Distributed Sensing of Signals Under a Sparse Filtering Model

Ali Hormati, Olivier Roy, Yue M. Lu, Martin Vetterli

Distributed Sensing of Signals Under a Sparse Filtering Model

Ali Hormati , Olivier Roy , Yue M. Lu and Martin Vetterli

Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.

Abstract:
We consider the task of recovering correlated vectors at a central decoder based on fixed linear measurements obtained by distributed sensors. Two different scenarios are considered: In the case of universal reconstruction, we look for a sensing and recovery mechanism that works for *all* possible signals, whereas in the case of almost sure reconstruction, we allow to have a small set (with measure zero) of unrecoverable signals. We provide achievability bounds on the number of samples needed for both scenarios. The bounds show that *only* in the almost sure setup can we effectively exploit the signal correlations to achieve effective gains in sampling efficiency. In addition, we propose an efficient and robust distributed sensing and reconstruction algorithm based on annihilating filters.

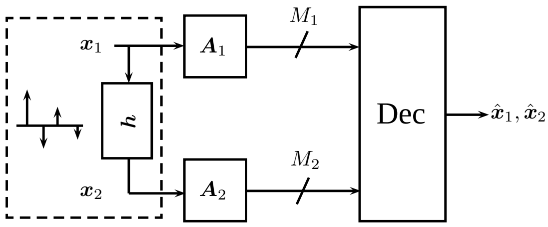


Figure 1: Distributed sensing setup. Signals $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are connected through an unknown sparse filter $\boldsymbol{h}$. The ith sensor ($i = 1, 2$) provides a M_i-dimensional observation of the signal $\boldsymbol{x}_i$ via a non-adaptive and fixed linear transform $\boldsymbol{A}_i$ to a central decoder.

1. Introduction

Consider two signals that are linked by an unknown filtering operation, where the filter is sparse in the time domain. Such models can be used, e.g., to describe the correlation between the transmitted and received signals in an unknown multi-path environment. We sample the two signals in a distributed setup: Each signal is observed by a different sensor, which sends a certain number of *non-adaptive* and *fixed* linear measurements of that signal to a central decoder. We study how the correlation induced by the above model can be exploited to reduce the number of measurements needed for perfect reconstruction at the central decoder, but *without* any inter-sensor communication during the sampling process.

Our setup is conceptually similar to the Slepian-Wolf problem in distributed source coding [6], which consists of correlated sources to be encoded separately and decoded jointly. While communication between the encoders is precluded, correlation between the measured data can be taken into account as an effective means to reduce the amount of information transmitted to the decoder. The main difference between our work and this classical distributed source coding setup is that we study a *sampling* problem and hence are only concerned about the number of sampling measurements we need to take, whereas the latter is about *coding* and hence uses bits as its “currency”. From the sampling perspective, our work is closely related to the problem of distributed compressed sensing, first introduced in [1] (see also [4, 5]). In that framework, jointly sparse data need to be reconstructed based on linear projections computed by distributed sensors. In this paper, we first introduce in Section 2. a novel correlation model for distributed signals. Instead of imposing any sparsity assumption on the signals themselves (as in [1]), we assume that the signals are linked by some unknown sparse filtering operation. Such models can be useful in describing the signal correlation in several practical scenarios (e.g. multi-path propagation and binaural audio recoding). In Section 3., we introduce two strategies for the design of the sampling system: In the *universal* strategy, we seek to successfully sense and recover *all* signals, whereas in the *almost sure* strategy, we allow to have a small set (with measure zero) of unrecoverable signals. We establish the corresponding achievability bounds on the number of samples needed for the two strategies mentioned above. These bounds indicate that the sparsity of the filter can be useful only in the almost sure strategy. Since the algorithms that achieves the bounds are computationally prohibitive, we introduce in Section 4., a concrete distributed sampling and reconstruction scheme that can recover the original signals in an efficient and robust way. Finally, Section 5. presents an application of our results in the area of binaural hearing aids. A preliminary version of this work was also presented at ICASSP 2009. In this paper, we add results on the achievability bound for the almost sure setup as well as a new section on applications.

