_e(G)\right) + \left(d_G - e_G\right) logP_e(G)$$

$$CM^{(0)} = \left(d_B - e_B\right) log\left(1 - P_e(B)\right) + e_B logP_e(B) + \left(d_G - e_G\right) log\left(1 - P_e(G)\right) + e_G logP_e(G)$$

Pairwise-error probability for the Gilbert-Elliott Channel Model

- The probability of error in the pairwise comparison of $CM^{(1)}$ and $CM^{(0)}$ is:

$$P_2(d) = P_r\left(CM^{(1)} \mathrm{f}\ CM^{(0)}\right)$$
$$= P_r\left\{\left(d_G + Cd_B\right) \mathrm{p}\ 2\left(e_G + Ce_B\right)\right\}$$

where C is the metric ratio defined as

$$C = \frac{log\left[\left(1 - P_e(B)\right)/P_e(B)\right]}{log\left[\left(1 - P_e(G)\right)/P_e(G)\right]}$$

- To evaluate $P_2(d)$, we need the probability distribution of being in the bad state d_B times out of d and the distribution for being in the good state d_G times out of d.

Pairwise-error Probability for the Gilbert-Elliott Channel Model

- We show that:

$$P_d(d_B) = \begin{cases} (1-b)^{d-1}\pi_G, & d_B = 0 \\ \left[P_d(d_B/GG) + P_d(d_B/GB)\right]\pi_G + \left[P_d(d_B/BG) + P_d(d_B/BB)\right]\pi_B, & 1 \le d_B < d \\ (1-g)^{d-1}\pi_B, & d_B = d \end{cases}$$

and

$$P_d(d_G) = \begin{cases} (1-g)^{d-1}\pi_B, & d_G = 0 \\ \left[P_d(d_G/GG) + P_d(d_G/GB)\right]\pi_G + \left[P_d(d_G/BG) + P_d(d_G/BB)\right]\pi_B, & 1 \le d_G < d \\ (1-b)^{d-1}\pi_G, & d_G = d \end{cases}$$

Pairwise-error Probability for the Gilbert-Elliott Channel Model

where

$$P_d(d_B/GG) = \sum_{i=2}^{\min(d_B+1,\,d-d_B)} \binom{d-d_B-1}{i-1}\binom{d_B-1}{i-2}(1-b)^{d-d_B-i}\, b^{i-1}(1-g)^{d_B-i+1}\, g^{i-1}$$

$$P_d(d_B/GB) = \sum_{i=1}^{\min(d_B,\,d-d_B)} \binom{d-d_B-1}{i-1}\binom{d_B-1}{i-1}(1-b)^{d-d_B-i}\, b^{i}(1-g)^{d_B-i}\, g^{i-1}$$

$$P_d(d_B/BG) = \sum_{i=1}^{\min(d_B,\,d-d_B)} \binom{d-d_B-1}{i-1}\binom{d_B-1}{i-1}(1-b)^{d-d_B-i}\, b^{i-1}(1-g)^{d_B-i}\, g^{i}$$

$$P_d(d_B/BB) = \sum_{i=2}^{\min(d_B,\,d-d_B+1)} \binom{d-d_B-1}{i-2}\binom{d_B-1}{i-1}(1-b)^{d-d_B-i+1}\, b^{i-1}(1-g)^{d_B-i}\, g^{i-1}$$

Pairwise-error Probability for the Gilbert-Elliott Channel Model

and

$$P_d\left(d_G/GG\right)=\sum_{i=2}^{\min\left(d_G,d-d_G+1\right)}\binom{d-d_G-1}{i-2}\binom{d_G-1}{i-1}\left(1-b\right)^{d_G-i}b^{i-1}\left(1-g\right)^{d-d_G+1}g^{i-1}$$

$$P_d\left(d_G/GB\right)=\sum_{i=1}^{\min\left(d_G,d-d_G\right)}\binom{d-d_G-1}{i-1}\binom{d_G-1}{i-1}\left(1-b\right)^{d_G-i}b^{i}\left(1-g\right)^{d-d_G-i}g^{i-1}$$

$$P_d\left(d_G/BG\right)=\sum_{i=1}^{\min\left(d_G,d-d_G\right)}\binom{d-d_G-1}{i-1}\binom{d_G-1}{i-1}\left(1-b\right)^{d_G-i}b^{i-1}\left(1-g\right)^{d-d_G-i}g^{i}$$

$$P_d\left(d_G/BB\right)=\sum_{i=2}^{\min\left(d_G+1,d-d_G\right)}\binom{d-d_G-1}{i-1}\binom{d_G-1}{i-2}\left(1-b\right)^{d_G-i+1}b^{i-1}\left(1-g\right)^{d-d_G-i}g^{i-1}$$

