New criteria for iterative decoding

Florence Alberge, Ziad Naja, Pierre Duhamel

New criteria for iterative decoding

Florence Alberge, Ziad Naja, Pierre Duhamel
International Conference on Acoustic Speech and Signal Processing 09

Laboratoire des Signaux et Systèmes
Supelec, Gif-sur-Yvette, France

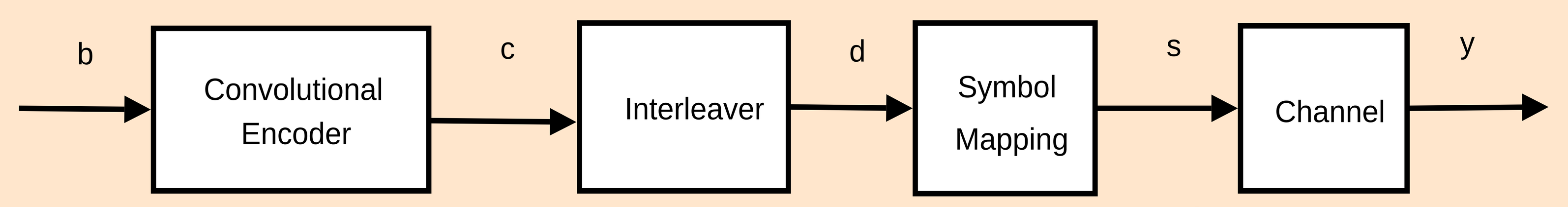
Univ. Paris-Sud, UMR 8506, Orsay
CNRS, Orsay, France

Context: Turbo-like algorithms with iterative decoding

Goal: Make the link between iterative decoding and classical optimization techniques. Improve the performance of iterative decoding.

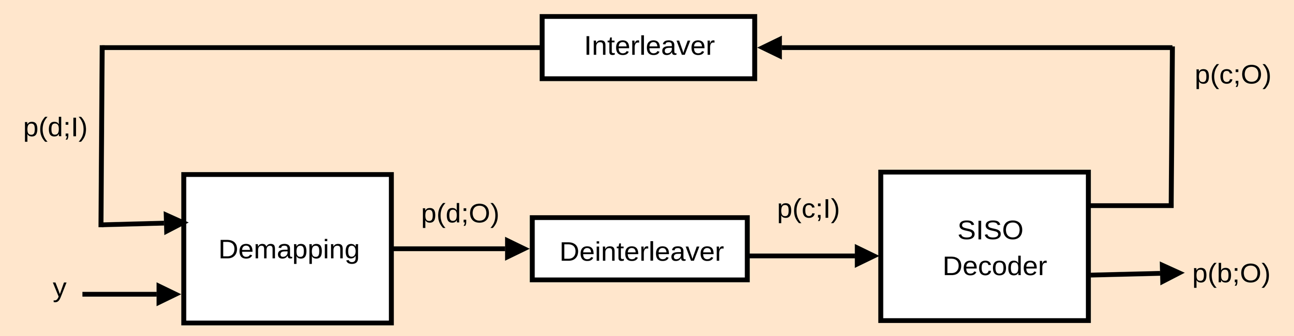
System model (Bit Interleaved Coded Modulation)(Zehavi, 1992; Li, 2002)

Transmission model



AWGN Channel $y_k = s_k + n_k$ $\mathbf{n}_k \rightsquigarrow \mathcal{N}(\mathbf{0}, \sigma_n^2)$
$s_k = Mapping(d_{km+1}, d_{km+2}, ..., d_{(k+1)m})$

BICM-ID receiver with soft-decision feedback



$$\textbf{APP (Demapping)} \quad p_{APP}(d_{km+i} = b) = K'_m \sum_{\mathbf{s}_k \in \Psi_b^i} p(\mathbf{y}_k \mid \mathbf{s}_k) \prod_j p(d_{km+j}; I)$$

