

A Power Efficient Sensing/Communication Scheme: Joint Source-Channel-Network Coding by Using Compressive Sensing

Soheil Feizi
MIT
Email: sfeizi@mit.edu

Muriel Médard
MIT
Email: medard@mit.edu

Abstract—We propose a joint source-channel-network coding scheme, based on compressive sensing principles, for wireless networks with AWGN channels (that may include multiple access and broadcast), with sources exhibiting temporal and spatial dependencies. Our goal is to provide a reconstruction of sources within an allowed distortion level at each receiver. We perform joint source-channel coding at each source by randomly projecting source values to a lower dimensional space. We consider sources that satisfy the restricted eigenvalue (RE) condition as well as more general sources for which the randomness of the network allows a mapping to lower dimensional spaces. Our approach relies on using analog random linear network coding. The receiver uses compressive sensing decoders to reconstruct sources. Our key insight is the fact that, compressive sensing and analog network coding both preserve the source characteristics required for compressive sensing decoding.

I. INTRODUCTION

The power budget and communication rates are two main limitations in wireless network applications. Information theory focuses much on minimizing transmission rates, but the complexity of coding processes. For example, in point-to-point channels, equivalence theory allows us to separate the problem of channel coding from that of joint network/source coding [1]. However, any approach that separates source coding from channel coding may suffer from potential inefficiencies since we reduce redundancy for source coding but later reintroduce it for channel coding. Moreover, even if we allow separate channel coding from source/network coding (which is necessary because of lack of separation [2]), the known schemes have generally high complexity inherited from the general Slepian-Wolf problem [3].

Compressive sensing ([4], [5], [6]) can be used to perform channel coding ([7]) as well as source coding, even for distributed sources ([7]). In both of these cases, the performance of the compressive sensing approach can be shown to be close to that of the information theoretic optimum. The fact that, a single approach may successfully be applied for channel coding and source coding renders it an attractive candidate for use in a joint problem. Moreover, the fact that, linear analog network coding is asymptotically optimal in the high SNR regime [8], together with the fact that, compressive sensing-based channel and source techniques are also linear, provides a compelling reason for considering a compressive sensing-based joint source-channel-network coding.

This material is based upon work supported by AFOSR under award No. 016974-002.

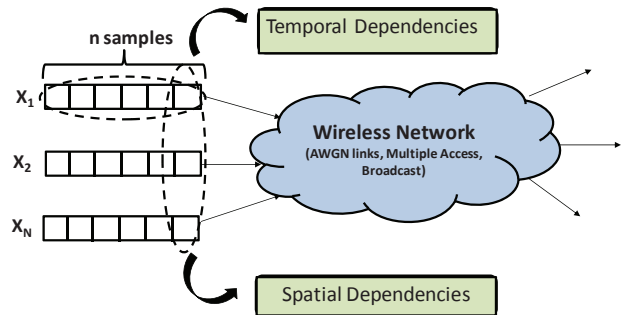


Fig. 1. The considered problem setup where sources have temporal and spatial dependencies, and receivers desire to reconstruct sources within an allowed distortion level.

Here, we propose a joint source-channel-network coding scheme, based on compressive sensing principles, for wireless networks with AWGN channels (that may include multiple access and broadcast). We assume that, sources exhibit temporal and spatial dependencies. We consider a multicast wireless network with multiple sources and receivers (see Figure 1). Our goal is to provide a reconstruction of all sources within an allowed distortion level at all receivers in a low-complexity manner. A key idea of our proposed scheme is to reshape inherent redundancies among source samples (temporal or spatial) in a joint source-channel-network coding scheme. Our scheme does not rely on the distribution knowledge of sources. Also, it provides an attractive relationship between the number of uses of the network and the reconstruction error.

A schematic view of our proposed framework is illustrated in Figure 2. In this framework, we perform joint source-channel coding at each source by randomly projecting source values to a lower dimensional space. We consider sources that spatially satisfy the restricted eigenvalue (RE) condition and may use sparse networks as well as more general sources, which use the randomness of the network to allow mapping to lower dimensional spaces. Our scheme relies on using random linear network coding in the real field. The receiver uses compressive sensing decoders (spatial and temporal ones) to reconstruct source signals.

Our main insight is that, the main characteristics required for operating compressive sensing are preserved under very

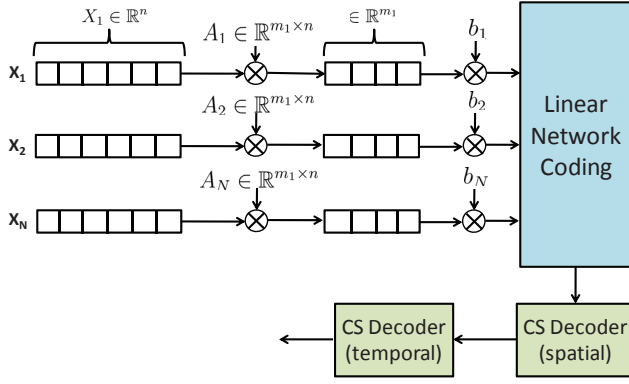


Fig. 2. A schematic view of our proposed joint source-channel-network coding scheme. Each source projects its high dimensional signal to a lower dimensional space by using a random projection matrix followed by a pre-coding on-off module. Intermediate nodes perform linear network coding. Each receiver uses compressive sensing decoders (LASSO) to reconstruct source signals.

general linear operations, which may themselves be arise from cascade of linear transformations in a network. We can therefore cascade compressive sensing techniques as well as linear operations tied to coding in the interior of the network. This preservation of the main characteristics of the sources under linear transformations lies at the core of both compressive sensing and analog network coding. This work constitutes a step in bringing together these two techniques, which have natural similarities.

