

MIT Open Access Articles

Empirical rate-distortion study of compressive sensing-based joint source-channel coding

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Rambeloarison, Muriel L., Soheil Feizi, Georgios Angelopoulos, and Muriel Medard. "Empirical Rate-Distortion Study of Compressive Sensing-Based Joint Source-Channel Coding." 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR) (November 2012).

As Published: <http://dx.doi.org/10.1109/ACSSC.2012.6489217>

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

Persistent URL: <http://hdl.handle.net/1721.1/90414>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



Empirical Rate-Distortion Study of Compressive Sensing-based Joint Source-Channel Coding

Muriel L. Rambeloarison, Soheil Feizi, Georgios Angelopoulos, and Muriel Médard
Research Laboratory of Electronics
Massachusetts Institute of Technology, Cambridge, MA
Email: {muriel, sfeizi, georgios, medard}@mit.edu

Abstract—In this paper, we present an empirical rate-distortion study of a communication scheme that uses compressive sensing (CS) as joint source-channel coding. We investigate the rate-distortion behavior of both point-to-point and distributed cases.

First, we propose an efficient algorithm to find the ℓ_1 -regularization parameter that is required by the Least Absolute Shrinkage and Selection Operator which we use as a CS decoder.

We then show that, for a point-to-point channel, the rate-distortion follows two distinct regimes: the first one corresponds to an almost constant distortion, and the second one to a rapid distortion degradation, as a function of rate. This constant distortion increases with both increasing channel noise level and sparsity level, but at a different gradient depending on the distortion measure. In the distributed case, we investigate the rate-distortion behavior when sources have temporal and spatial dependencies. We show that, taking advantage of both spatial and temporal correlations over merely considering the temporal correlation between the signals allows us to achieve an average of a factor of approximately $2.5\times$ improvement in the rate-distortion behavior of the joint source-channel coding scheme.

I. INTRODUCTION

Compressive sensing (CS) is a novel technique which allows to reconstruct signals using much fewer measurements than traditional sampling methods by taking advantage of the sparsity of the signals to be compressed. Previous works related to the rate-distortion analysis of CS have been focused on its performance related to image compressing [1] and quantized CS measurements [2]. References [3], [4] and [5] derive bounds for the rate-distortion, while [6] presents a rate-distortion analysis by representing the compressive sensing problems using a set of differential equations derived from a bipartite graph. In a recent work [7], a joint source-channel-network coding scheme is proposed using compressive sensing for wireless network with AWGN channels. In this scheme, the sources exhibit both temporal and spatial dependencies, and the goal of the receivers is to reconstruct the signals within an allowed distortion level.

In this paper, we focus on the empirical rate-distortion behavior of this CS-based joint source-channel coding scheme using Least Absolute Shrinkage and Selection Operator (LASSO) [8] as a CS decoder and propose an algorithm to find the ℓ_1 -regularization parameter central to the LASSO optimization. We consider a point-to-point channel and illustrate how the rate-distortion varies as a function of channel noise level and sparsity level of the original signal. We also

investigate a distributed case, which highlights the significant advantage of taking the spatial and temporal dependencies of the sources we consider.

Our study shows that the rate-distortion behavior exhibits two distinct regimes for a point-to-point channel. For a number of CS measurements greater than some optimal value m^* , the distortion is almost constant. On the other hand, when fewer measurements than m^* are taken, the distortion degrades very rapidly with respect to the rate. Increased channel noise and sparsity level both influence the value of the distortion for the first regime, which increases accordingly.

For the distributed case, we consider a network with sources that have temporal and spatial dependencies. When both types of correlations are taken in consideration, we observe that the rate-distortion behavior of the network is on average 2.5 times better than that when only temporal dependencies are considered.

II. BACKGROUND AND PROBLEM SETUP

In this section, we review the fundamentals of compressive sensing (CS), introduce the cross-validation algorithm we use, and introduce the notation and parameters for our simulations.