2. The Correlation Model

Consider two signals $x_1(t)$ and $x_2(t)$, where $x_2(t)$ can be obtained as a filtered version of $x_1(t)$. In particular, we assume that

$$x_2(t) = (x_1 * h)(t)\,, \tag{1}$$

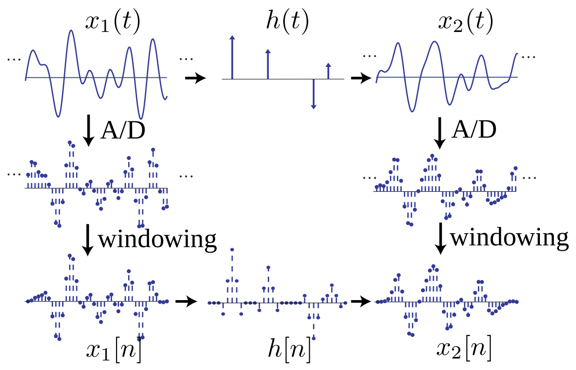


Figure 2: The continuous-time sparse filtering operation and its discrete-time counterpart.

where $h(t) = \sum_{k=1}^{K} c_k \delta(t - t_k)$ is a stream of K Diracs with *unknown* delays $\{t_k\}_{k=1}^{K}$ and coefficients $\{c_k\}_{k=1}^{K}$.
In this work, we study a finite-dimensional discrete version of the above model. As shown in Figure 2, we assume that the original continuous signal $x_1(t)$ is bandlimited to $[-\sigma, \sigma]$. Sampling $x_1(t)$ at uniform time interval T leads to a discrete sequence of samples $x_{s1}[n] \stackrel{\text{def}}{=} x_1(nT)$, where the sampling rate $1/T$ is set to be above the Nyquist rate σ/π. To obtain a finite-length signal, we subsequently apply a temporal window to the infinite sequence $x_{s1}[n]$ and get

$$x_1[n] \stackrel{\text{def}}{=} x_{s1}[n]\, w_N[n], \quad \text{for } n = 0, 1, ..., N-1,$$

where $w_N[n]$ is a smooth temporal window of length N. Note that when N is large enough, we can neglect the windowing effect, since $\widehat{w}_N(\omega)/(2\pi)$ approaches a Dirac function $\delta(\omega)$ as $N \to \infty$.
Applying the above procedure to $x_2(t)$ and using (1), we have

$$X_2[m] \approx \frac{1}{T} \widehat{x}_2 \left(\frac{2\pi m}{NT} \right) \approx X_1[m] H[m], \tag{2}$$

where

$$H[m] \stackrel{\text{def}}{=} \sum_{k=1}^{K} c_k e^{-j2\pi m t_k/(NT)}. \tag{3}$$

The above relationship implies that the finite-length signals $x_1[n]$ and $x_2[n]$ can also be approximately modeled as the input and output of a *discrete-time* filtering operation[1]. In general, the location parameters $\{t_k\}$ in (3) can be arbitrary real numbers, and consequently, the discrete-time filter $h[n]$ is no longer sparse (see Figure 2 for a typical impulse response of $h[n]$). However, when the sampling interval T is small enough, we can assume that the real-valued delays $\{t_k\}$ are close enough to the sampling grid, i.e., $t_k/T \approx n_k$ for some integers $\{n_k\}$. We will follow this assumption[2] throughout the paper.

Definition 1 (Correlation Model) *The signals of interest are two vectors $\boldsymbol{x}_1 = (x_1[0], \ldots, x_1[N-1])^T$ and $\boldsymbol{x}_2 = (x_2[0], \ldots, x_2[N-1])^T$, linked to each other through a circular convolution*

$$x_2[n] = (x_1 \circledast h)[n] \quad \textit{for } n = 0, 1, \ldots, N-1, \tag{4}$$

where $\boldsymbol{h} = (h[0], \ldots, h[N-1])^T \in \mathbb{R}^N$ is an unknown K-sparse vector, that is, $\|\boldsymbol{h}\|_0 = K$.

