

DISTRIBUTED CODING OF SOURCES WITH BURSTY CORRELATION

Elsa Dupraz[†], Francesca Bassi[†], Thomas Rodet[†], Michel Kieffer^{†,*, \diamond}

[†]LSS – CNRS – SUPELEC – Univ Paris-Sud, 91192 Gif-sur-Yvette, France

* On leave at LTCI – CNRS Télécom ParisTech, 75013 Paris, France, \diamond Institut Universitaire de France

ABSTRACT

This paper focuses on the performance of a Wyner-Ziv coding scheme for which the correlation between the source and the side information is modeled by a hidden Markov model with Gaussian emission. Such a signal model takes the memory of the correlation into account and is hence able to describe the bursty nature of the correlation between sources in applications such as sensor networks, video coding etc.

This paper provides bounds on the rate-distortion performance of a Wyner-Ziv coding scheme for such model. It proposes a practical coding scheme able to exploit the memory in the correlation. Finally, the contribution to each part of the coding and decoding scheme is analysed.

Index Terms— Wyner-Ziv coding, Hidden Markov Model, Rate-distortion analysis, MCMC

1. INTRODUCTION

Although results on the theoretical performance of lossy distributed source coding schemes are since long established [1, 2], many efforts remain to be made towards the characterization of setups involving real sources. This paper aims to account for the potentially bursty nature of the correlation between sources in Wyner-Ziv setups.

Bursty correlation may appear, for example, in the context of sensor networks [3]. Two closely located sensors measure some physical quantity (temperature, pressure). The respective readings X_k and Y_k are, in general, strongly correlated, but perturbations in the local sensing environment may occasionally cause bursts of weakly correlated pairs. Simple correlation models for X_k and Y_k , such as Gaussian [4] or Gauss-Markov [5], may not be sufficient to capture such a behavior.

The instantaneous correlation level (good, bad) between X_k and Y_k is represented in this paper by the realization of some hidden state variable S_k . The state sequence $\{S_k\}_{k=1}^{+\infty}$, as detailed in Section 2, is modeled as a Markov process, which allows to account for the temporal dependence of consecutive correlation levels. Similar signal representations have been already considered in the context of lossless source coding [6]. For the binary case, in particular, the dependence between the sources has been modeled using a Gilbert-Elliot

channel [7] and practical coding solutions have been proposed [8, 9]. To the best of our knowledge no performance analysis or practical coding schemes have been proposed for the asymmetric lossy setup.

The rest of the paper is organized as follows. The signal model is introduced in Section 2. In Section 3 the rate-distortion function of the Wyner-Ziv setup is characterized, establishing an upper bound to the rate loss with respect to a genie-aided setup, where the instantaneous realization of the state variable is available at the encoder and at the decoder. In Section 4 the implementation of practical solutions is discussed, and a Monte-Carlo Markov Chain (MCMC) [10] decoder based on a Minimum Mean Square Error (MMSE) estimator is proposed. Its performance is compared to the theoretical bounds in Section 5.

2. SIGNAL MODEL

The two correlated sources generate real-valued symbols, corresponding to the realizations of the random sequences $\{X_k\}_{k=1}^{+\infty}$ and $\{Y_k\}_{k=1}^{+\infty}$. Let X_k and Y_k represent the source and the side information symbols, respectively. Their instantaneous dependence is modeled using the additive channel $X_k = Y_k + Z_k$. The correlation noise Z_k and the side information Y_k are assumed independent. The symbols in $\{Y_k\}_{k=1}^{+\infty}$ are independent, identically distributed according to $\mathcal{N}(0, \sigma_y^2)$. The correlation noise sequence $\{Z_k\}_{k=1}^{+\infty}$ is distributed according to a hidden Markov model, with hidden state S_k and Gaussian emission. The state variable S_k takes values in $\{0, 1\}$. The sequence $\{S_k\}_{k=1}^{+\infty}$ is a time-invariant Markov process with transition probability matrix P , with

$$P_{ij} = \Pr(S_k = j | S_{k-1} = i) \quad \forall (i, j) \in \{0, 1\}^2, \quad (1)$$

and with stationary probability $p_i = \Pr(S_k = i)$. Let σ_i^2 be the variance associated with the realization $S = i$, so that $(Z_k | S_k = i)$ is distributed according to $\mathcal{N}(0, \sigma_i^2)$. It is assumed $\sigma_0^2 < \sigma_1^2$. In the following, let $\mathbf{X}_{1:N} = (X_1 \dots X_N)$. We consider the quadratic distortion measure $d(\mathbf{X}_{1:N}, \hat{\mathbf{X}}_{1:N}) = \frac{1}{N} \|\mathbf{X}_{1:N} - \hat{\mathbf{X}}_{1:N}\|^2$.