This paper is organized as follows. In Section II, we review some prior results in compressive sensing and random projections. We propose our scheme in Section III and characterize its performance. Proofs are presented in Section IV. We conclude the paper in Section V.

II. COMPRESSIVE SENSING BACKGROUND

In this section, we review some prior results in compressive sensing. Let $\mathbf{Y} \in \mathbb{R}^N$ be an unknown signal vector. We say this signal is k -sparse if k of its coordinates are non-zero and the rest are zero. Note that, the assumption of having $N - k$ components of \mathbf{Y} to be exactly zero may seem unrealistic for practical cases. This assumption is called a hard sparsity assumption. Weakly sparse models are also considered in the compressive sensing literature [9]. Roughly, a vector \mathbf{Y} is weakly sparse if it can be approximated by a sparse vector. In this paper, we consider hard sparsity models although all discussions can be extended to weakly sparse signals.

Let $G \in \mathbb{R}^{m \times N}$ be a measurement matrix where $m \ll N$. The observation vector \mathbf{Z} is defined by the following linear model:

$$\mathbf{Z} = \mathbf{G}\mathbf{Y} + \mathbf{W},$$

where $\mathbf{W} \in \mathbb{R}^m$ is a noise vector. We assume $w_i \sim \mathcal{N}(0, \sigma^2)$, where $\mathcal{N}(0, \sigma^2)$ represents a Gaussian random variable with mean zero and variance σ^2 .

By obtaining observations \mathbf{Z} and knowing measurement

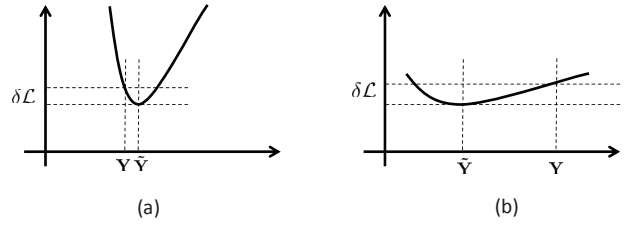


Fig. 3. To control the reconstruction error, the loss function should have high curvature around the optimal solution of LASSO (equation 1) along the directions of $\mathcal{C}(\mathcal{S}; \alpha)$.

matrix \mathbf{G} , the goal is to reconstruct $\tilde{\mathbf{Y}}$ such that

$$\|\mathbf{Y} - \tilde{\mathbf{Y}}\|_{l_2}^2 \leq D,$$

where D is a distortion allowance threshold. If the signal \mathbf{Y} is a sparse signal (i.e., $k \ll N$) and the measurement matrix satisfies the *Restricted Eigenvalue (RE) Condition* [10] (which will be explained later) or more restrictively, *Restricted Isometry Property* ([11]), then having m measurements where $m \ll N$ is sufficient to recover the original sparse vector. In fact, $\tilde{\mathbf{Y}}$ is a solution for the following convex optimization called LASSO ([12]):

$$\min_{\tilde{\mathbf{Y}}} \frac{1}{2m} \|\mathbf{Z} - \mathbf{G}\tilde{\mathbf{Y}}\|_{l_2}^2 + \xi \|\tilde{\mathbf{Y}}\|_{l_1}, \quad (1)$$

where $\|\cdot\|_{l_p}$ represents the l_p norm of a vector and $\xi > 0$ is a regularization parameter. If $\tilde{\mathbf{Y}}$ is a solution of this optimization, then, [10] shows that,

$$\|\mathbf{Y} - \tilde{\mathbf{Y}}\|_{l_2}^2 \leq \frac{\delta}{\gamma^2} \sigma^2 \frac{k \log(N)}{m} \quad (2)$$

with high probability, where δ is a constant and γ is a parameter related to the restricted eigenvalue condition of the matrix \mathbf{G} .

Now, we define the restricted eigenvalue condition (RE) for the measurement matrix $\mathbf{G} \in \mathbb{R}^{m \times N}$ ([10]). Suppose \mathcal{S} is the support of a sparse vector \mathbf{Y} (i.e., \mathcal{S} is the set of non-zero indices of \mathbf{Y} where $|\mathcal{S}| = k$). \mathcal{S}^c is the complement set of \mathcal{S} . We define the subset

$$\mathcal{C}(\mathcal{S}; \alpha) = \{\mathbf{Y} \in \mathbb{R}^N : \|\mathbf{Y}_{\mathcal{S}^c}\|_{l_1} \leq \alpha \|\mathbf{Y}_{\mathcal{S}}\|_{l_1}\}.$$

This corresponds to the cone of vectors where l_1 norm on their support dominates the l_1 norm off the support. For $\alpha = 1$, this cone is the same as the one used in restricted null-space property ([13]).