[1] Note that in order to be unambiguous in the positions $\{t_k\}$, we need to ensure that $NT > \max_k \{t_k\}$.

[2] We introduce this assumption (i.e. $t_k/T = n_k$ for some $n_k \in \mathbb{Z}$) mainly for the simplicity it brings to the theoretical analysis in later parts of this paper. It is however not an inherent limitation of our work.

3. Bounds

3.1 Universal Recovery

Let $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ be the sampling matrices used by the two sensors, and $\boldsymbol{A}$ be the block-diagonal matrix with $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ on the main diagonal. We first focus on finding those $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ such that every $\boldsymbol{x}^T = (\boldsymbol{x}_1^T, \boldsymbol{x}_2^T)$ is uniquely determined by its sampling data $\boldsymbol{A}\boldsymbol{x}$. Here we denote by $\mathcal{X}$ the set of all stacked vectors $\boldsymbol{x}$ such that its components $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ satisfy (4) for some K-sparse vector $\boldsymbol{h}$.

Definition 2 (Universal Achievability) *We say a sampling pair (M_1, M_2) is achievable for universal reconstruction if there exists fixed measurement matrices $\boldsymbol{A}_1 \in \mathbb{R}^{M_1 \times N}$ and $\boldsymbol{A}_2 \in \mathbb{R}^{M_2 \times N}$ such that the set*

$$\mathcal{B}(\boldsymbol{A}_1, \boldsymbol{A}_2) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathcal{X} : \exists \boldsymbol{x}' \in \mathcal{X} \textit{ with } \boldsymbol{x} \neq \boldsymbol{x}' \textit{ but } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}'\} \tag{5}$$

is empty.

Intuition suggests that, due to the correlation between the vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the minimum number of samples needed to perfectly describe all possible vectors $\boldsymbol{x}$ can made smaller than the total number of coefficients $2N$. The following proposition shows that, surprisingly, this is not the case.

Proposition 1 *A sampling pair (M_1, M_2) is achievable for universal reconstruction if and only if $M_1 \geq N$ and $M_2 \geq N$.*

Proof Let us consider two stacked vectors $\boldsymbol{x}^T = (\boldsymbol{x}_1^T, \boldsymbol{x}_2^T)$ and $\boldsymbol{x}'^T = (\boldsymbol{x}_1'^T, \boldsymbol{x}_2'^T)$, each following the correlation model (4). They can be written under the form

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{I}_N \\ \boldsymbol{C} \end{bmatrix} \boldsymbol{x}_1 \quad \text{and} \quad \boldsymbol{x}' = \begin{bmatrix} \boldsymbol{I}_N \\ \boldsymbol{C}' \end{bmatrix} \boldsymbol{x}_1',$$

where $\boldsymbol{C}$ and $\boldsymbol{C}'$ are circulant matrices with vectors $\boldsymbol{h}$ and $\boldsymbol{h}'$ as the first column, respectively. It holds that

$$\boldsymbol{x} - \boldsymbol{x}' = \begin{bmatrix} \boldsymbol{I}_N & -\boldsymbol{I}_N \\ \boldsymbol{C} & -\boldsymbol{C}' \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_1' \end{bmatrix}.$$

Moreover, we have that

$$\text{rank} \begin{bmatrix} \boldsymbol{I}_N & -\boldsymbol{I}_N \\ \boldsymbol{C} & -\boldsymbol{C}' \end{bmatrix} = N + \text{rank}\,(\boldsymbol{C} - \boldsymbol{C}').$$

When $\boldsymbol{C} - \boldsymbol{C}'$ is of full rank, the above matrix is of rank $2N$. This happens, for example, when $K = 1$ with $\boldsymbol{C} = 2\boldsymbol{I}_N$ and $\boldsymbol{C}' = \boldsymbol{I}_N$. In this case, $\boldsymbol{x} - \boldsymbol{x}'$ can take any possible values in $\mathbb{R}^{2N}$. Hence, a necessary (and sufficient) condition for the set (5) to be empty is that the block-diagonal matrix $\boldsymbol{A}$ is a $M \times 2N$-dimensional matrix of full rank, with $M \geq 2N$. In particular, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ must be full rank matrices of size $M_1 \times N$ and $M_2 \times N$, respectively, with $M_1, M_2 \geq N$. Note that, in the centralized scenario, the full rank condition would still require to take at least $2N$ measurements.