Pairwise-error Probability for the Gilbert-Elliott Channel Model

- Thus

$$P_2(d) = \sum_{d_B + dG = d} \sum_{e_B} \sum_{e_G} \binom{d_B}{e_B} P_e(B)^{e_B} \left(1 - P_e(B)\right)^{d_B - e_B} P_d(d_B) \binom{d_G}{e_G} P_e(G)^{e_G} \left(1 - P_e(G)\right)^{d_G - e_G} P_d(d_G)$$

- We apply this union bound technique to obtain upper bounds on the bit-error rate of a turbo-coded DS-CDMA system.

Performance Analysis of Turbo-Coded DS-CDMA System on a Gilbert-Elliott Channel Model

- System model
  - We consider an asynchronous binary PSK direct-sequence CDMA system that allow K users to share a channel. The received signal at a given receiver is given by

$$r(t) = \sqrt{2P}\sum_{k=1}^{K} \alpha_k(t)\, a_k(t-\tau_k)\, b_k(t-\tau_k)\cos(\omega_c t + \theta_k) + n(t)$$

  - The output of the matched filter at each sampling instant is

$$Z_i = \sqrt{\frac{P}{2}}\, b_{i,0} T \alpha_i + \sqrt{\frac{P}{2}} \sum_{\substack{k=1 \\ k\neq i}}^{K} \alpha_k \left[b_{k,-1} R_{k,i}(\tau_k) + b_{k,0} \hat{R}_{k,i}(\tau_k) \right] \cos\phi_k + n'(t)$$

Performance Analysis of Turbo-coded DS-CDMA System on a Gilbert-Elliott Channel Model

- Using gausssian approximation, the SNR at the output of the receiver is

$$\square = \frac{\left(E\{Z_i/\square_i\}\right)^2}{\operatorname{var}\{Z_i/\square_i\}} = \frac{\square_i^2}{\frac{1}{3N}\sum_{\substack{k=1 \\ k \neq i}}^{K} E\{\square_k^2\} + \frac{N_0}{2P.T}}$$

and its expected value is

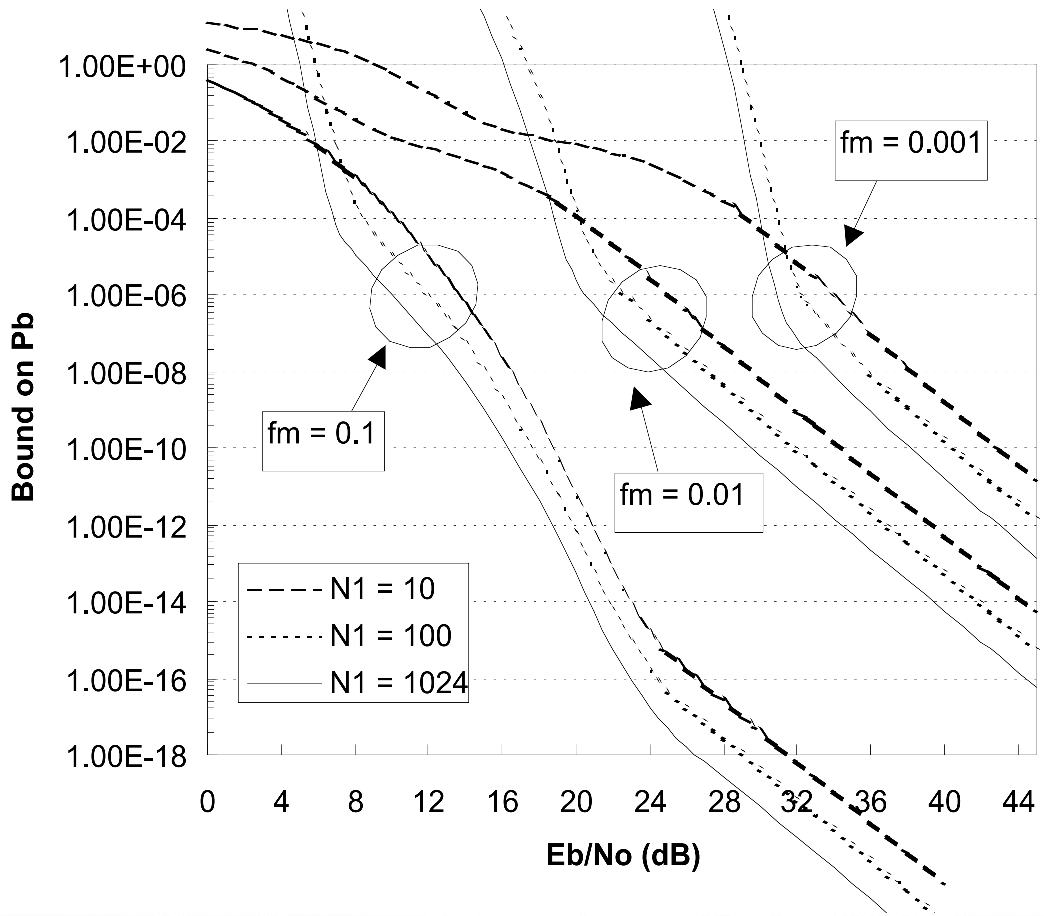
$$E\{\square\} = \square = \frac{1}{\frac{(K-1)}{3N} + \frac{N_0}{2P.T.E\{\square_i^2\}}}$$