This signal model extends the one considered in [11], where $\{Z_k\}_{k=1}^{+\infty}$ is a Bernoulli-Gaussian process.

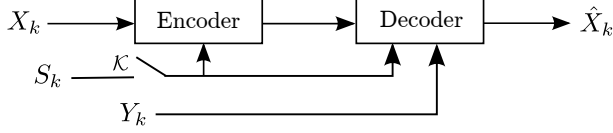


Fig. 1. Setup 1 (\mathcal{K} closed) and Setup 2 (\mathcal{K} open).

3. BOUNDS FOR THE RATE-DISTORTION CHARACTERISTIC

The aim of this section is the analysis of the Wyner-Ziv rate-distortion function $R_{X|Y}(D)$ for the signal model presented in Section 2. The Wyner-Ziv setup (Setup 2) is depicted in Figure 1, for Switch \mathcal{K} open. Since $R_{X|Y}(D)$ cannot be expressed in closed form, we characterize the upper bound to the rate loss in $R_{X|Y}(D)$ with respect to the theoretical performance $R_{X|Y,S}(D)$ of a genie-aided setup, where the realization of S_k is made available both at the encoder and at the decoder. The genie-aided setup (Setup 1) is depicted in Figure 1, for Switch \mathcal{K} closed.

Remarking that the random sequences defined in Section 2 are stationary and ergodic processes, the per-symbol rate-distortion function $R_{X|Y,S}(D)$ for Setup 1 is derived from [12] as

$$R_{X|Y,S}(D) = \lim_{N \rightarrow \infty} \frac{1}{N} R_{\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}}(D), \quad (2)$$

where $R_{\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}}(D)$ is defined by

$$R_{\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}}(D) = \inf I(\mathbf{X}_{1:N}; \mathbf{U}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}). \quad (3)$$

The minimization in (3) is on the pairs $(f_{\mathbf{U} | \mathbf{X}, \mathbf{S}}(\mathbf{u} | \mathbf{x}, \mathbf{s}), F_D(\cdot, \cdot, \cdot))$ such that the Markov chain $\mathbf{U} \leftrightarrow (\mathbf{X}, \mathbf{S}) \leftrightarrow (\mathbf{Y}, \mathbf{S})$ is satisfied, and $E[d(\mathbf{X}, F_D(\mathbf{U}, \mathbf{Y}, \mathbf{S}))] \leq D$ holds. Let $R_{X|Y,S}^M(D)$ be the rate-distortion function associated to Setup 1 when the state variable S is assumed memoryless, Bernoulli distributed, as in [11].

Similarly, the rate-distortion function $R_{X|Y}(D)$ for Setup 2 is defined as

$$R_{X|Y}(D) = \lim_{N \rightarrow \infty} \frac{1}{N} R_{\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}}(D) \quad (4)$$

where

$$R_{\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}}(D) = \inf I(\mathbf{X}_{1:N}; \mathbf{U}_{1:N} | \mathbf{Y}_{1:N}). \quad (5)$$

The minimization in (5) is on the pairs $(f_{\mathbf{U} | \mathbf{X}}(\mathbf{u} | \mathbf{x}), F_D(\cdot, \cdot))$ such that the Markov chain $\mathbf{U} \leftrightarrow \mathbf{X} \leftrightarrow \mathbf{Y}$ is satisfied, and $E[d(\mathbf{X}, F_D(\mathbf{U}, \mathbf{Y}))] \leq D$ holds. Let $R_{X|Y}^M(D)$ be the rate-distortion function associated to Setup 2 when the state variable S is assumed memoryless, Bernoulli distributed.