3.2 Almost Sure Recovery

As shown in Proposition 1, universal recovery is a rather strong requirement to satisfy since we have to take at least N samples at each sensor, without being able to exploit the existing correlation. In many situations, however, it is sufficient to consider a weaker requirement, which aims at finding measurement matrices that permit the perfect recovery of *almost all* signals from $\mathcal{X}$.

Definition 3 (Almost Sure Achievability) *We say a sampling pair (M_1, M_2) is achievable for almost sure reconstruction if there exist* fixed *measurement matrices $\boldsymbol{A}_1 \in \mathbb{R}^{M_1 \times N}$ and $\boldsymbol{A}_2 \in \mathbb{R}^{M_2 \times N}$ such that the set $\mathcal{B}(\boldsymbol{A}_1, \boldsymbol{A}_2)$, as defined in* (5)*, is of probability zero.*

The above definition for the almost sure recovery depends on the probability distribution of the signal $\boldsymbol{x}_1$ and the sparse filter $\boldsymbol{h}$. In what follows, it is sufficient to assume that the signal $\boldsymbol{x}_1$ and the non-zero coefficients of the filter $\boldsymbol{h}$ have non-singular[3] probability distributions over $\mathbb{R}^N$ and $\mathbb{R}^K$, respectively. The following proposition gives an achievability bound of the number of samples needed for the almost sure reconstruction.

Proposition 2 *A sampling pair (M_1, M_2) is achievable for almost sure reconstruction if*

$$\begin{aligned} M_1 &\geq \min\{K + r, N\}, \\ M_2 &\geq \min\{K + r, N\}, \\ \textit{and} \quad M_1 + M_2 &\geq \min\{N + K + r, 2N\}, \end{aligned} \tag{6}$$

where $r = 1 + \textit{mod}\,(K, 2)$.

Proof Due to space limitations, we just provide the sketch of the proof which is constructive in nature. First, let the two sensors take the Fourier transform of their signals and send the first $(K + r + 1)/2$ frequency components to the central decoder. By dividing the two sets of measurements (Note that the denominator should not be zero, which is guaranteed almost surely), the decoder calculates the necessary Fourier elements of the K-sparse filter $\boldsymbol{h}$ in order to reconstruct it almost surely. Then, the sensors transmit complementary subsets of frequency indices up to the Nyquist frequency. Knowing the filter $\boldsymbol{h}$ and the frequency content of one of the signals at some index, the decoder computes the corresponding frequency content of the other signal using (4).

Proposition 2 shows that, in contrast to the universal scenario, the correlation between the signals by means of the sparse filter provides a big saving in the almost sure setup, especially when $K \ll N$. This is depicted as the solid line in Figure 3.
Unfortunately, the algorithm that attains the bound in (6) is combinatorial in nature and thus, computationally prohibitive [1]. In the following, we propose a novel distributed sensing algorithm based on annihilating filters. This algorithm needs effectively K more measurements with respect to the achievability region for the almost sure reconstruction but exhibits polynomial complexity of $O(KN)$.

[3] By a non-singular distribution, we mean any continuous distribution such that the probability that the random variables lie in a low-dimensional subspace is zero.

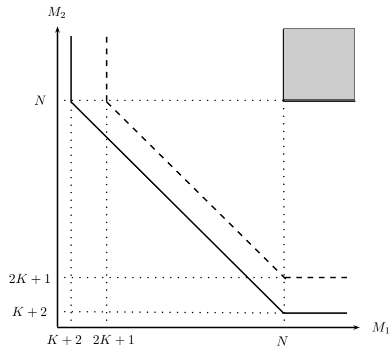


Figure 3: Achievable sampling region for universal reconstruction (shaded area), sampling pairs achieved for almost sure reconstruction for K odd (solid line) and sampling pairs achieved for almost sure reconstruction by the proposed algorithm based on annihilating filters (dashed line).