A. Compressive Sensing

Let $\mathbf{X} \in \mathbb{R}^N$ be a k -sparse vector and let $\Phi \in \mathbb{R}^{m \times N}$ be measurement matrix such that $\mathbf{Y} = \Phi\mathbf{X}$ is the noiseless observation vector, where $\mathbf{Y} \in \mathbb{R}^m$. \mathbf{X} can be recovered by using $m \ll n$ measurements if Φ obeys the *Restricted Eigenvalue (RE) Condition* [7].

We consider noisy measurements, such that the measurement vector is $\mathbf{Y} = \Phi\mathbf{X} + \mathbf{Z}$, where \mathbf{Z} is a zero-mean random Gaussian channel noise vector.

It was shown in [8] that CS reconstruction can be formulated as a Least Absolute Shrinkage and Selection Operator (LASSO) problem, which is expressed as

$$\tilde{\mathbf{X}} = \arg \min_{\mathbf{X}} \frac{1}{2m} \|\mathbf{Y} - \Phi\mathbf{X}\|_{\ell_2}^2 + \lambda \|\mathbf{X}\|_{\ell_1} \quad (1)$$

where $\lambda \geq 0$ is the ℓ_1 -regularization parameter. By definition, given a vector \mathbf{X} and a solution $\tilde{\mathbf{X}}$, the LASSO problem involves a ℓ_1 -penalization estimation, which shrinks the estimates of the coefficients of $\tilde{\mathbf{X}}$ towards zero relative to their maximum likelihood estimates [8]. Equation (1) thus outputs a solution $\tilde{\mathbf{X}}$ that is desired to have a number of non-zero coefficients close to k , while maintaining a high-fidelity

reconstruction of the original signal. Thus, as λ is increased, so is the number of coefficients forced to zero.

In the next section, we propose an algorithm to choose λ using cross-validation, based on work by [9] and [10].

B. Cross-validation with modified bisection

As explained in [11], cross-validation is a statistical technique which allows to choose a model which best fits a set of data. It operates by dividing the available data into a training set to learn the model and a testing set to validate the model. The goal is then to select the model that best fits both the training and testing set.

We use a modified version of this algorithm to choose the value of λ which minimizes the energy of the relative error between some original signal and its reconstruction. As such, the $m \times N$ measurement matrix Φ in (1) is separated into a training and a cross-validation matrix, as shown in (2),

$$\Phi \in \mathbb{R}^{m \times N} \rightarrow \begin{cases} \Phi_{tr} \in \mathbb{R}^{m_{tr} \times N} \\ \Phi_{cv} \in \mathbb{R}^{m_{cv} \times N} \end{cases} \quad (2)$$

where $m_{tr} + m_{cv} = m$. In order for the cross-validation to work, Φ_{tr} and Φ_{cv} must be properly normalized and have the same distribution as Φ . For the purpose of the schemes we consider, we fix the number of cross-validation measurements at 10% of the total number of measurements, so $m_{cv} = \text{round}(0.1 m)$, which provide a reasonable trade-off between complexity and performance of the algorithm [9].

Algorithm 1 summarizes the cross-validation technique used to find the best value of λ for the rate-distortion simulations.

Algorithm 1 Cross-validation with modified bisection method

- 1: $\mathbf{Y}_{cv} = \Phi_{cv}\mathbf{X} + \mathbf{Z}_{cv}$
 - 2: $\mathbf{Y}_{tr} = \Phi_{tr}\mathbf{X} + \mathbf{Z}_{tr}$
 - 3: $\lambda = \lambda_{\text{init}}$
 - 4: Let ϵ be an empty vector with coefficients ϵ_i
 - 5: **while** $i \leq \text{MaxIterations}$ **do**
 - 6: Solve $\tilde{\mathbf{X}}_{tr}^{[\lambda]} = \arg \min_{\mathbf{X}} \frac{1}{2m} \|\mathbf{Y}_{tr} - \Phi_{tr}\mathbf{X}\|_{\ell_2}^2 + \lambda \|\mathbf{X}\|_{\ell_1}$
 - 7: $\epsilon_i \leftarrow \|\mathbf{Y}_{cv} - \Phi_{cv}\tilde{\mathbf{X}}_{tr}^{[\lambda]}\|_{\ell_2}$
 - 8: $\lambda \leftarrow \lambda/1.5$
 - 9: **end while**
 - 10: $\lambda^* = \arg \min_{\lambda} \epsilon = \arg \min_{\lambda} \|\mathbf{Y}_{cv} - \Phi_{cv}\tilde{\mathbf{X}}_{tr}^{[\lambda]}\|_{\ell_2}$
-