3.1. Setup 1, closed-form expression of $R_{X|Y,S}(D)$

Proposition 1 For the signal model defined in Section 2,

$$R_{X|Y,S}(D) = \sum_{i \in \{0,1\}} p_i \max \left(0, \frac{1}{2} \log_2 \left(\frac{\sigma_i^2}{D'} \right) \right), \quad (6)$$

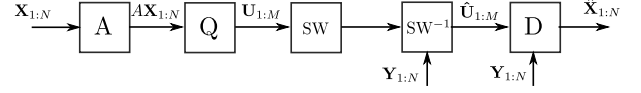


Fig. 2. Proposed coding scheme

where D' is such that $\sum_{i \in \{0,1\}} p_i \min(D', \sigma_i^2) \leq D$.

The proposition follows from the fact that the conditioning on $\mathbf{S}_{1:N}$ in (3) breaks the temporal dependence between the symbols in the sequence $\mathbf{X}_{1:N}$. As a result, $R_{X|Y,S}(D) = R_{X|Y,S}^M(D)$. The closed-form expression of $R_{X|Y,S}^M(D)$ is derived in [11]. This setup does not suffer rate loss compared to the joint coding case. Such a result is difficult to derive for Setup 2, as we only have an upper bound on performance.

3.2. Setup 2, rate loss for $R_{X|Y}(D)$

A lower bound to the rate-distortion function $R_{X|Y}(D)$ for Setup 2 is naturally provided by the performance of the genie-aided setup (Setup 1), $R_{X|Y,S}(D)$.

Proposition 2 For the signal model defined in Section 2,

$$R_{X|Y}(D) \leq R_{X|Y,S}(D) + L_{X|Y}(D) + \Lambda_{X|Y} \quad (7)$$

with

$$L_{X|Y}(D) = \frac{1}{2} \log_2 (1 + D/\sigma_\Phi^2) \quad (8)$$

$$\Lambda_{X|Y} = \min (H(S_k | S_{k-1}), h(\mathbf{Z}) - h(Z_k | S_k)), \quad (9)$$

and where $h(\mathbf{Z})$ is the differential entropy rate $h(\mathbf{Z}) = \lim_{N \rightarrow \infty} \frac{1}{N} h(\mathbf{Z}_{1:N})$.

The rate loss term (8) vanishes as $D \rightarrow 0$, and is due to the fact that no rate adaptation is possible at the encoder when the realization of S_k is not available. The rate loss term (9) is due to the uncertainty on S_k at the decoder side.

Sketch of the proof to Proposition 2 Consider a particular test-channel $(f_{\mathbf{U} | \mathbf{X}}(\mathbf{u} | \mathbf{x}), F_D(\cdot, \cdot))$ chosen in the minimization set of (5). The probability density function $f_{\mathbf{U} | \mathbf{X}}(\mathbf{u} | \mathbf{x})$ is chosen as Gaussian, by letting $\mathbf{U}_{1:N} = \mathbf{X}_{1:N} + \Phi_{1:N}$, with $\Phi_{1:N}$ independent on $\mathbf{X}_{1:N}$. The symbols Φ_k are independent, identically distributed according to $\mathcal{N}(0, \sigma_\Phi^2)$.

Since the MMSE estimator of $\mathbf{X}_{1:N}$ from $(\mathbf{U}_{1:N}, \mathbf{Y}_{1:N})$ is not known in closed form, the reconstruction function $\hat{\mathbf{X}}_{1:N} = F_D(\mathbf{U}_{1:N}, \mathbf{Y}_{1:N})$ is chosen as the LMMSE estimator. The variance σ_Φ^2 is chosen to match the target distortion D . The right part of (7) is obtained by evaluating the mutual information in (5) for this test channel, and taking the limit as $N \rightarrow \infty$ as in (4).

4. TOWARDS A PRACTICAL SETUP

The coding chain for the model introduced in Section 2 is depicted in Figure 2. It works on source sequences of length

N . We first apply a square invertible mixing matrix $A \in \mathbb{R}^{N \times N}$. The encoder then produces a quantized version $\mathbf{U}_{1:N}$ of $A\mathbf{X}_{1:N}$

$$\mathbf{U}_{1:N} = A\mathbf{X}_{1:N} + \Phi_{1:N}. \quad (10)$$

We assume a dithered lattice quantizer [13] in high resolution regime. Let σ_ϕ^2 denote the second moment (per dimension) of the lattice [13]. The encoder transmits the quantization index without losses to the decoder. We assume an idealized Slepian-Wolf (SW) coding chain, able to achieve the optimal rate given in [14] for ergodic sources.