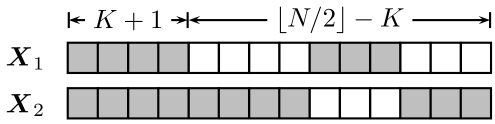


Figure 4: Sensors 1 and 2 both send the first $K + 1$ DFT coefficients of their observation, but only complementary subsets of the remaining frequency components.

4. Distributed Sensing Algorithm

The proposed distributed sensing scheme is based on a frequency-domain representation of the input signals. Let us denote by $\boldsymbol{X}_1 \in \mathbb{C}^N$ and $\boldsymbol{X}_2 \in \mathbb{C}^N$ the DFTs of the vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively. The circular convolution in (4) can be expressed as

$$\boldsymbol{X}_2 = \boldsymbol{H} \odot \boldsymbol{X}_1\,, \tag{7}$$

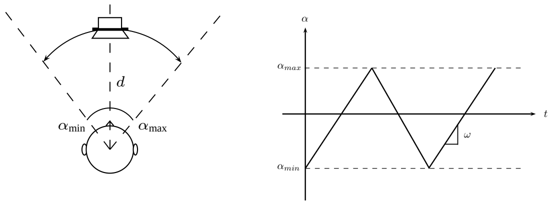
where $\boldsymbol{H} \in \mathbb{C}^N$ is the DFT of the filter $\boldsymbol{h}$ and $\odot$ denotes the element-wise product. Our approach consists of two main steps:

1. Finding filter $\boldsymbol{h}$ by sending the first $K + 1$ (1 real and K complex) DFT coefficients of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.
2. Sending the remaining frequency indices by sharing them among the two sensors.

The decoder first finds the filter $\boldsymbol{h}$ using only the first $K+1$ DFT coefficients of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. To this end, the decoder first computes

$$H[m] = \frac{X_2[m]}{X_1[m]} \quad \text{and} \quad H[-m] = H^*[m] \tag{8}$$

provided that $X_1[m]$ is non-zero for $m = 0, 1, \ldots, K$. This happens almost surely if the distribution of $\boldsymbol{x}_1$ is, for example, non-singular. Then, it finds the K-sparse filter with an annihilating filter approach; see [7] for details. The senors also transmit complementary subsets (in terms of frequency indexes) of the remaining DFT coefficients of their signals ($N - 2K - 1$ real values in total). This is illustrated in Figure 4. The first $K + 1$ DFT coefficients allow to almost surely reconstruct the filter $\boldsymbol{h}$. The missing frequency components of $\boldsymbol{x}_1$ (resp. $\boldsymbol{x}_2$) are then recovered from the available DFT coefficients of $\boldsymbol{x}_2$ (resp. $\boldsymbol{x}_1$) using the relation (7).



(a) (b)

Figure 5: Audio Experiment Setup. (a) A sound source travels at a distance of d meter in front of the head. (b) Angular position of the sound source with respect to time.

Note that in order to compute $X_1[m]$ from $X_2[m]$, the frequency components of the filter $H[m]$ should be nonzero. This is insured almost surely with our assumption that the nonzero elements of the filter $\boldsymbol{h}$ are chosen according to a non-singular distribution in $\mathbb{R}^K$. In terms of achievability, we have thus shown the following result.