Given an original signal \mathbf{X} , the cross-validation and the training measurement vectors \mathbf{Y}_{cv} and \mathbf{Y}_{tr} are generated by taking the CS measurements and corrupting them with zero-mean Gaussian channel noise, represented by \mathbf{Z}_{cv} and \mathbf{Z}_{tr} (Lines 1 and 2). The initial value of λ that is investigated is one that we know leads to the all-zero reconstructed signal $\tilde{\mathbf{X}}_{tr}^{[\lambda]} = \mathbf{0}$ (Line 3). For a chosen number of repetitions, an estimation $\tilde{\mathbf{X}}_{tr}^{[\lambda]}$ of the reconstructed signal is obtained by decoding \mathbf{Y}_{tr} (Line 6) and the cross-validation error is computed (Line 7). The next value for λ to be investigated is obtained by dividing the current value by 1.5. The optimal value λ^* is then the one that minimizes the cross-validation error (Line 10).

In the field of CS, cross-validation mainly used with homotopy continuation algorithms such as LARS [12], which iterate over an equally-spaced range of decreasing values for λ . While this iterative process allows for better accuracy for smaller range steps, it comes at the cost of a latency which increases with the number of values of λ tested, due to the time-consuming decoding (Line 6). In our scheme, we circumvent this latency issue by considering a decreasing geometrical sequence of values of λ , which still guarantees that we find a solution for λ^* of the same order as the one predicted by an homotopy continuation algorithm, but in a fraction of the time. Indeed, we are able to obtain a solution after a maximum of 15 iterations of Lines 6 to 8, by using a method comparable to the bisection method [13] to obtain the values of λ to be tested. However, in order to improve the accuracy, we choose a common ratio of 1.5^{-1} instead of 2^{-1} . By abuse of notation, we refer to this technique as a ‘‘cross-validation with modified bisection method.’’

C. Simulations setup

In this section, we define the signal and measurement matrix models that were used for the simulations, the distortion measures used to obtain the rate-distortion results, as well as the software we use.

1) *Signal model and measurement matrix:* We consider a k -sparse signal \mathbf{X} of length $N = 1024$, and define its sparsity ratio as $k/N = \alpha$. \mathbf{X} is formed of spikes of magnitudes ± 1 and ± 0.5 , where each magnitude has a probability of $\alpha/4$.

We choose the measurement matrix Φ with a Rademacher distribution defined as follows

$$\Phi_{ij} = \frac{1}{\sqrt{m}} \begin{cases} -1 & \text{with probability } 0.5 \\ +1 & \text{with probability } 0.5 \end{cases} \quad (3)$$

where m is the number of measurements taken. It is shown in [14] that the RE condition holds for this type of matrix.

2) *Distortion measures:* We consider two distortion measures: the mean-squared error (*MSE*) and a scaled version of the percent root-mean-square difference (*PRD*) [15] often used to quantify errors in biomedical signals [15] and defined as follows:

$$PRD = \frac{\sqrt{\sum_{n=1}^N |\mathbf{X} - \tilde{\mathbf{X}}|^2}}{\sqrt{\sum_{n=1}^N |\mathbf{X}|^2}} \quad (4)$$

where \mathbf{X} is the original signal of length N and $\tilde{\mathbf{X}}$ its reconstruction.

The simulations were implemented in MATLAB using the software `cvx` [16], a modeling system for convex optimization which uses disciplined convex programming to solve (1) [17].