Definition 1. The matrix \mathbf{G} satisfies the restricted eigenvalue (RE) condition over \mathcal{S} with parameter (γ, α) if

$$\frac{\frac{1}{m} \|\mathbf{G}\mathbf{Y}\|_{l_2}^2}{\|\mathbf{Y}\|_{l_2}^2} \geq \gamma > 0 \quad (3)$$

for all $\mathbf{Y} \in \mathcal{C}(\mathcal{S}; \alpha) - \{0\}$.

The proof of equation (2) can be found in reference ([10]). By abuse of notation, we say a square matrix satisfies the

restricted eigenvalue condition with parameter (γ, α) if all of its sub-matrices with $|\mathcal{S}|$ rows satisfies the RE condition with parameter (γ, α) . Also, roughly we say a matrix satisfy the RE condition, if γ is large enough. In the following, we develop some intuitions about the compressive sensing decoder of equation (1) and the RE condition.

The restricted eigenvalue condition bounds away eigenvalues of the matrix $\frac{1}{m}\mathbf{G}^T\mathbf{G}$ from zero in the direction of the considered cone. Note that, to be able to control error of LASSO in equation (1), around the optimal point, the loss function should have high curvature to translate small l_2 loss to small error in \mathbf{Y} (see Figure 3). For this loss function, Hessian (curvature) can be calculated as $\frac{1}{m}\mathbf{G}^T\mathbf{G}$. Therefore, we want to bound eigenvalues of $\frac{1}{m}\mathbf{G}^T\mathbf{G}$ away from zero to have high curvature of the loss function. However, since \mathbf{G} is a fat matrix (i.e., $m \ll N$), some of its eigenvalues are zero. That is why we require to have eigenvalues away from zero for a subset $\mathcal{C}(\mathcal{S}; \alpha)$ of vectors. This subset is well-aligned with the directions of the Hessian in which loss is observed ([10]).

It is shown in different references that random matrices satisfy the RE condition. Reference [14] shows a connection between compressive sensing, n -widths, and the Johnson-Lindenstrauss Lemma. Through this connection, [15] proposes some data-base friendly matrices (binary matrices) satisfying the RE condition.

Our goal is in this paper is to use random projection principles to obtain a power-rate efficient sensing-communication scheme. New results and bounds in compressive sensing can potentially improve results of this paper. One can find more details on compressive sensing results in [4], [5], [6], etc.

In the next section, we present main results of this paper.

III. MAIN RESULTS

In this section, we present main results of this paper. First, we present a Lemma demonstrating how the RE condition changes when matrices are cascaded (multiplied to each other). Then, after introducing the problem setup and some notations, we present our proposed joint source-channel-network coding scheme. Finally, we characterize the performance of this proposed scheme. Proofs are presented in the next section.

A. Background Results

In this section, we present a Lemma which is central to implementing the cascading linear transformations which form the basis of our approach.

Lemma 2 (Cascading Lemma). *Suppose $\mathbf{G} \in \mathbb{R}^{m \times N}$ satisfies the RE condition with parameter (γ, α) . Say $\mathbf{C}_1 \in \mathbb{R}^{m \times m}$ and $\mathbf{C}_2 \in \mathbb{R}^{N \times N}$. Suppose minimum absolute eigenvalues of \mathbf{C}_1 and \mathbf{C}_2 are λ_1 and λ_2 , respectively. Also, say for $\mathbf{Y} \in \mathcal{C}(\mathcal{S}; \alpha)$, $\mathbf{Y}\mathbf{C}_2 \in \mathcal{C}(\mathcal{S}; \alpha)$. Then, $\mathbf{C}_1\mathbf{G}$ and $\mathbf{G}\mathbf{C}_2$ satisfy the RE condition with parameters $(\gamma\lambda_1^2, \alpha)$ and $(\gamma\lambda_2^2, \alpha)$, respectively.*

The proof of this Lemma is presented in Section IV. Intuitively, this lemma says, if \mathbf{G} , which satisfies the RE

condition, is cascaded with other matrices whose eigenvalues are bounded away from zero, the overall matrix would satisfy the RE condition.

B. The problem Setup and Notations

We consider a wireless network with N sources and l receivers. Each source has n correlated samples to transmit to all receivers (see Figure 1). Also, we assume that, there are temporal and spatial correlations among source samples. Vector $\mathbf{X}_i \in \mathbb{R}^n$ represents samples of source i and $\mathbf{X} \in \mathbb{R}^{N \times n}$ is a matrix of all source samples. x_i^t represents a sample of source i at time t .