Prior $\times$ Channel probability $\rightarrow$ Marginalization

$$\textbf{Extrinsic (Demapping)} \quad p(d_{km+i} = b; O) = K_m \sum_{\mathbf{s}_k \in \Psi_b^i} p(\mathbf{y}_k \mid \mathbf{s}_k) \prod_{j \neq i} p(d_{km+j}; I)$$

$$\textbf{APP (Decoder)} \quad p_{APP}(c_l = b) = K'_c \sum_{\mathbf{c} \in \mathcal{R}_b^l} \mathbf{I}_{\mathcal{C}}(\mathbf{c}) \prod_j p(c_j; I) \quad (\mathbf{I}_{\mathcal{C}} \textit{ indicator function of the code})$$

Prior $\times$ Indicator function $\rightarrow$ Marginalization

$$\textbf{Extrinsic (Decoder)} \quad p(c_l = b; O) = K_c \sum_{\mathbf{c} \in \mathcal{R}_b^l} \mathbf{I}_{\mathcal{C}}(\mathbf{c}) \prod_{j \neq l} p(c_j; I)$$

Information geometry and iterative decoding

Basic tools and Notations (Walsh, 2006)

$(\mathbf{B_i}) \in \{0,1\}^N$ **binary representation** of integer i gathered into matrix $\mathbf{B} = (\mathbf{B_0}, \mathbf{B_1}, ..., \mathbf{B_{2^N-1}})^T$.

PMF $\eta = (\mathbf{Pr}[\chi = \mathbf{B_0}], \mathbf{Pr}[\chi = \mathbf{B_1}], ..., \mathbf{Pr}[\chi = \mathbf{B_{2^N-1}}])^{\mathbf{T}}$

Log-coordinates of PMF η $(\theta_i)_{0 \leq i \leq 2^N-1} = \ln(Pr[\chi = \mathbf{B_i}]) - \ln(Pr[\chi = \mathbf{B_0}])$

Bitwise log-probability ratio $(\lambda_j)_{0 \leq j \leq N-1} = log(\frac{Pr[\chi_j=1]}{Pr[\chi_j=0]})$

For **factorisable probability measures**, the log-coordinates take the form $\theta = \mathbf{B}\lambda$.

Link with iterative decoding

- **Demapping subblock**

$$p_{APP}(d_{km+i} = b) = K'_m \textstyle\sum_{\mathbf{s}_k \in \Psi_b^i} p(\mathbf{y}_k \mid \mathbf{s}_k) \; \prod_j p(d_{km+j}; I)$$
$$\downarrow \theta_{\mathbf{m}} \qquad \downarrow_{Log-coordinates} \mathbf{B}\lambda_{\mathbf{1}}$$

$$p_{APP}(d_{km+i} = b) = K''_m \; p(d_{km+i}; I) \; p(d_{km+i}; O)$$
$$\downarrow \mathbf{B}\lambda_{\mathbf{1}} \qquad \downarrow_{Log-coordinates} \mathbf{B}\lambda_{\mathbf{2}}$$

The **demapping sub-block** solves, with respect to λ_2, $p_{\mathbf{B}(\lambda_1+\lambda_2)} = p_{\mathbf{B}\lambda_1+\theta_{\mathbf{m}}}$

- **Decoding sub-block**

$$p_{APP}(c_l = b) = K'_c \textstyle\sum_{\mathbf{c} \in \mathcal{R}_b^l} \mathbf{I}_{\mathcal{C}}(\mathbf{c}) \; \prod_j p(c_j; I)$$
$$\downarrow \theta_{\mathbf{c}} \qquad \downarrow_{Log-coordinates} \mathbf{B}\lambda_{\mathbf{2}}$$

$$p_{APP}(c_l) = K''_c \; p(c_l; I) \; p(c_l; O)$$
$$\downarrow \mathbf{B}\lambda_{\mathbf{2}} \qquad \downarrow_{Log-coordinates} \mathbf{B}\lambda_{\mathbf{1}}$$

The **decoding sub-block** solves, with respect to λ_1, $p_{\mathbf{B}(\lambda_1+\lambda_2)} = p_{\mathbf{B}\lambda_2+\theta_{\mathbf{c}}}$

- **Global criterion**
Let $D_{FD}(\mathbf{p}, \mathbf{q}) = \sum_j p_j \ln\left(\frac{p_j}{q_j}\right) + \sum_j (1-p_j) \ln\left(\frac{1-p_j}{1-q_j}\right)$ denote the **Fermi-Dirac entropy** (Kullback-Leibler distance for bit probabilities).