III. JOINT CS-BASED SOURCE-CHANNEL CODING FOR A POINT-TO-POINT CHANNEL

In this section, we evaluate the performance of a joint source-channel coding scheme using compressive sensing (CS) proposed in [7]. The signal and measurement models are defined in Section II-C. The sensing-communication scheme is performed in the following steps:

a) *Step 1 (Encoding)*: The CS encoding is done by taking m measurements of the signal \mathbf{X} of length $N = 1024$ using a measurement matrix $\Phi \in \mathbb{R}^{m \times N}$ distributed as in (3) to obtain a measurement vector $\mathbf{Y} = \Phi \mathbf{X}$.

b) *Step 2 (Transmission through channel)*: The measurement vector \mathbf{Y} is transmitted through a channel, which is either noiseless or noisy. If it is noisy, the standard deviation of the noise level is defined as a percentage of power of the signal \mathbf{Y} . For our simulations, we consider 5% and 10% channel noise. The signal reaching the receiver is $\mathbf{Z} = \mathbf{Y} + \mathbf{W} = \Phi \mathbf{X} + \mathbf{W}$, where $\mathbf{W} \in \mathbb{R}^m$ is additive zero-mean random Gaussian noise.

c) *Step 3 (Decoding)*: At the receiver, the LASSO decoder outputs an reconstructed signal $\tilde{\mathbf{X}}$ of \mathbf{X} by solving the following complex optimization

$$\tilde{\mathbf{X}} = \arg \min_{\mathbf{X}} \frac{1}{2m} \|\mathbf{Z} - \Phi \mathbf{X}\|_{\ell_2}^2 + \lambda^* \|\mathbf{X}\|_{\ell_1} \quad (5)$$

where we use Algorithm 1 to find λ^* .

Rate is calculated as m/N and we compare how both the channel noise level and the sparsity of the original signal affect the rate-distortion behavior of the scheme, for the PRD and MSE distortion measures. In these simulations, each point has been achieved by averaging the distortion values obtained by running each setting (channel noise, m , and sparsity ratio) 15 times.

A. Rate distortion as a function of noise level

We observe the rate-distortion behavior at 3 channel noise levels: noiseless, 5% and 10% channel noise. Figure 1 shows the rate-distortion in terms of *PRD* and *MSE* for a sparsity ratio $k/N = 0.075$.

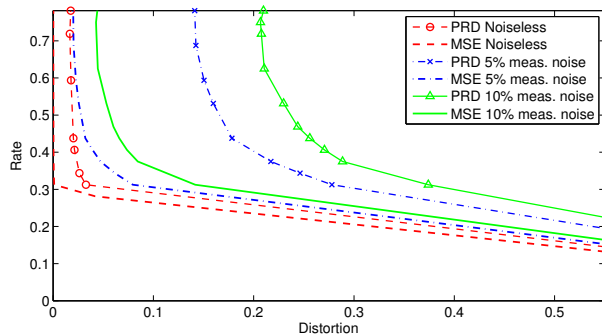


Fig. 1. Rate-Distortion for sparsity ratio $k/N = 0.075$

As seen in Figure 1, we can distinguish two regimes in the rate-distortion curves: the first one corresponds to an almost constant distortion D^* after the number of measurements exceeds some critical value m^* . As expected, both m^* and D^* increase slightly with increasing channel noise. However, we observe that this increase is much more important when *PRD* is used a distortion measure.

The second observed regime demonstrates a rapid degradation of the distortion, as the number of measurements is insufficient to properly reconstruct the original signal. This

rapid degradation corresponds to the settings of the simulations where the number of measurements is inferior to m^* .

B. Rate distortion as a function of sparsity level

We observe the rate-distortion behavior at 4 sparsity ratios $k/N = [0.01, 0.025, 0.05, 0.075]$ and present the corresponding rate-distortion curves in Figures 2 and 3. Both of these sets of curves correspond to a level of channel noise of 5%.