We have the following assumptions on source samples capturing their temporal and spatial redundancies:

Assumption 3. *Sources have temporal and spatial dependencies:*

1) *Each source has temporal dependency: there exists a matrix $\Phi \in \mathbb{R}^{n \times n}$ such that $\mathbf{X}_i = \Phi\eta_i$, where η_i is k_1 sparse and $k_1 \ll n$.*

2) *Source samples have spatial dependency: suppose $\mathbf{Y} \in \mathbb{R}^N$ is a vector containing $y_j = \mathbf{A}_j^T \mathbf{X}_j$ for some vector $\mathbf{A}_j \in \mathbb{R}^n$, for $1 \leq j \leq N$. There exists a matrix $\Psi \in \mathbb{R}^{N \times N}$ such that, $\mathbf{Y} = \Psi\mu$, where μ is k_2 sparse, and $k_2 \ll N$.*

Note that, matrices Φ and Ψ are not required to be known at sources, but at receivers only. This is an important issue in hardware implementations of sensors ([16]). We will show that, our proposed scheme is generic.

Each receiver desires to obtain all sources within an allowed distortion level D . In other words, for all $1 \leq i \leq N$ and all receivers, we should have

$$\frac{1}{n} \|\mathbf{X}_i - \tilde{\mathbf{X}}_i\|_{l_2}^2 \leq D, \quad (4)$$

with high probability, where $\tilde{\mathbf{X}}_i$ is the reconstructed signal of source i .

Links of the network are AWGN channels. Nodes in the network are allowed to perform linear network coding in the real field (see analog network coding references [17], [8], etc.). We assume that, the network can deliver m linear combinations of source samples to each receiver at each use of the network. In other words, at each use of the network, a receiver obtains $\mathbf{Z} \in \mathbb{R}^m$ where, $\mathbf{Z} = \mathbf{G}\mathbf{Y} + \mathbf{W}$. $\mathbf{G} \in \mathbb{R}^{m \times N}$ is the network coding matrix, $\mathbf{Y} \in \mathbb{R}^N$ is a vector that sources transmit at that time (i.e., source i transmits y_i). \mathbf{W} is the noise vector where we assume, $w_i \sim N(0, \sigma^2)$, for all $1 \leq i \leq m$. The quantity m can be viewed as a real field version of the min-cut rate of the network defined in the finite field ([18], [19]). We assume that, the number of source samples (i.e., $n \times N$) is much larger than m . We define C_{use} as the number of network uses to deliver all source samples to receivers so that, the reconstruction distortion requirement (equation 4) holds for each receiver. We characterize C_{use} for our proposed sensing-communication framework and compare it with a naive information theoretic approach.

C. Joint Source-Channel-Network Coding Scheme

In this section, we present our proposed joint source-channel-network coding scheme based on compressive sensing principles. We also characterize the rate-distortion performance of our proposed scheme in Theorem 6.

Our proposed sensing-communication scheme can be performed in the following steps, as illustrated in Figure 2. We will explain each step in more detail.

Algorithm 4. *Our proposed joint source-channel-network coding can be performed in following steps.*

- **Step 1 (Temporal Pre-Coding):** Each source projects its n dimensional signal to a lower dimensional space by using random projection matrices.
- **Step 2 (Spatial Pre-Coding):** Each source i transmits with certain probability b_i , at each time.
- **Step 3 (Network Coding):** Nodes in the network perform analog random linear network coding.
- **Step 4 (Decoding):** Each receiver uses compressive sensing decoders (LASSO) to reconstruct source signals.

These steps are shown in a schematic way in Figure 2. In the following, we explain each step in more details:

Step 1 (Temporal Pre-Coding): Each source i projects its sample vector \mathbf{X}_i to a lower dimensional space, by using a matrix multiplication. That is, $\mathbf{Y}_i = \mathbf{A}_i \mathbf{X}_i$, where $\mathbf{X}_i \in \mathbb{R}^n$ and $\mathbf{A}_i \in \mathbb{R}^{m_1 \times n}$. Note that, $m_1 \ll n$, and m_1 is chosen according to Theorem 6 to control the reconstruction error. Matrix \mathbf{A}_i is a random matrix that satisfies the RE condition with parameter (γ_1, α) .

Step 2 (Spatial Pre-Coding): At time t , each source transmits y_i^t (the t -th component of \mathbf{Y}_i) with probability b_i . In the next step, we will show how by choosing appropriate b_i , the transmission power can be decreased further.

Step 3 (Network Coding): Nodes in the network perform random linear network coding over the real field. At each use of the network, it can deliver m linear combinations of transmitted samples to receivers. Therefore, m_2/m network transmissions are required to gather m_2 linear combinations of y_i^t at each receiver, with high probability. Note that, $m_2 \ll N$ is chosen according to Theorem 6 to control the reconstruction error. By abuse of notation, we drop superscript t . Hence, a receiver obtains:

$$\mathbf{Z} = \mathbf{G}\mathbf{B}\mathbf{Y} + \mathbf{W} = \mathbf{G}\mathbf{B}\Psi\boldsymbol{\mu} + \mathbf{W}, \quad (5)$$

where $\mathbf{Y} \in \mathbb{R}^N$ contains y_i for $1 \leq i \leq N$, $\mathbf{B} \in \mathbb{R}^{N \times N}$ is a diagonal matrix whose i -th diagonal element is one with probability b_i , and zero otherwise. $\mathbf{G} \in \mathbb{R}^{m_2 \times N}$ is the network coding matrix for that receiver. Note that, by using Assumption 3, $\mathbf{Y} = \Psi\boldsymbol{\mu}$, where $\boldsymbol{\mu}$ is k_2 sparse.

