A Game-Theoretic Interpretation of Iterative Decoding System model (Bit Interleaved Coded Modulation)

Florence Alberge

A Game-Theoretic Interpretation of Iterative Decoding

Florence Alberge

19th European Signal Processing Conference (EUSIPCO 2011)

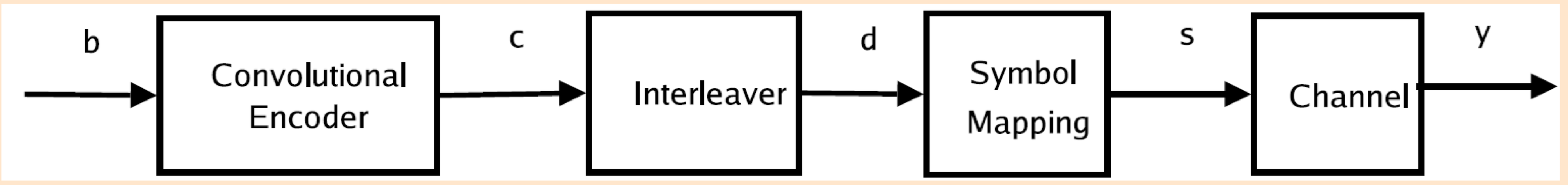
Univ. Paris-Sud, UMR 8506, Orsay
CNRS, Orsay, France

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

Goal: Iterative Decoding as a distributed optimization strategy - Interpretation of turbo-like Iterative Decoding (Game) - Results on Convergence.

System model (Bit Interleaved Coded Modulation)

Transmission model



Noisy Memoryless Channel $y_k = h_k s_k + n_k$

$s_k = Mapping(d_{kM+1}, d_{kM+2}, ..., d_{(k+1)M})$

n_b information bits, n encoded bits

Maximum Likelihood Decoding (MLD)

- Searching for optimal sequence of information bits $\widehat{\mathbf{b}}_{MLD}$

$$\widehat{\mathbf{b}}_{MLD} = \arg\max_{\mathbf{b}\in\{0,1\}^{n_b}} p(\mathbf{y} \mid \mathbf{b})$$

- Searching for optimal sequence of encoded/interleaved bits $\widehat{\mathbf{d}}_{MLD}$

$$\widehat{\mathbf{d}}_{MLD} = \arg\max_{\mathbf{d}\in\{0,1\}^{n}} \underbrace{p_{ch}(\mathbf{y} \mid \mathbf{d})}_{Channel\ probability} \underbrace{I_{co}(\mathbf{d})}_{Indicator\ function}$$

- Searching for optimal weighting $\widehat{\mathbf{p}}_{MLD}$

$$\widehat{\mathbf{p}}_{MLD}(\mathbf{d}) = \arg\max_{\mathbf{p}\in\mathcal{E}_s} \sum_{\mathbf{d}} \mathbf{I}_{co}(\mathbf{d})\mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\mathbf{p}(\mathbf{d}) \quad \mathcal{E}_s \text{ : fully-factorized PMFs}$$

Solution is $\widehat{\mathbf{p}}_{MLD}(\mathbf{d}) = \delta_{\widehat{\mathbf{d}}_{MLD}}(\mathbf{d})$

Implementation untractable $\Rightarrow$ Suboptimal optimization

Suboptimal Decoding

Numerical value of $n \Rightarrow$ Solution 1: Compute bit-marginals
Interleaver $\Rightarrow$ Solution 2: Consider separately decoding and demapping
Leads to another expression for the MLD

$$\left(\widehat{\mathbf{l}}_{MLD}(\mathbf{d}), \widehat{\mathbf{q}}_{MLD}(\mathbf{d})\right) = \arg\max_{\mathbf{l},\mathbf{q}\in\mathcal{E}_s} \sum_{d_k}\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\mathbf{q}(\mathbf{d})\mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\mathbf{l}(\mathbf{d})$$

with optimal solution

$$\widehat{\mathbf{l}}_{MLD}(\mathbf{d}) = \widehat{\mathbf{q}}_{MLD}(\mathbf{d}) = \delta_{\widehat{\mathbf{d}}_{MLD}}(\mathbf{d})$$

but still untractable...