Proposition 3 *A sampling pair (M_1, M_2) is achievable for almost sure reconstruction using the efficient annihilating filter method if*

$$\begin{aligned} M_1 &\geq \min\{2K+1, N\}, \\ M_2 &\geq \min\{2K+1, N\}, \\ \textit{and} \quad M_1 + M_2 &\geq \min\{N+2K+1, 2N\}. \end{aligned}$$

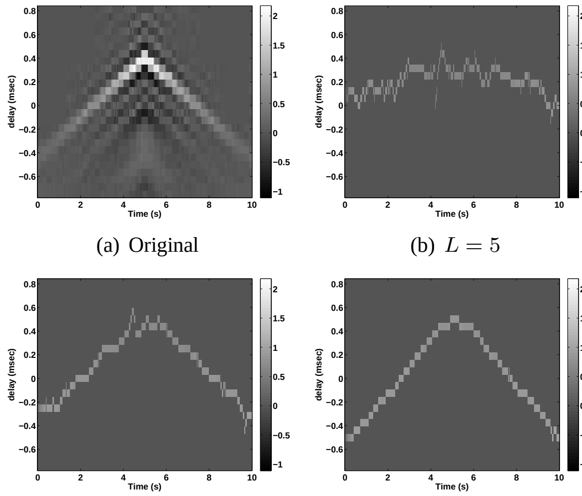
In the presence of noise or model mismatch, we add robustness to the system by sending $L+1$ DFT coefficients of x_i ($i = 1, 2$) with $L \geq K$ to the decoder. We denoise the measurements by using the denoising algorithm due to Cadzow; for details see [3]. Then the annihilating filter method uses the denoised measurements to estimate the sparse filter.

5. Application

In a practical scenario, we consider the signals recorded by two hearing aids mounted on the left and right ears of the user. We assume that the signals of the two hearing aids are related thorough a filtering operation. We refer to this filter as binaural filter. In the presence of a single source in far field, and neglecting reverberations and the head-shadow effect [2], the signal recorded at hearing aid 2 is simply a delayed version of the one observed at hearing aid 1. Hence, the binaural filter can be assumed to have sparsity factor $K = 1$. In the presence of reverberations and head shadowing, the filter from one microphone to the other is no longer sparse which introduces model mismatch. Despite this model mismatch, the transfer function between the two received signals should be approximately sparse, with the main peak indicating the desired relative delay.

In our setup, a single sound source located at distance $d = 1$ meter from the head of a KEMAR mannequin, moves back and forth between two angles $\alpha_{\min} = -45°$ and $\alpha_{\max} = 45°$. The angular speed of the source is $\omega = 18$ deg/sec. The sound is recorded by the microphones of the two hearing aids, located at the ears of the mannequin. We want to retrieve the binaural filter between the two hearing aids at hearing aid 1, from limited data transmitted by hearing aid 2. Then, the main peak of the binaural filter indicates the relative delay between the two received signals, which can be used to localize the source.

Figure 6 demonstrates the localization performance of the algorithm. Figure 6(a) shows the evolution of the original binaural impulse response over time. Figures 6(b)- 6(d) exhibits the sparse approximation to the filter, using different number of measurements. This clearly demonstrates the effect of the over-sampling factor on the robustness of the reconstruction algorithm.



(a) Original (b) $L = 5$

(c) $L = 15$ (d) $L = 25$

Figure 6: Tracking the binaural impulse response. Each column in the image corresponds to the binaural impulse response at the time mentioned on the x axis. (a) Original binaural filter. (b)-(d) Tracking the evolution of the main peak with different values of the oversampling factor L.

6. Conclusions

A general formulation of the distributed sensing problem has been proposed where the two signals are connected through an unknown sparse filter. In this context, both universal and almost sure reconstruction were addressed together with their corresponding achievable bounds. In addition, a distributed sensing scheme was presented, together with a method to make it robust to model mismatch. Our future research will focus on investigating more the applications of the proposed methods in the distributed sensing context.

References:

[1] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk. Distributed compressed sensing. Technical Report ECE-0612, Electrical and Computer Engineering Department, Rice University, Dec. 2006.

[2] J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA, 1997.

[3] J. A. Cadzow. Signal enhancement – A composite property mapping algorithm. *IEEE Trans. Acoust., Speech, Signal Process.*, 36(1):49–67, Jan. 1988.

[4] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, Feb. 2006.

[5] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, Apr. 2006.

[6] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory*, 19:471–480, Jul. 1973.

[7] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, 50(6):1417–1428, Jun. 2002.