To be able to use compressive sensing principles and decoders, the matrix $\mathbf{G}\mathbf{B}\Psi$ should satisfy the RE condition explained in II. We consider two cases: (1) Ψ satisfies the RE condition, (2) Ψ does not satisfy the RE condition.

In case (1), since Ψ satisfies the RE condition, by using Lemma 2, $\mathbf{G}\mathbf{B}\Psi$ would satisfy the RE condition. Since the rank of $\mathbf{G}\mathbf{B}\Psi$ needs to be m_2 with high probability, we

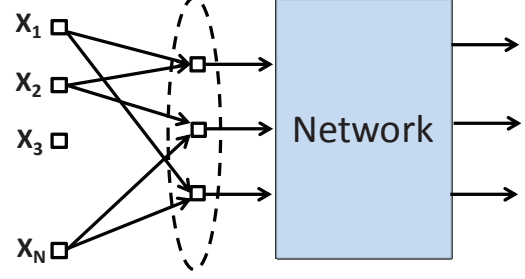


Fig. 4. An example of a network topology where the network coding matrix can be designed to satisfy the RE condition. Hence, the network can transform sources with spatial dependencies which do not satisfy the RE condition, to a proper lower dimensional space.

choose $b_i = 1$ with probability m_2/N , and $b_i = 0$ with probability $1 - m_2/N$. In other words, only a fraction of sources are transmitting. Also, by performing random linear network coding over the real field, rank of the matrix \mathbf{G} would be m_2 , with high probability ([20]). Hence, for a general network structure, receivers can use compressive sensing decoders to reconstruct source signals within an allowed distortion level.

In case (2) where Ψ does not satisfy the RE condition, we choose $b_i = 1$ for all i (i.e., all sources are transmitting). Hence, \mathbf{B} is the identity matrix. In this case, if \mathbf{G} satisfies the RE condition, by using Lemma 2, $\mathbf{G}\Psi$ would satisfy the RE condition, and therefore, receivers can use compressive sensing decoders to reconstruct source signals.

Intuitively, to have \mathbf{G} satisfy the RE condition, the network should be dense to randomly mix source samples. For networks that are not sufficiently dense, such as tree networks, \mathbf{G} never satisfies the RE condition. The following example presents a family of networks for which \mathbf{G} can be designed to satisfy the RE condition.

Example 5. *Consider the network topology with N sources depicted in Figure 4. Sources are connected randomly to m intermediate nodes. Expected in-degree of each of these intermediate nodes is greater than or equal to $N/3$ (i.e., high in-degree nodes). These intermediate nodes are connected to another network. The transform matrix of this network (i.e., $\mathbf{G} \in \mathbb{R}^{m_2 \times N}$) can be decoupled into two matrices: the one for the first stage of the network $\mathbf{G}_1 \in \mathbb{R}^{m \times N}$ and, the one for the rest, $\mathbf{G}_2 \in \mathbb{R}^{m_2 \times m}$. That is, $\mathbf{G} = \mathbf{G}_2\mathbf{G}_1$. If sources choose network coding coefficients uniformly from $\{-1, 1\}$, as shown in reference [15], the matrix \mathbf{G}_1 satisfies the RE condition. By performing random linear network coding, the rank of \mathbf{G}_2 is m with high probability. Therefore, by using Lemma 2, \mathbf{G} satisfies the RE condition.*

Step 4 (Decoding): Each receiver first uses a compressive sensing decoder as in equation (1) to reconstruct \mathbf{Y}^t for each $1 \leq t \leq m_1$ as follows:

$$\min_{\mu} \frac{1}{2m_2} \|\mathbf{Z} - \mathbf{G}\mathbf{B}\Psi\mu\|_{l_2}^2 + \xi \|\mu\|_{l_1}. \quad (6)$$

Therefore, for each source i , $\tilde{\mathbf{Y}}_i$ is obtained where $\tilde{\mathbf{Y}}_i = \mathbf{A}_i\Phi\mathbf{X}_i^s + \mathbf{U}_i$. \mathbf{U}_i is the error occurred in the reconstruction of \mathbf{Y}^t . Then, another compressive sensing decoder is used to reconstruct $\tilde{\mathbf{X}}_i$:

$$\min_{\eta_i} \frac{1}{2m_1} \|\tilde{\mathbf{Y}}_i - \mathbf{A}_i\Phi\eta_i\|_{l_2}^2 + \xi \|\eta_i\|_{l_1}. \quad (7)$$

The number of network uses of this proposed sensing-communication scheme is equal to $\mathcal{C}_{\text{use}} = \frac{m_1 m_2}{m}$.