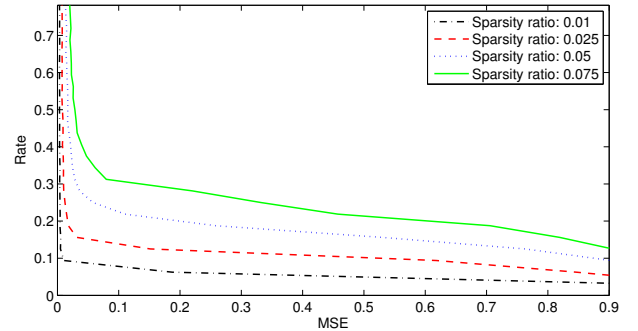


Fig. 2. Rate-Distortion for channel noise level of 5% with *MSE* as distortion measure

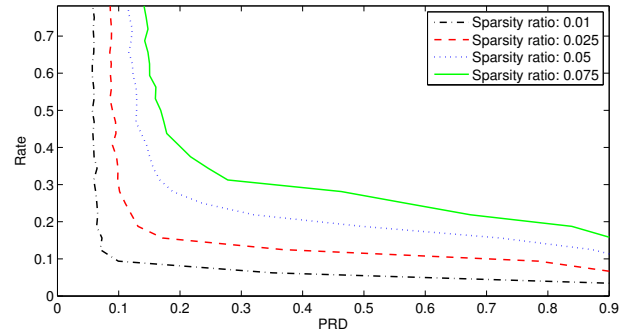


Fig. 3. Rate-Distortion for channel noise level of 5% with *PRD* as distortion measure

For a given noise level, we observe an upper-right shift of the curves for increasing sparsity ratio. In particular, we can see that the value of m^* increases almost linearly with the sparsity ratio. We also notice that the value of m^* increases much sharply when *MSE* is used as a distortion measure. As before, we can observe that the changes in rate-distortion curves are much distinguishable when the distortion measure is *PRD*.

IV. JOINT CS-BASED SOURCE-CHANNEL CODING FOR A DISTRIBUTED CASE

In this section, we evaluate the performance of the compressive sensing-based joint source-channel coding scheme for a distributed case. We consider a single-hop network depicted in Figure 4 with two sources s_1 and s_2 , whose samples exhibit both spatial and temporal redundancies [7]. The temporal redundancy refers to the fact that each signal is sparse; the

spatial redundancy refers to the fact that the difference between the two signals at the two sources is sparse.

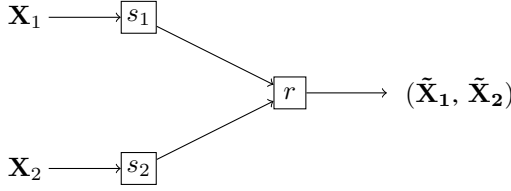


Fig. 4. Single-hop network for distributed cases

In our simulations, \mathbf{X}_1 is k_1 -sparse and $\mathbf{X}_2 = \mathbf{X}_1 + \mathbf{E}$, where \mathbf{E} is a k_2 -sparse error signal; we assume that $k_1 \gg k_2$. The goal is to reconstruct both $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$ at the receiver r . We present two ways of performing these reconstructions, and in both cases, the total rate and the distortion were respectively calculated as following

$$R_{total} = \frac{m_1 + m_2}{N} \quad (6)$$

$$D_{total} = D_1 + D_2 \quad (7)$$

where m_i is the number of compressive sensing measurements taken at source s_i and D_i is the distortion measured between the original and reconstructed signal \mathbf{X}_i and $\tilde{\mathbf{X}}_i$. For both of the cases, we present the results of the simulations for when the measurements are subjected to both no noise and 5% noise.

A. Case 1: Only temporal dependency is considered

In this case, we treat s_1 and s_2 as if there were two independent sources, that is \mathbf{X}_1 and \mathbf{X}_2 are compressed and decompressed independently. Algorithm 2 summarizes how this process is done.