The demapping sub-block solves the minimization problem $\min_{\lambda_2} D_{FD}(p_{\mathbf{B}\lambda_1+\theta_{\mathbf{m}}}, p_{\mathbf{B}(\lambda_1+\lambda_2)})$.

The decoding sub-block solves the minimization problem $\min_{\lambda_1} D_{FD}(p_{\mathbf{B}\lambda_2+\theta_{\mathbf{c}}}, p_{\mathbf{B}(\lambda_1+\lambda_2)})$.

New criteria

An hybrid proximal point algorithm

Goal: Link the two (independent) criteria using **proximal point technique** (Luque, 1984).

$$\textbf{Demapping} \quad \min_{\lambda_2} J_{\theta_m}(\lambda_1, \lambda_2) = \min_{\lambda_2} D_{FD}(\mathbf{p}_{\mathbf{B}\lambda_1+\theta_{\mathbf{m}}}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)}) + \mu_m D_{FD}(\mathbf{p}_{\mathbf{B}(\lambda_1^{(\mathbf{k})}+\lambda_2^{(\mathbf{k})})}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)})$$
$$\textbf{Decoding} \quad \min_{\lambda_1} J_{\theta_c}(\lambda_1, \lambda_2) = \min_{\lambda_1} D_{FD}(\mathbf{p}_{\mathbf{B}\lambda_2+\theta_{\mathbf{c}}}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)}) + \mu_c D_{FD}(\mathbf{p}_{\mathbf{B}(\lambda_1^{(\mathbf{k})}+\lambda_2^{(\mathbf{k+1})})}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)})$$

Optimal choice of the step-size :

$$\textbf{Choose } \mu_{\mathbf{m}} \textbf{ such that } J_{\theta_m}(\lambda_1^{(\mathbf{k})}, \lambda_2^{(\mathbf{k+1})}) \leq J_{\theta_c}(\lambda_1^{(\mathbf{k})}, \lambda_2^{(\mathbf{k})})$$
$$\textbf{Choose } \mu_{\mathbf{c}} \textbf{ such that } J_{\theta_c}(\lambda_1^{(\mathbf{k+1})}, \lambda_2^{(\mathbf{k+1})}) \leq J_{\theta_m}(\lambda_1^{(\mathbf{k})}, \lambda_2^{(\mathbf{k+1})})$$

Both criteria **decrease with the iterations**. Convergence towards the same stationary point than the classical iterative decoding.