$\mathbf{A}_i\Phi$ satisfies the RE condition with parameters $(\gamma_1\lambda_1^2, \alpha)$ when \mathbf{A}_i satisfies the RE condition with parameters (γ_1, α) and λ_1 is the minimum absolute eigenvalue of the matrix Φ (according to Lemma 2). The matrix $\mathbf{G}\mathbf{B}\Psi$ satisfies the RE condition with parameters $(\gamma_2\lambda_2^2, \alpha)$ (i.e., in case 1, matrix $\mathbf{B}\Psi_S$ satisfies the RE condition with parameter (γ_2, α) and λ_2 is the minimum absolute eigenvalue of \mathbf{G}^S where \mathbf{G}^S contains columns of \mathbf{G} corresponding to active sources. In case 2, \mathbf{G} satisfies the RE condition with parameters (γ_2, α) and λ_1 is the minimum absolute eigenvalue of the matrix Ψ). Let λ_3 and λ_4 be maximum eigenvalues of matrices Ψ and Φ , respectively. By using these parameters, define the following constant:

$$c = \delta \frac{(\lambda_3\lambda_4)^2}{(\gamma_1\gamma_2)^2(\lambda_1\lambda_2)^4}, \quad (8)$$

where δ is a constant. The following Theorem illustrates how to choose m_1 and m_2 in the proposed joint source-channel-network coding scheme to reconstruct each source signal within an allowed distortion level:

Theorem 6. *A joint source-channel-network coding scheme of Algorithm 4 satisfies the distortion requirement of equation (4) with high probability if*

$$\mathcal{C}_{\text{use}} = c \frac{k_1 k_2 \log(n) \log(N) \sigma^2}{m D}, \quad (9)$$

where $k_1 \ll n$ and $k_2 \ll N$ represent temporal and spatial sparsity of sources, σ^2 is the noise variance, D is the allowed reconstruction distortion level, and c is a constant defined in equation (8).

The proof of this Theorem is presented in Section IV. This theorem provides a way of choosing parameters m_1 and m_2 in Algorithm 4, where each source signal can be reconstructed within an allowed distortion level (equation (4)).

Remarks:

- Our proposed scheme is efficient: Sources use naturally occurring redundancies to perform joint source-channel coding. Therefore, there is no need of performing source coding and channel coding separately which would cause inefficiencies. Also, in the case that Ψ satisfies the RE condition, only a fraction of sources are required

to transmit, which provides additional power saving in transmissions.

- Our proposed scheme is broadly structure independent: Matrices Φ and Ψ are not required to be known at sources, but at receivers. It is important in the hardware implementation of sensors for different applications where measured signals can have different correlation structures ([16]).
- Our proposed scheme has a continuous rate-distortion performance: as it can be seen from Theorem 6, if the noise variance σ^2 increases, by using the same sensing-communication scheme, the reconstruction distortion would increase in a continuous manner (i.e., there is no performance drop off). In fact, our proposed scheme can be viewed as a *noise modulating* sensing-communication scheme.
- Our proposed scheme has low complexity decoders. Unlike Slepian-Wolf ([3]) decoders of information theoretic coding schemes, decoders used in our scheme are convex optimizations and therefore have lower complexity.
- It is insightful to compare the rate-distortion performance of our scheme with a naive information theoretic scheme ignoring source correlations. Following the rate-distortion Theorem ([1]), to represent a Gaussian random variable $\mathcal{N}(0, \sigma^2)$ within a distortion level D , one needs $R(D) = \frac{1}{2} \log(\frac{\sigma^2}{D})$ bits. Suppose sources transmit their samples ignoring correlations. Since at one network use, the network can deliver m linear combinations of source samples, $\frac{nN}{m}$ network uses are required. Moreover, since the noise variance is σ^2 but the allowed distortion level is D , roughly speaking, one needs at least $R(D)$ extra factor of transmissions. Overall,

$$\frac{nN}{m} \log\left(\frac{\sigma^2}{D}\right)$$

network uses are required. Note that, $k_1 \ll n$ and $k_2 \ll N$. Comparing this with the one of our proposed scheme (Theorem 6) illustrates that, although our scheme is power efficient, it performs closely to an optimal scheme where only an informative subspace of source samples are transmitted.

- Note that, projection matrices \mathbf{A}_i should be known at receivers. Transmitting these matrices may cause a rate overhead in a practical sensing-communication scheme. In practical implementations, since this matrix is a random matrix, each source can use a pseudo random matrix generated (i.e., pseudo-random bit sequence (PRBS) generators) and transmit its starting seed, instead of transmitting the whole matrix. This would have a negligible effect in the sensing-communication performance. Analyzing our proposed scheme with pseudo random matrices are beyond the scope of this paper.
- Note that, the result of Theorem 6 is not deterministic and holds with high probability. There is always a small chance that, reconstructed signals have higher distortion than the allowed threshold.
- As the last point, our proposed framework allows to

have *distributed sensing-communication*. That is, suppose there is one signal that receivers desire to reconstruct it. Each sensor has a low power budget to take all measurements and perform codings by itself. However, if several sensors (sources) are used, instead, each can take few measurements from the signal. By using our proposed scheme, receivers can reconstruct the original signal by using these measurements. In fact, our framework makes a virtual *strong* sensor by using several *weak* sensors.