Algorithm 2 Distributed case 1

- 1: $\mathbf{Y}_1 = \Phi_1 \mathbf{X}_1 + \mathbf{Z}_1$
 - 2: $\mathbf{Y}_2 = \Phi_2 \mathbf{X}_2 + \mathbf{Z}_2$
 - 3: Decompress \mathbf{Y}_1 to obtain $\tilde{\mathbf{X}}_1$ by solving
$$\tilde{\mathbf{X}}_1 = \arg \min_{\mathbf{X}_1} \frac{1}{2m} \|\mathbf{Y}_1 - \Phi_1 \mathbf{X}_1\|_{\ell_2}^2 + \lambda^* \|\mathbf{X}_1\|_{\ell_1}$$
 - 4: Decompress \mathbf{Y}_2 to obtain $\tilde{\mathbf{X}}_2$ by solving
$$\tilde{\mathbf{X}}_2 = \arg \min_{\mathbf{X}_2} \frac{1}{2m} \|\mathbf{Y}_2 - \Phi_2 \mathbf{X}_2\|_{\ell_2}^2 + \lambda^* \|\mathbf{X}_2\|_{\ell_1}$$
-

The signals that r receives are shown in Lines 1 and 2 of Algorithm 2, where \mathbf{Z}_i represents an additive zero-mean Gaussian noise associated with the channel. $\Phi_1 \in \mathbb{R}^{m_1 \times N}$ and $\Phi_2 \in \mathbb{R}^{m_2 \times N}$ are random matrices similar to (3).

Lines 3 and 4 of the algorithm correspond to the CS LASSO decoding performed at r to obtain estimates of the original signals \mathbf{X}_1 and \mathbf{X}_2 .

B. Case 2: Both spacial and temporal dependencies are considered

In this case, we take advantage of the spatial correlation between \mathbf{X}_1 and \mathbf{X}_2 , as shown in Algorithm 3.

Algorithm 3 Distributed case 2

- 1: $\mathbf{Y}_1 = \Phi_1 \mathbf{X}_1 + \mathbf{Z}_1$
 - 2: Decompress \mathbf{Y}_1 to obtain $\tilde{\mathbf{X}}_1$ by solving
$$\tilde{\mathbf{X}}_1 = \arg \min_{\mathbf{X}_1} \frac{1}{2m} \|\mathbf{Y}_1 - \Phi_1 \mathbf{X}_1\|_{\ell_2}^2 + \lambda^* \|\mathbf{X}_1\|_{\ell_1}$$
 - 3: $\mathbf{Y}_2 = \Phi_2 \mathbf{X}_2 + \mathbf{Z}_2$
 - 4: $\mathbf{Y}_2 = \Phi_2 (\mathbf{X}_1 + \mathbf{E}) + \mathbf{Z}_2$, and we already have an estimate for \mathbf{X}_1
 - 5: Let $\mathbf{Y}_E = \mathbf{Y}_2 - \Phi_1 \tilde{\mathbf{X}}_1$
 - 6: Thus $\mathbf{Y}_E = \Phi_2 \mathbf{E} + \mathbf{Z}_E$
 - 7: Decompress \mathbf{Y}_E to obtain $\tilde{\mathbf{E}}$ by solving
$$\tilde{\mathbf{E}} = \arg \min_{\mathbf{E}} \frac{1}{2m} \|\mathbf{Y}_E - \Phi_1 \tilde{\mathbf{X}}_1\|_{\ell_2}^2 + \lambda^* \|\mathbf{E}\|_{\ell_1}$$
 - 8: Hence $\tilde{\mathbf{X}}_2 = \tilde{\mathbf{X}}_1 + \tilde{\mathbf{E}}$
-

Lines 1 and 3 of Algorithm 3 corresponds to the signal received at r from source s_1 and s_2 respectively, where as before $\Phi_i \in \mathbb{R}^{m_i \times N}$ is generated using (3) and \mathbf{Z}_i is a random Gaussian noise vector corresponding to the noisy channel between s_i and r . We set $m_1 \gg m_2$. The receiver then uses the LASSO decoder to obtain $\tilde{\mathbf{X}}_1$ (Line 2). Given the spatial dependency between \mathbf{X}_1 and \mathbf{X}_2 , Lines 3 and 4 are equivalent for \mathbf{Y}_2 . The measurement vector \mathbf{Y}_E can thus be defined (Line 5), and decoded to obtain an estimate for the error \mathbf{E} (Line 7). Line 8 shows how $\tilde{\mathbf{X}}_2$ is computed as the sum $\tilde{\mathbf{X}}_1 + \tilde{\mathbf{E}}$.