In this section we focus on the design of the decoder which outputs the reconstruction $\hat{\mathbf{X}}_{1:N}$ of the source sequence. Since the distortion measure is quadratic, we consider MMSE estimation. Although the MMSE estimator cannot be characterized in analytical form, an asymptotically optimal implementation can be achieved via MCMC based methods. The decoder observes $\mathbf{Y}_{1:N}$ and the noisy description (10) of the source. The estimator considers the elements in Φ as Gaussian, *i.i.d.* distributed, with variance σ_ϕ^2 (worst-case assumption). We want to jointly estimate $\mathbf{X}_{1:N}$ and $\mathbf{S}_{1:N}$ from $\mathbf{Y}_{1:N}$ and $\mathbf{U}_{1:N}$. The joint posterior distribution can be expressed as

$$P(\mathbf{X}_{1:N}, \mathbf{S}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{U}_{1:N}) \propto P(\mathbf{U}_{1:N} | \mathbf{X}_{1:N}) P(\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}) P(\mathbf{S}_{1:N}) \quad (11)$$

with

$$P(\mathbf{U}_{1:N} | \mathbf{X}_{1:N}) \sim \mathcal{N}(A\mathbf{X}_{1:N}, \sigma_\phi^2 I_N) \quad (12)$$

$$P(\mathbf{X}_{1:N} | \mathbf{Y}_{1:N}, \mathbf{S}_{1:N}) \sim \mathcal{N}(\mathbf{Y}_{1:N}, R_s) \quad (13)$$

$$P(\mathbf{S}_{1:N}) = P(S_1) \prod_{k=2}^N P(S_k | S_{k-1}) \quad (14)$$

and $R_s = E[\mathbf{Z}_{1:N} \mathbf{Z}_{1:N}^T | \mathbf{S}_{1:N}]$. We estimate the joint posterior mean using a Gibbs sampler from the family of the MCMC algorithms. After a burning period, this algorithm provide samples of the joint posterior law. It means that we generate randomly sequences $\mathbf{X}_{1:N}$ and $\mathbf{S}_{1:N}$ according to their joint posterior distribution. The posterior mean is then estimated by avering over J samples. We cannot sample directly the joint posterior law (11) because it is too complex. Thus we use a Gibbs sampler algorithm. The principle of the Gibbs sampler is to sample under the conditional law iteratively. It is shown in [10] that asymptotically this algorithm provides true samples under the joint law.

Each iteration j of the algorithm can be decomposed into two steps described as follows.

1. Sample $X_{1:N}^{(j)}$ according to

$$P(\mathbf{X}_{1:N} | \mathbf{S}_{1:N} = \mathbf{s}_{1:N}^{(j-1)}, \mathbf{U}_{1:N} = \mathbf{u}_{1:N}, \mathbf{Y}_{1:N} = \mathbf{y}_{1:N}) = \frac{1}{(2\pi)^{N/2} R_x^{1/2}} \exp\left(-\frac{(\mathbf{X}_{1:N} - \mathbf{m}_X)^T R_x^{-1} (\mathbf{X}_{1:N} - \mathbf{m}_X)}{2}\right)$$

with

$$\begin{aligned} R_x &= (A^T A / \sigma_\phi^2 + R_s^{-1})^{-1} \\ \mathbf{m}_X &= R_x (A^T \mathbf{u}_{1:N} / \sigma_\phi^2 + R_s^{-1} \mathbf{y}_{1:N}) \end{aligned} \quad (15)$$

2. Sample $S_{1:N}^{(j)}$ sequentially in order to use the Markov chain property of $\mathbf{S}_{1:N}$ according to

$$P(S_1^{(j)} = 1 | X_1 = x_1^{(j)}, Y_1 = y_1, U_1 = u_1) = \left(1 + \frac{P(S_1 = 1)}{P(S_1 = 0)} \sqrt{\frac{\sigma_0^2}{\sigma_1^2}} e^{\left(-\frac{1}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) (x_1^{(j)} - y_1)^2\right)}\right)^{-1}$$

and $\forall k = 2 \dots N$,

$$P(S_k^{(j)} = 1 | X_k = x_k^{(j)}, Y_k = y_k, U_k = u_k, S_{k-1} = s_{k-1}^{(j)}) = \left(1 + \frac{P(S_k = 1 | s_{k-1}^{(j)})}{P(S_k = 0 | s_{k-1}^{(j)})} \sqrt{\frac{\sigma_0^2}{\sigma_1^2}} e^{\left(-\frac{1}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) (x_k^{(j)} - y_k)^2\right)}\right)^{-1}$$