IV. PROOFS

In this section, we present proofs.

A. Proof of Lemma 2

We first show that, $\mathbf{C}_1\mathbf{G}$ satisfies the RE condition with parameters $(\gamma\lambda_1^2, \alpha)$ where \mathbf{G} satisfies the RE condition with parameters (γ, α) and λ_1 is the minimum absolute eigenvalue of \mathbf{C}_1 . We have,

$$\begin{aligned} \frac{1}{m}\|\mathbf{C}_1\mathbf{G}\mathbf{Y}\|_{l_2}^2 &= \frac{1}{m}\mathbf{Y}^T\mathbf{G}^T\mathbf{C}_1^T\mathbf{C}_1\mathbf{G}\mathbf{Y} \\ &= \frac{1}{m}(\mathbf{G}\mathbf{Y})^T(\mathbf{C}_1^T\mathbf{C}_1)(\mathbf{G}\mathbf{Y}) \\ &\stackrel{(1)}{\geq} \frac{1}{m}\lambda_1^2\|\mathbf{G}\mathbf{Y}\|_{l_2}^2 \\ &\stackrel{(2)}{\geq} \gamma\lambda_1^2\|\mathbf{Y}\|_{l_2}^2 \end{aligned}$$

Inequality (1) comes from a linear algebra lemma and inequality (2) comes from the definition of the RE condition for the matrix \mathbf{G} .

In the following, we prove the second half of this Lemma. We assume that, for $\mathbf{Y} \in \mathcal{C}(\mathcal{S}; \alpha)$, $\mathbf{Y}\mathbf{C}_2 \in \mathcal{C}(\mathcal{S}; \alpha)$. Also, suppose λ_2 is the minimum absolute eigenvalue of the matrix \mathbf{C}_2 . Therefore,

$$\begin{aligned} \frac{1}{m}\|\mathbf{G}\mathbf{C}_2\mathbf{Y}\|_{l_2}^2 &= \frac{1}{m}\mathbf{Y}^T\mathbf{C}_2^T\mathbf{G}^T\mathbf{G}\mathbf{C}_2\mathbf{Y} \\ &= \mathbf{Y}^T\left(\frac{1}{m}\mathbf{C}_2^T\mathbf{G}^T\mathbf{G}\mathbf{C}_2\right)\mathbf{Y} \\ &= (\mathbf{C}_2\mathbf{Y})^T\left(\frac{1}{m}\mathbf{G}^T\mathbf{G}\right)(\mathbf{C}_2\mathbf{Y}) \\ &\stackrel{(3)}{\geq} \gamma\|\mathbf{C}_2\mathbf{Y}\|_{l_2}^2 \\ &\stackrel{(4)}{\geq} \gamma\lambda_2^2\|\mathbf{Y}\|_{l_2}^2 \end{aligned}$$

Inequality (3) comes from the definition of the RE condition for the matrix \mathbf{G} . Note that, in this part, we use the assumption that, for $\mathbf{Y} \in \mathcal{C}(\mathcal{S}; \alpha)$, $\mathbf{Y}\mathbf{C}_2 \in \mathcal{C}(\mathcal{S}; \alpha)$. Inequality (4) comes from a linear algebra lemma. This completes the proof.

B. Proof of Theorem 6

In this part, we prove Theorem 6. Suppose $\mathbf{A}_i\Phi$ satisfies the RE condition with parameters $(\gamma_1\lambda_1^2, \alpha)$ (i.e., \mathbf{A}_i satisfies the RE condition with parameters (γ_1, α) and λ_1 is the minimum absolute eigenvalue of the matrix Φ). Then, use

Lemma 2). Also, say the matrix $\mathbf{G}\mathbf{B}\Psi$ satisfies the RE condition with parameters $(\gamma_2\lambda_2^2, \alpha)$ (i.e., in case 1, matrix $\mathbf{B}\Psi_{\mathcal{S}}$ satisfies the RE condition with parameter (γ_2, α) and λ_2 is the minimum absolute eigenvalue of $\mathbf{G}^{\mathcal{S}}$ where $\mathbf{G}^{\mathcal{S}}$ contains columns of \mathbf{G} , corresponding to active sources. In case 2, \mathbf{G} satisfies the RE condition with parameters (γ_2, α) and λ_1 is the minimum absolute eigenvalue of the matrix Ψ). Hence, by using the LASSO decoder of equation (6) and the error bound of equation (2), we have

$$\|\mu - \tilde{\mu}\|_{l_2}^2 \leq \frac{\delta_1}{\gamma_2^2\lambda_2^4} \frac{k_2 \log(N)}{m_2} \sigma^2. \quad (10)$$

Hence,

$$\begin{aligned} \|y_i - \tilde{y}_i\|_{l_2}^2 &\stackrel{(1)}{\leq} \|\mathbf{Y} - \tilde{\mathbf{Y}}\|_{l_2}^2 \\ &\stackrel{(2)}{=} \|\Psi\mu - \Psi\tilde{\mu}\|_{l_2}^2 \\ &\stackrel{(3)}{\leq} \lambda_3^2\|\mu - \tilde{\mu}\|_{l_2}^2 \\ &\stackrel{(4)}{\leq} \frac{\delta_1\lambda_3^2}{\gamma_2^2\lambda_2^4} \frac{k_2 \log(N)}{m_2} \sigma^2. \end{aligned} \quad (11)$$

Inequality (1) is true for all vectors. Equality (2) uses the spatial correlation matrix property where $\mathbf{Y} = \Psi\mu$. Inequality (3) uses a linear algebra lemma. Inequality (4) uses equation (10).