The compared performance of the two algorithms for the distributed case are shown on Figures 5 to 8 for a noiseless and 5% channel noise settings. We observe that, for the noiseless channel, at a rate of 0.5, we obtain on average a factor of $2.5 \times$ improvement when using Algorithm 3 over Algorithm 2 with PRD as a distortion measure. When using MSE , an average improvement of almost $3 \times$ is obtained for the same setting.

When the channel is noisy, the similar average improvements at a rate of 0.5 are respectively factor of $2 \times$ and $2.5 \times$ for PRD and MSE . These results prove that taking advantage of the spatial and temporal correlations between the two signals allows to achieve a much improved rate-distortion behavior.

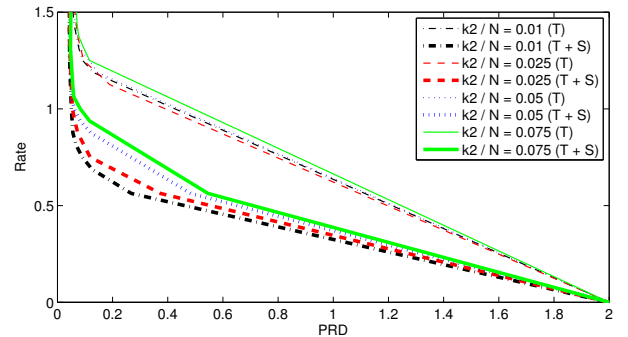


Fig. 5. Distributed Case: Noiseless channel with PRD as distortion measure, (T) is temporal correlation only case; (T+S) is temporal and spatial correlation case

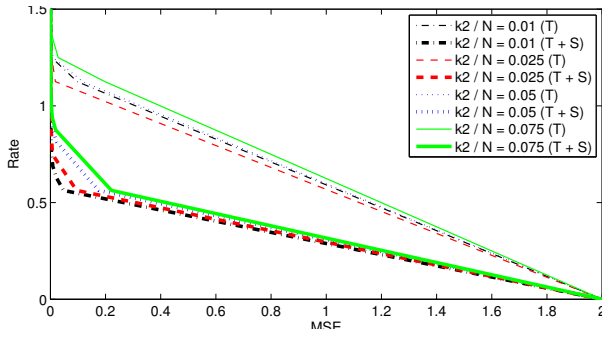


Fig. 6. Distributed Case: Noiseless channel with MSE as distortion measure, (T) is temporal correlation only case; (T+S) is temporal and spatial correlation case

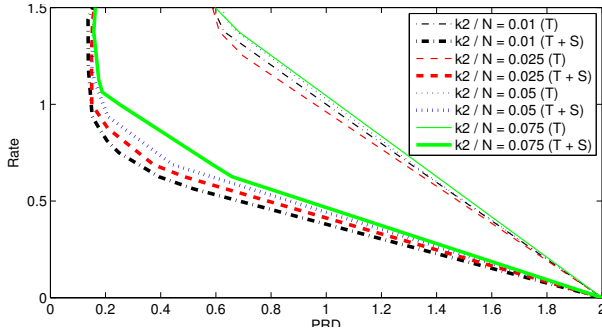


Fig. 7. Distributed Case: 5% channel noise with PRD as distortion measure, (T) is temporal correlation only case; (T+S) is temporal and spatial correlation case

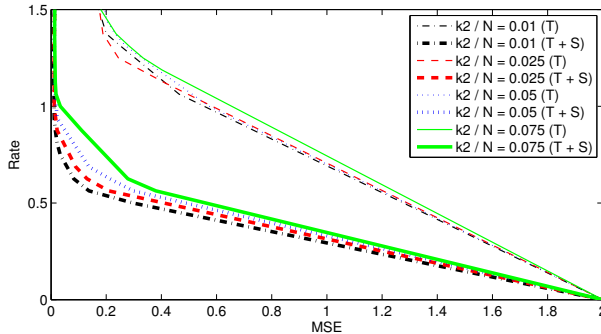


Fig. 8. Distributed Case: 5% channel noise with MSE as distortion measure, (T) is temporal correlation only case; (T+S) is temporal and spatial correlation case

V. CONCLUSIONS

In this paper, we empirically evaluated the rate-distortion behavior of a joint source-channel coding scheme, based on compressive sensing for both a point-to-point channel and a distributed case.