Performance Analysis of Turbo-Coded DS-CDMA System on a Gilbert-Elliott Channel Model

- Simulation model
  - Code of rate 1/3, rate one-half RSC component encoders of memory 2 and octal generator (7, 5).
  - Gold spreading sequence of length $N = 63$.
  - The number of users is 10 and the frame size is 1024.
  - Perfect channel estimation and power control.
  - The product $f_D T_S$ is considered as an independent parameter, f_m, and simulations are performed for $f_m = 0.1$, 0.01and 0.001.
  - The threshold is 10 dB for $f_m = 0.1$, 0.01 and 14 dB for $f_m = 0.001$.

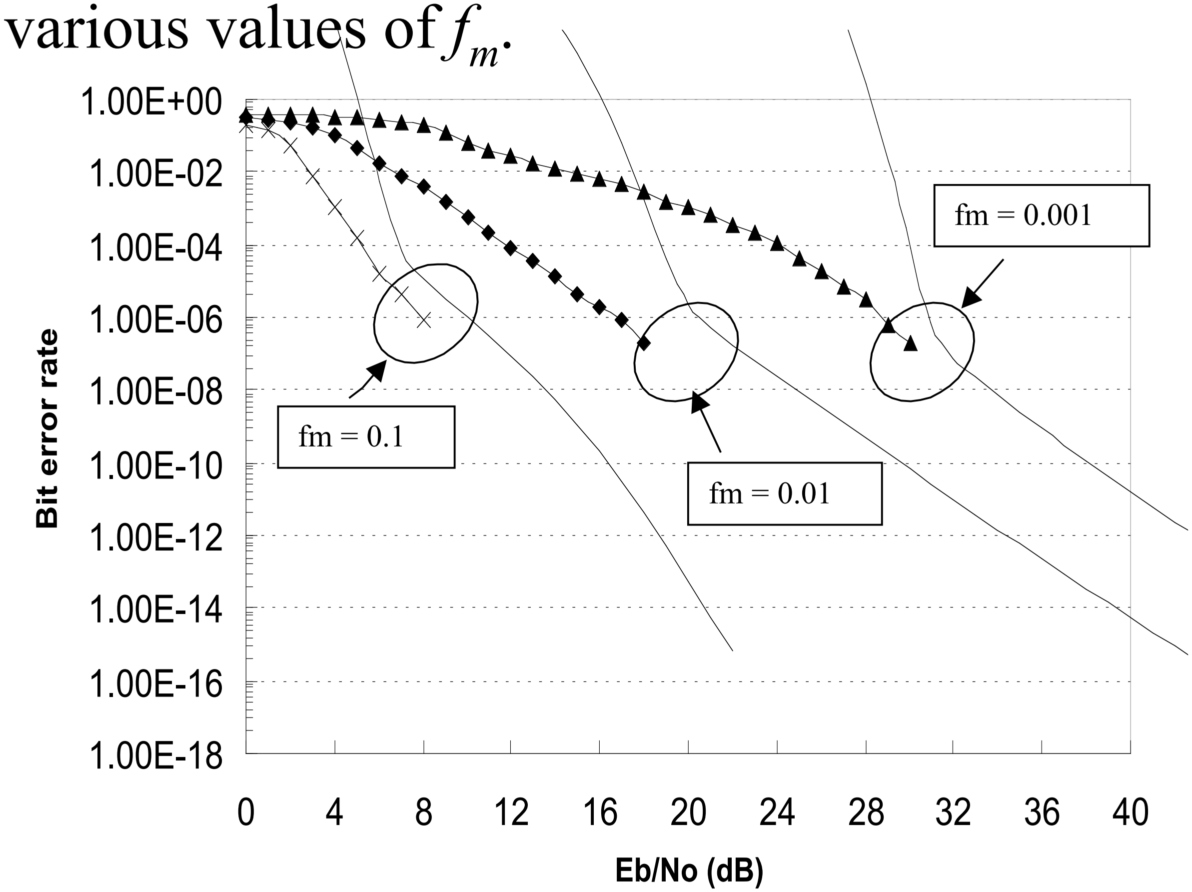
Results

Graph 3: Bounds on the BER for different values of $N1$ and f_m



Results

Graph 4: Transfer function bound (solid lines) versus simulation for various values of f_m.



The Effect of Imperfect Interleaving for the GE Channel