The algorithm is initialized with a state sequence $\mathbf{S}_{1:N}^{(0)}$ sampled symbol by symbol from the *a priori* distribution of $\mathbf{S}_{1:N}$ (see Section 2). To finish an approximation of the MMSE estimate $\hat{\mathbf{X}}_{1:N}$ is calculated from the sampled sequences $\mathbf{X}_{1:N}^{(j)}$ as $\hat{\mathbf{X}}_{1:N} = \frac{1}{J} \sum_{j=b}^J \mathbf{X}_{1:N}^{(j)}$ where J is the number of iterations and b is the end of the burning period.

5. EXPERIMENTS

This section evaluates the performance of the proposed MCMC decoding scheme. The memory of the state variable S is characterized by $\mu = 1 - p_{01} - p_{10}$, see [7], where $p_{01} = \Pr(S_k = 1 | S_{k-1} = 0)$ and $p_{10} = \Pr(S_k = 0 | S_{k-1} = 1)$. The variance of Y_k , is chosen as $\sigma_y^2 = 1$ and the parameters associated to source S_k are $\sigma_0^2 = 0.04$, $\sigma_1^2 = 1$, $p_0 = 0.9$, $\mu = 0.8$. The transition probability matrix is derived from $p_0, p_1 = 1 - p_0$ and μ . The matrix A is a DCT matrix. Blocks of $N = 100$ source samples and $J = 2000$ iterations are considered.

We aim at evaluating the performance of the proposed decoder taking into account the characteristics of the coding scheme described in Figure 2. We assume asymptotic dithered lattice quantization, with $N \rightarrow \infty$. We first assume that the SW chain achieves the theoretical performance for ergodic sources, as in Section 4. As this assumption is not realistic, we provide also comparison with the performance of a system exploiting a state of the art SW chain for *i.i.d.* sources.

Under the assumption of ideal SW chain, the rate needed to transmit $\mathbf{U}_{1:N} = A\mathbf{X}_{1:N} + \Phi_{1:N}$ losslessly is given by [11]

$$\begin{aligned} R(\sigma_\phi^2) &= \frac{1}{N} h(\mathbf{U}_{1:N} | \mathbf{Y}_{1:N}) - \frac{1}{2} \log_2(2\pi e \sigma_\phi^2) \quad (16) \\ &= \frac{1}{N} h(\mathbf{Z}_{1:N} + \Phi_{1:N}) - \frac{1}{2} \log_2(2\pi e \sigma_\phi^2). \quad (17) \end{aligned}$$

An upper bound to $h(\mathbf{Z}_{1:N} + \Phi_{1:N})$ is evaluated numerically. For the fixed value σ_ϕ^2 , we hence evaluate the performance of the estimator. We generate 500 realizations of $\mathbf{X}_{1:N}$, $\mathbf{Y}_{1:N}$ and $\Phi_{1:N}$, and we decode the sequence to obtain $\hat{\mathbf{X}}_{1:N}$. The distortion $D(\sigma_\phi^2)$ is hence given by $D(\sigma_\phi^2) = \frac{1}{500} \sum_{i=1}^{500} \|\mathbf{X}_{1:N} - \hat{\mathbf{X}}_{1:N}\|_2^2$.

In the second case, the SW chain is designed to work with *i.i.d.* sources, and regards $\mathbf{Z}_{1:N}$ as composed by symbols distributed according to $\mathcal{N}(0, \sigma_z^2)$. This induces performance loss due to the fact that the temporal correlation between consecutive Z_k is not exploited to achieve further compression. In this case, practical coding solutions exist, providing rate $R(\sigma_\phi^2) \leq \frac{1}{2} \log_2 \left(1 + \sigma_z^2 / \sigma_\phi^2 \right)$. The distortion $D(\sigma_\phi^2)$ is evaluated as before.

To finish, we also compare our MCMC based algorithm to another solution that is a derived version of the Fast Bayesian Matching Pursuit Algorithm (FBMP) [15] adapted to the case with memory. All the results are presented in Figure 3 and compared to the theoretical lower and upper bounds proposed in (7). We observe only a little difference between the lower bound and the upper bound, meaning both are close to the true rate-distortion function. We see that the proposed decoding algorithm induces no loss in performance when it follows an ideal SW chain, unlike the FBMP algorithm. We also observe that the SW chain designed for the Gaussian approximation induces a loss of approximately 3 dB compared to the ideal one. A practical solution adapted to our model may be between the two curves.