We first proposed an efficient algorithm to choose the ℓ_1 -regularization parameter λ from the LASSO, which we used as a compressive sensing decoder. This algorithm, which combines cross-validation and modified bisection, offers a reasonable trade-off between accuracy and computation time.

Using the values of λ obtained with this algorithm, we

characterized the rate-distortion behavior of the joint source-channel scheme in a point-to-point channel using two distortion measures and showed that there exists an optimal sampling rate above which the distortion remains relatively constant, and below which it degrades sharply.

We then studied a single-hop network with two spatially and temporally correlated sparse sources and a receiver which uses compressive sensing decoders to reconstruct the source signals. We observed the effect of these signal correlations on the rate-distortion behavior of the scheme and showed that taking both spatial and temporal correlation in consideration allows us to achieve a factor of $2.5\times$ improvement in rate-distortion compared to only taking temporal correlation.

REFERENCES

- [1] A. Schulz, L. Velho, and E. da Silva, "On the Empirical Rate-Distortion Performance of Compressive Sensing," in *2009 16th IEEE International Conference on Image Processing (ICIP 2009)*, November 2009, pp. 3049–3052.
- [2] W. Dai, H. V. Pham, and O. Milenkovic, "Distortion-Rate Functions for Quantized Compressive Sensing," in *2009 IEEE Information Theory Workshop on Networking and Information Theory (ITW 2009)*, June 2009, pp. 171–175.
- [3] B. Mulgrew and M. Davies, "Approximate Lower Bounds for Rate-Distortion in Compressive Sensing Systems," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011)*, no. 3849-3852, April 2008.
- [4] J. Chen and Q. Liang, "Rate Distortion Performance Analysis of Compressive Sensing," in *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, 2011, pp. 1–5.
- [5] A. K. Fletcher, S. Rangan, and V. K. Goyal, "On the Rate-Distortion Performance of Compressive Sensing," in *2007 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007)*, vol. 3, April 2007, pp. 885–888.
- [6] F. Wu, J. Fu, Z. Lin, and B. Zeng, "Analysis on Rate-Distortion Performance of Compressive Sensing for Binary Sparse Source," in *Data Compression Conference*, March 2009, pp. 113–122.
- [7] S. Feizi and M. Médard, "A Power Efficient Sensing/Communication Scheme: Joint Source-Channel-Network Coding by Using Compressive Sensing." Annual Allerton Conference on Communication, Control, and Computing, 2011.
- [8] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [9] R. Ward, "Compressed Sensing with Cross Validation," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 5773–5782, December 2009.
- [10] P. Boufounos, M. F. Duarte, and R. G. Baraniuk, "Sparse Signal Reconstruction from Noisy Compressive Measurements using Cross Validation," *IEEE/SP 14th Workshop on Statistical Signal Processing*, pp. 299–303, 2007.
- [11] P. Refailzadeh, L. Tang, and H. Liu, "Cross-validation," *Encyclopedia of Database Systems*, pp. 532–538, 2009.
- [12] B. Efron, J. Johnstone, I. Hastie, and R. Tibshirani, "Least Angle Regression," *Annals of Statistics*, vol. 32, pp. 407–499, 2004.
- [13] R. L. Burden and J. D. Faires, *Numerical Analysis*. PWS Publishers, 1985.
- [14] D. Achlioptas, "Database-friendly random projections: Johnson-Lindenstrauss with binary coins," *Journal of Computer and System Sciences*, vol. 66, no. 4, pp. 671–687, 2003.
- [15] F. Chen, F. Lim, O. Abari, A. Chandrakasan, and V. Stojanović, "Energy-Aware Design for Compressed Sensing Systems for Wireless Sensors under Performance and Reliability Constraints," *to be published*, 2011.
- [16] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 1.21," <http://cvxr.com/cvx>, April 2011.
- [17] —, "Graph Implementations for Nonsmooth Convex Programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes on Control and Information Sciences, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.