n convenient APPROXIMATIONS: $\mathcal{C}_{MLD} \rightarrow \tilde{\mathcal{C}}_1 \ \dots \ \tilde{\mathcal{C}}_k \ \dots \ \tilde{\mathcal{C}}_n$ specific for each bit-marginal

$$\underbrace{\sum_{d_k}\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\mathbf{q}(\mathbf{d})\mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\mathbf{l}(\mathbf{d})}_{\textit{bit-marginals of a product } (C_{MLD},\ optimal)} \approx \underbrace{\sum_{d_k}\left(\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\mathbf{q}(\mathbf{d})\right)\left(\sum_{\mathbf{d}:d_k} \mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\mathbf{l}(\mathbf{d})\right)}_{\textit{product of the bit-marginals } (\tilde{C}_k,\ suboptimal)}$$

- Suggests a distributed optimization strategy
- Average cost function: $\tilde{\mathcal{C}} = \sum_{i=1}^{n} \tilde{\mathcal{C}}_k \Rightarrow$ Score function for the distributed optimization strategy
- $\tilde{\mathcal{C}}$ is a weighted sum:

$$\tilde{\mathcal{C}}(\mathbf{l},\mathbf{q}) = n\mathcal{C}_{MLD} + \sum_{r=1}^{n-1}(n-r)N_r(\mathbf{l},\mathbf{q})$$

where $N_r(\mathbf{l},\mathbf{q}) = \sum_{(\mathbf{d},\mathbf{d}')\in\mathbb{S}_r} \mathbf{I}_{co}(\mathbf{d})\mathbf{q}(\mathbf{d})\mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d}')\mathbf{l}(\mathbf{d}')$ and $\mathbb{S}_r$ denote the set of pairs $(\mathbf{d},\mathbf{d}')$ such that $d_H(\mathbf{d},\mathbf{d}') = r$ (Hamming distance)

Distributed Optimization

Maximize separately the n cost functions $\tilde{\mathcal{C}}_k$

$$\left(\widehat{l}_k, \widehat{q}_k\right) = \arg\max_{l_k,q_k\in\mathcal{F}} \sum_{d_k} q_k(d_k)l_k(d_k)\left(\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\Pi_{i\neq k}q_i(d_i)\right)\left(\sum_{\mathbf{d}:d_k} \mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\Pi_{i\neq k}l_i(d_i)\right)$$

Linear in $q_k(d_k)l_k(d_k) \Rightarrow \widehat{l}_k\widehat{q}_k \in \{0,1\}$: HARD decision

SOFT decision preferred: $\Rightarrow \widehat{q}_k\widehat{l}_k \propto \left(\underbrace{\underbrace{\left(\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\Pi_{i\neq k}q_i(d_i)\right)}_{\textit{extrinsic (decoder)}}\underbrace{\left(\sum_{\mathbf{d}:d_k} \mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\Pi_{i\neq k}l_i(d_i)\right)}_{\textit{extrinsic (demapper)}}}_{\textit{A Posteriori Probability}}\right)^{\beta}$

Equivalently with LLR :

$$\lambda_{\mathbf{l},k} + \lambda_{\mathbf{q},k} = \beta\log\left(\frac{\sum_{\mathbf{d}:d_k} \mathbf{I}_{co}(\mathbf{d})\Pi_{i\neq k}q_j(d_j)}{\sum_{\mathbf{d}:\overline{d}_k} \mathbf{I}_{co}(\mathbf{d})\Pi_{i\neq k}q_i(d_i)}\right) + \beta\log\left(\frac{\sum_{\mathbf{d}:d_k} \mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\Pi_{i\neq k}l_i(d_i)}{\sum_{\mathbf{d}:\overline{d}_k} \mathbf{p}_{ch}(\mathbf{y} \mid \mathbf{d})\Pi_{i\neq k}l_i(d_i)}\right) \quad (1)$$