6. CONCLUSION

This work introduces an HMM driven correlation model in the context of lossy source coding with side information at the decoder. This model may capture the bursty nature of source correlation, *e.g.*, in sensor networks. Bounds on the rate-distortion function are obtained for this model, and a practical decoder is proposed. It was evaluated by considering ideal quantizer and SW encoder and decoder.

Future work will be dedicated to the design of adapted quantizer and SW chain to our model with memory. Moreover, until now, only the cases with either perfect estimate of the states or no knowledge of the states have been considered. The next step may be to extend the results to imperfect knowledge of the states, or imperfect knowledge of the source correlation distribution.

7. REFERENCES

- [1] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.
- [2] R. Zamir, "The rate-loss in the Wyner-Ziv problem," *IEEE Trans. on Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, 1996.
- [3] R. Cristescu, B. Bekerull-Lozano, and M. Vetterli, "Networked

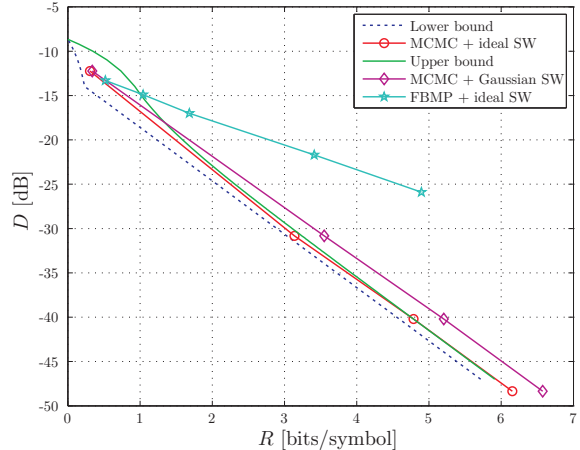


Fig. 3. Performance evaluation

Slepian-Wolf: theory, algorithms, and scaling laws," *IEEE Trans. on Inf. Theory*, vol. 51, no. 12, pp. 4057–4073, 2005.

- [4] M. Gastpar, P.L. Dragotti, and M. Vetterli, "The distributed Karhunen-Loeve transform," *IEEE Trans. on Inf. Theory*, vol. 52, no. 12, pp. 1–10, 2006.
- [5] T. Chu and Z. Xiong, "Coding of Gauss-Markov sources with side information at the decoder," in *Proc. IEEE WSSP*, 2003, pp. 26–29.
- [6] J. Del Ser, P.M. Crespo, and O. Galdos, "Asymmetric joint source-channel coding for correlated sources with blind HMM estimation at the receiver," *EURASIP Jnl on wirel. comm. and netw.*, vol. 2005, no. 4, pp. 483–492, 2005.
- [7] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Trans. on Inf. Theory*, vol. 35, no. 6, pp. 1277–1290, 1989.
- [8] J. Garcia-Frias, "Decoding of low-density parity-check codes over finite-state binary Markov channels," *IEEE Trans. on Comm.*, vol. 52, no. 11, pp. 1840–1843, 2004.
- [9] V. Toto-Zarasoia, A. Roumy, and C. Guillemot, "Hidden Markov model for distributed video coding," *Proc. of the ICIP, Hong Kong*, pp. 3709–3712, 2010.
- [10] C.P. Robert and G. Casella, *Monte-Carlo Statistical Methods*, Springer Texts in Statistics. Springer, New York, 2000.
- [11] F. Bassi, *Wyner-Ziv coding with uncertain side information quality*, Ph.D. thesis, Université Paris-Sud, 2010.
- [12] K. Iwata, "An information-spectrum approach to rate-distortion function with side information," in *Proc. IEEE ISIT*, 2000.
- [13] Zamir, R. and Feder, M., "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. on Inf. Theory*, vol. 38, no. 2, pp. 428–436, 1992.
- [14] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources (Corresp.)," *IEEE Trans. on Inf. Theory*, vol. 21, no. 2, pp. 226–228, 1975.
- [15] P. Schniter, L.C. Potter, and J. Ziniel, "Fast Bayesian matching pursuit," *Proc. ITA*, 2008.