Say $\sigma_u^2 = \|y_i - \tilde{y}_i\|_{l_2}^2$. For the decoder of equation (7), we can write,

$$\begin{aligned} \|x_i^t - \tilde{x}_i^t\|_{l_2}^2 &\stackrel{(5)}{\leq} \|\mathbf{X}_i - \tilde{\mathbf{X}}_i\|_{l_2}^2 \\ &\stackrel{(6)}{=} \|\Phi\eta - \Phi\tilde{\eta}\|_{l_2}^2 \\ &\stackrel{(7)}{\leq} \lambda_4^2\|\eta - \tilde{\eta}\|_{l_2}^2 \\ &\stackrel{(8)}{\leq} \frac{\delta_2\lambda_4^2}{\gamma_1^2\lambda_1^4} \frac{k_1 \log(n)}{m_1} \sigma_u^2 \\ &\stackrel{(9)}{\leq} c \frac{k_1 k_2 \log(n) \log(N)}{m_1 m_2} \sigma^2 \end{aligned}$$

where $c = \delta_1\delta_2 \frac{(\lambda_3\lambda_4)^2}{(\gamma_1\gamma_2)^2(\lambda_1\lambda_2)^4}$.

Inequality (5) is true for all vectors. Equality (6) uses the temporal correlation matrix property where $\mathbf{X}_i = \Phi\eta$. Inequality (7) uses a linear algebra lemma. Inequality (8) uses equation (2). Inequality (9) uses equation (11). Having the distortion requirement $\frac{1}{n}\|\mathbf{X}_i - \tilde{\mathbf{X}}_i\|_{l_2}^2 \leq D$ and $C_{\text{use}} = \frac{m_1 m_2}{D}$ completes the proof of Theorem 6.

V. CONCLUSIONS

In this paper, we proposed a joint source-channel-network coding scheme, based on compressive sensing principles, for wireless networks with AWGN channels (that may include multiple access and broadcast). A key idea of our proposed scheme is to reshape inherent redundancies among source samples (temporal or spatial) in joint source-channel-network coding to have a power/rate efficient sensing communication scheme. By characterizing the performance of the proposed scheme, we showed that, our proposed scheme

- (1) is power and rate efficient, that is, although it performs closely to optimal rates, it provides power efficiencies.
- (2) has low decoding complexity where decoders are convex optimizations.
- (3) is broadly structure independent, that is, sensing, coding and communication processes in sensors are application independent.
- (4) and has continuous rate-distortion performance, that is, if channel qualities change, the reconstruction quality will change in a continuous manner.

In this framework, we perform joint source-channel coding at each source by randomly projecting source values to a lower dimensional space. We consider sources that spatially satisfy the restricted eigenvalue (RE) condition and may use sparse networks as well as more general sources which use the randomness of the network to allow mapping to lower dimensional spaces. The receiver uses compressive sensing decoders to reconstruct source signals.

REFERENCES

- [1] I. Csiszar and J. Körner, in *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, New York, 1981.
- [2] A. Ramamoorthy, K. Jain, P. Chou, and M. Effros, "Separating distributed source coding from network coding," *Information Theory, IEEE Transactions on*, vol. 52, no. 6, pp. 2785–2795, 2006.
- [3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.
- [4] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [5] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, p. 969, 2007.
- [6] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [7] S. Feizi, M. Medard, and M. Effros, "Compressive sensing over networks," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 1129–1136.
- [8] I. Maric, A. Goldsmith, and M. Médard, "Analog network coding in the high-snr regime," in *2010 IEEE Wireless Network Coding Workshop*, Boston, US, 2010.
- [9] D. Donoho, I. Johnstone, A. Maleki, and A. Montanari, "Compressed sensing over l_p -balls: Minimax mean square error," *arXiv:1103.1943v2*, 2011.
- [10] G. Raskutti, M. Wainwright, and B. Yu, "Restricted eigenvalue properties for correlated gaussian designs," *The Journal of Machine Learning Research*, vol. 11, pp. 2241–2259, 2010.
- [11] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [12] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [13] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *American Mathematical Society*, vol. 22, no. 1, pp. 211–231, 2009.
- [14] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
- [15] 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.
- [16] F. Chen, A. Chandrakasan, and V. Stojanovic, "A signal-agnostic compressed sensing acquisition system for wireless and implantable sensors," in *Custom Integrated Circuits Conference (CICC), 2010 IEEE*. IEEE, pp. 1–4.
- [17] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2007, pp. 397–408.
- [18] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, 2000.
- [19] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. on networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
- [20] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.