Can be recast as a distributed optimization process with strictly concave utility functions:

$$U_k(\lambda_{\mathbf{l},k}, \lambda_{\mathbf{q},k}, \lambda_{\mathbf{l},-k}, \lambda_{\mathbf{q},-k}) = -\parallel \lambda_{\mathbf{l},k} + \lambda_{\mathbf{q},k} - \beta\log\left(\frac{f_{d_k}(\mathbf{q}, I_{co})f_{d_k}(\mathbf{l}, p_{ch}(\mathbf{y} \mid \mathbf{d}))}{f_{\overline{d}_k}(\mathbf{q}, I_{co})f_{\overline{d}_k}(\mathbf{l}, p_{ch}(\mathbf{y} \mid \mathbf{d})}\right) \parallel^2 \quad (2)$$

Iterative Decoding as a Game

Definition 1 (Game)**.** A game $\mathcal{G}$ is defined as a triplet $\mathcal{G} = (K, \{S_i\}_{i\in K}, \{u_i\}_{i\in K})$ where $K = \{1, 2, ..., n\}$ is a finite set of players, $\forall i \in K$, $\{S_i\}$ is the set of strategy of player i and u_i its utility function.

Proposition 1. Iterative decoding is a game $\mathcal{G}_{soft}$ with n players (**encoded bits**), set of strategy $\{S_i\}_{i\in K} = \mathbb{R}^2$, and utility function $\{U_i\}_{i\in K}$ defined in (**??**).

Definition 2 (Nash Equibrium)**.** A profile s^* is a (pure) NE for $\mathcal{G}$ if $\forall i \in K$, $\forall s_i' \in S_i$, $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i', s_{-i}^*)$.

Proposition 2. Let $(\lambda_{\mathbf{l}^*}, \lambda_{\mathbf{q}^*})$ be a solution of (**??**). Then, $(\lambda_{\mathbf{l}^*}, \lambda_{\mathbf{q}^*})$ is a NE of the game $\mathcal{G}_{soft}$.

Proposition 3. Let $(\lambda_{\mathbf{l}^*}, \lambda_{\mathbf{q}^*})$ be a solution of (**??**). Then, $(\lambda_{\mathbf{l}^*}, \lambda_{\mathbf{q}^*})$ is an induced equilibrium of the game $\mathcal{G}_{hard}$ with $K = \{0, 1, ..., n\}$, $\{S_i\}_{i\in K} = [0,1]\times[0,1]$ and with utility function $\tilde{\mathcal{C}}_i$ meaning that $\forall i \in K$, $\forall s_i' = (l_i(d_i), q_i(d_i)) \in S_i$,

$$\tilde{\mathcal{C}}_i(H(s_i^*), s_{-i}^*) \geq \tilde{\mathcal{C}}_i(s_i', s_{-i}^*)$$

where $H(\mathbf{s}_i^*)$ is a hard decision operator.

Definition 3 (Social Welfare)**.** The social welfare W of a game is defined as the sum of the utilities of all players, $W = \sum_{i=1}^{n} u_i$.

- Game $\mathcal{G}_{soft} \Rightarrow W_{soft} = 0$ for any NE (USELESS).
- Game $\mathcal{G}_{hard} \Rightarrow W_{hard} = \tilde{\mathcal{C}} \Rightarrow$ a posteriori evaluation of the optimization process and of the vicinity to the MLD solution.
- There always exists at least one NE (proved in (**?**) when $\beta = 1$).

Convergence

Iterative decoding = (**??**) + Jacobi/Gauss-Seidel scheduling (**?**).
Convergence proofs based on the Jacobian ∇F (**?**) $\rightarrow$ Key property for ∇F: STRICTLY DIAGONAL DOMINANCE.