An hybrid minimum entropy algorithm

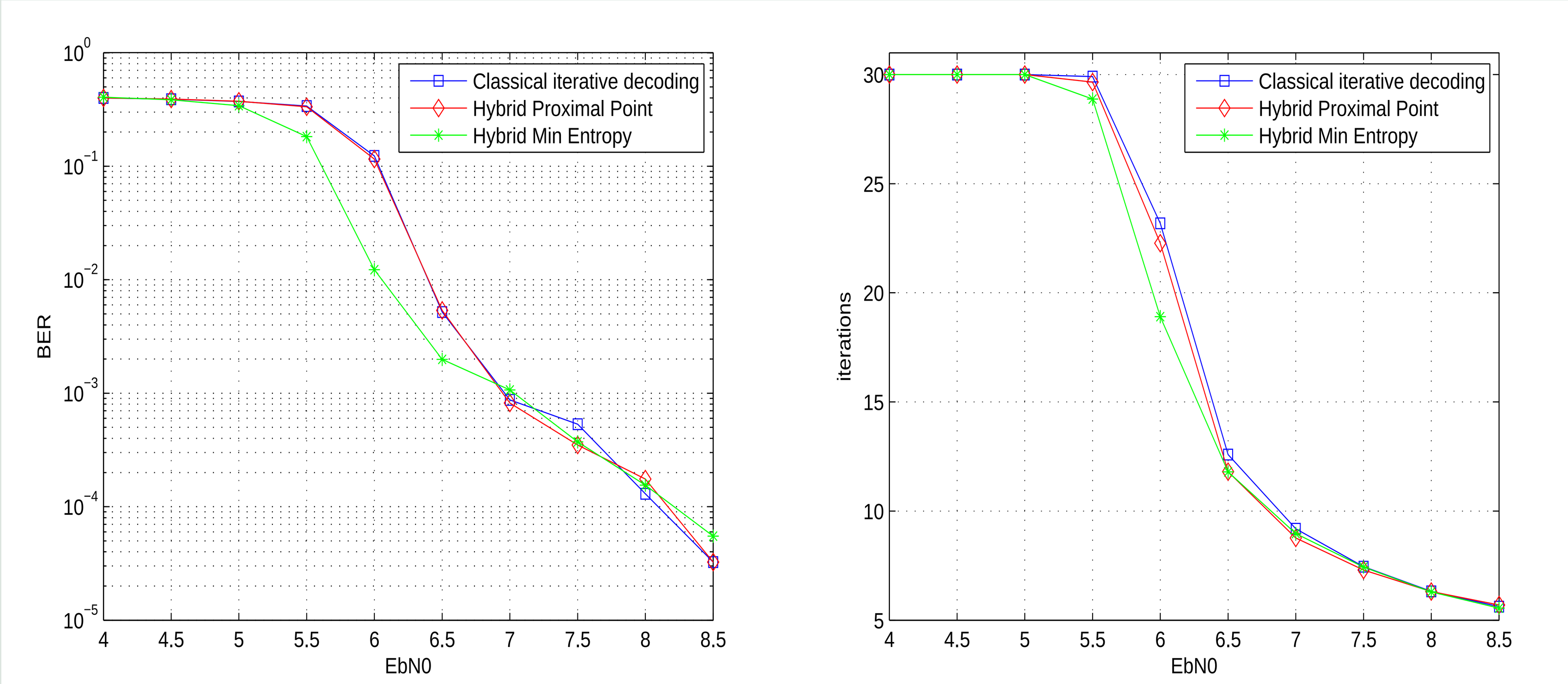
Goal: Improve the performance of the iterative decoding.

Rationale: The entropy of the APP gives a measure of the **reliability** of the decision : $E_{B(\lambda_1+\lambda_2)} \rightarrow 0$ means that the iterative decoding is **confident about its decisions** (Kocarev, 2006).

$$\textbf{Demapping} \quad \min_{\lambda_2} D_{FD}(\mathbf{p}_{\mathbf{B}\lambda_1+\theta_{\mathbf{m}}}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)}) + \eta_m E_{B(\lambda_1+\lambda_2)}$$
$$\textbf{Decoding} \quad \min_{\lambda_1} D_{FD}(\mathbf{p}_{\mathbf{B}\lambda_2+\theta_{\mathbf{c}}}, \mathbf{p}_{\mathbf{B}(\lambda_1+\lambda_2)}) + \eta_c E_{B(\lambda_1+\lambda_2)}$$

Simulation

The generator polynomial of the encoder is $g = [111; 001; 100]$. The bits are mapped using subset partitioning to a 8-PSK modulation. The length of the coded bit sequence is $L_c = 6000$. The step-sizes η_m and η_c in the HMEA are both chosen equal to 0.05.



Left: BER versus EbN0 – Right: Iteration number versus EbN0

References

E. Zehavi. 8-PSK trellis codes for a Rayleigh fading channel. *IEEE. Trans on Commun.*, 40:873–883, May 1992.

X. Li, A. Chindapol and J.A. Ritcey. Bit interleaved coded modulation with iterative decoding and 8-PSK signaling. *IEEE. Trans on Commun.*, 50:1250–1257, Aug. 2002.

J.M. Walsh, P.A. Regalia and C.R. Johnson. Turbo decoding as Iterative Constrained Maximum Likelihood sequence Decoding. *IEEE Trans on Inform Theory*, 52:5426–5437, Dec. 2006.

F.J. Luque. Asymptotic convergence analysis of the proximal point algorithm. *SIAM Journal on Control and Optimization*, 22(2), pp. 277–293, 1984.

L. Kocarev, F. Lehmann, G.M. Maggio, B. Scanavino, Z. Tatsev and A. Vardy. Nonlinear dynamics of iterative decoding systems: Analysis and Applications. *IEEE Trans. on Inform. Theory*, 52(4): 1366–1384, 2006.