- An effective method to cope with burst errors is to insert an interleaver between the channel encoder and the channel.
- How effective an interleaver is depends on its depth, m.
- The size of the interleaver is typically determined by how much delay can be tolerated.
- We show that interleaving a code to degree m has exactly the same effect as transmitting at a lower rate or increased symbol duration of $T.m$.
- Thus, the GE channel with an interleaver will be equivalent to a GE channel where the corresponding transition probabilities are the m-step transition probabilities of the original model.

The Effect of Imperfect Interleaving for the GE Channel

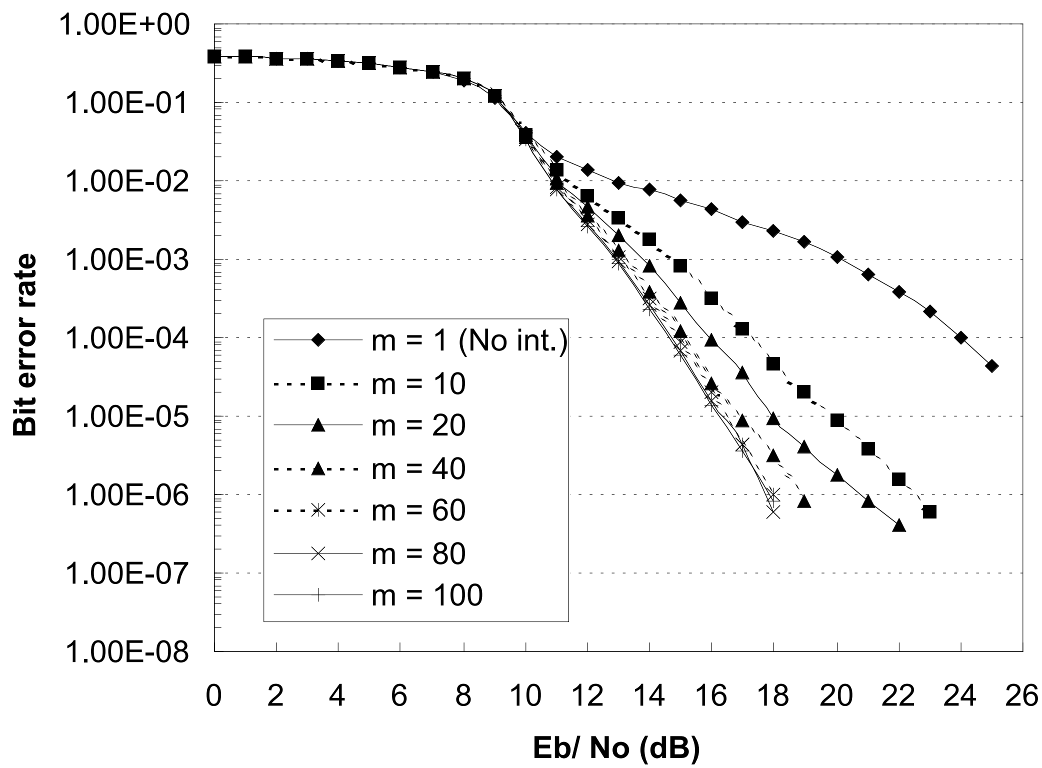
- The m-step transition probabilities are obtained by applying the Chapman-Kolmogrov equation to our two-state Markov chain.
- We obtain:

$$P^m(G/G) = \frac{g}{b+g} + \frac{b}{b+g}(1 \square b - g)^m \qquad P^m(G/B) = \frac{g}{b+g} \square \frac{g}{b+g}(1 \square b - g)^m$$

$$P^m(B/G) = \frac{b}{b+g} \square \frac{b}{b+g}(1 \square b - g)^m \qquad P^m(B/B) = \frac{b}{b+g} + \frac{g}{b+g}(1 \square b - g)^m$$

Results

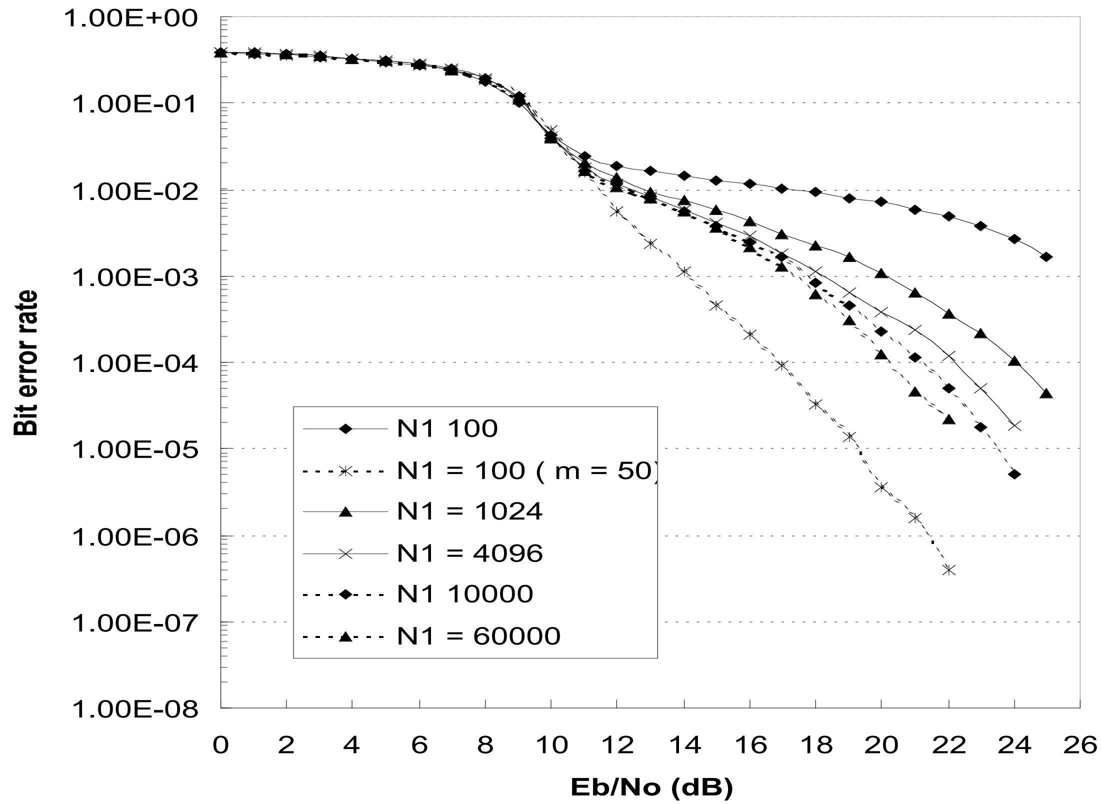
Graph 5: Simulated bit error rate on the interleaved Gilbert-Elliott channel model for different values of the interleaver depth. $fm = 0.001$.



Results

Graph 6: Comparison of the performance of a combined small code interleaver with

channel interleaver and larger code interleavers of various sizes.



Conclusions

- In this presentation, we present stopping criteria for Turbo decoding.
- We model a slowly varying Rayleigh fading channel by the Gilbert-Elliott channel model.
- We then use this model to analytically evaluate the performance of a Turbo-coded DS-CDMA system.
- We analyze the effect of imperfect interleaving for the Gilbert-Elliott channel model.
- We show that a combination of a small code interleaver with a channel interleaver could outperform codes with very large interleavers, making Turbo codes suitable for even delay-sensitive services.