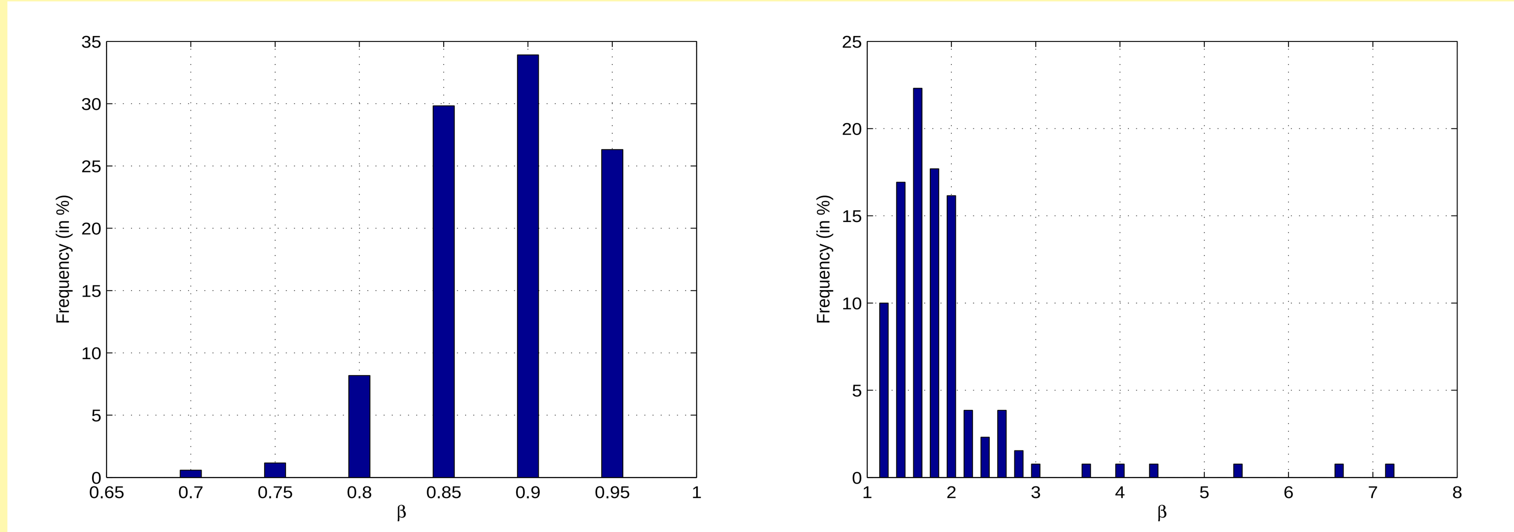
Definition 4. Let **A** denote a $n \times n$ matrix with elements a_{ij} in $\mathbb{R}$. **A** is a strictly diagonal dominant matrix if

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \quad \forall i \in \{1, ..., n\}$$

Proposition 4. It always exist $\beta_0 > 0$ such that $\forall\beta \leq \beta_0$, $\nabla F(\lambda_{\mathbf{l}}, \lambda_{\mathbf{q}})$ is a strictly diagonally dominant matrix for all $(\lambda_{\mathbf{l}}, \lambda_{\mathbf{q}}) \in \mathbb{R}^n \times \mathbb{R}^n$.

Illustration

BICM - $(5, 7)$ convolutional code of rate $1/2$ - $n_b = 400$ (a frame). Set partitioning - 16-QAM.



(left) EbN0=5dB - β_0 distribution among non-converging sequences in classical BICM , (right) EbN0=7dB among converging sequences in classical BICM

Conclusion

- $\tilde{\mathcal{C}}$: a posteriori score function (optimization process / MLD solution)
- β useful to improve convergence rate
- Convergence is not an issue

References

F. Alberge, Z. Naja and P. Duhamel. From Maximum Likelihood to Iterative Decoding. *In ICASSP Proc*. Prague, Czech Republic, May 22-27, 2011.

J.Osborne and A. Rubinstein. *A course in game theory*. MIT Press, 1994.

S. Lasaulce, M. Debbah and E. Altman. Methodologies for Analysing Equilibria in Wireless Games. *IEEE Sig. Proc. Magazine*, vol. 26, no. 5, pp. 357-378, 2009.

T. Richardson. The Geometry of Turbo-decoding Dynamics. *IEEE Trans on inform. Theory*, vol. 46, no. 1, pp. 9-23, 2000.

J.J. Moré. Nonlinear Generalizations of Matrix Diagonal Dominance with Application to Gauss-Seidel Iterations. *SIAM J. Numer Anal.*, vol. 9, no. 2, pp. 357-378, 1972.