COMPRESSED ACQUISITION AND PROGRESSIVE RECONSTRUCTION OF MULTI-DIMENSIONAL CORRELATED DATA IN WIRELESS SENSOR NETWORKS

Markus Leinonen, Marian Codreanu, Markku Juntti

Department of Communications Engineering and Centre for Wireless Communications
University of Oulu, P.O. Box 4500, 90014, Oulu, Finland
email: markus.leinonen@ee.oulu.fi

ABSTRACT

This paper addresses compressed acquisition and progressive reconstruction of spatially and temporally correlated signals in wireless sensor networks (WSNs) via compressed sensing (CS). We propose a novel method based on sliding window processing, where the sink periodically collects CS measurements of sensor samples, and then, instantaneously reconstructs current WSN samples by exploiting the spatio-temporal correlation via Kronecker sparsifying bases. By using previous estimates as prior information, the method can progressively improve the reconstruction accuracy of the signal ensemble. Furthermore, the method can control the trade-off between decoding delay and complexity. Numerical results demonstrate that the proposed method can recover WSN data samples from CS measurements with higher reconstruction accuracy, yet with lower decoding delay and complexity, as compared to the state of the art methods.

Index Terms— Compressed sensing, spatio-temporal correlation, sliding window processing, joint signal recovery, Kronecker sparsifying bases, multi-hop wireless sensor networks

1. INTRODUCTION

Wireless sensor networks (WSNs) consisting of multiple battery-powered sensors have been frequently proposed for monitoring or measuring various types of natural phenomena, e.g., light, temperature or humidity. Typically, the observed sensor readings have both temporal and spatial dependency. This dependency can be utilized by a joint data acquisition and reconstruction method, compressed sensing (CS) [1–6], which allows a signal of length N to be accurately recovered from its $M < N$ linear measurements.

Especially, the CS has established a promising foundation for the development of energy efficient data gathering methods in multi-hop WSNs [7–13]. By exploiting spatial (i.e., inter-signal) or temporal (i.e., intra-signal) correlation of the sensors via CS, the amount of data traffic for data delivery can be reduced. For spatially correlated data, compressed acquisition can be performed, e.g., by linearly combining sensor measurements along multi-hop routing [9–11, 13] or by collecting only a fraction of sensor readings [12, 13]. For temporally correlated data, each sensor can transmit CS measurements of a block of its buffered data samples [13–15]. However, as a drawback, this induces decoding delay for the estimates.

Apart from these works dealing with correlated signals in single dimension, CS has been also proposed for utilizing joint intra- and inter-signal dependencies. In [7, 16], joint sparsity models were proposed for representing particular joint dependencies of signal ensembles. Recently, Kronecker compressed sensing (KCS) [17] was proposed for the acquisition of correlated multi-dimensional signals via Kronecker sparsifying bases, which encode distinct correlation patterns of each signal dimension into single basis. As the KCS features blocks of data, the increased decoding delay and complexity may arise as substantial restrictions for an application.

In this paper, we consider compressed acquisition and progressive reconstruction of signals with spatio-temporal correlation in multi-hop data gathering WSNs. As the main contribution, we propose a novel WSN data acquisition method based on sliding window processing and Kronecker sparsifying bases [17], where the sink can instantaneously reconstruct current data samples from periodically collected CS measurements. Furthermore, the method uses multiple previously decoded estimates to progressively improve the accuracy of signal estimates and can control the trade-off between the decoding delay and complexity, making it suitable for delay-constrained applications. Related works on CS reconstruction with prior information utilization include [18–20], which consider general sparse dynamic signal frameworks. The numerical experiments show that our proposed method can recover WSN data samples with higher reconstruction accuracy, yet with lower decoding delay and complexity, as compared to the state of the art methods.

The paper is organized as follows. The WSN model is defined in Section 2. The CS background is recapitulated in Section 3. Section 4 discusses CS acquisition of multi-dimensional correlated signals. Our proposed method is derived in Section 5. Finally, Section 6 presents numerical results and Section 7 concludes the paper.

2. NETWORK MODEL

We consider a single-sink multi-hop WSN consisting of a set of randomly deployed sensors indexed by integers $\mathcal{N} = \{1, \ldots, N\}$. At each sensing period, they monitor a phenomenon to acquire T data samples. The WSN readings are assumed to encompass spatial and temporal dependency, as illustrated in Fig. 1 for $N = 10$ and $T = 6$.

The sensors, capable of transmitting, receiving and relaying data, communicate in a pre-defined, fixed multi-hop routing structure. We assume appropriate medium access control and scheduling for the system to support feasible data traffic across the wireless links. The modeling of detailed packet transmission protocols and physical layer parameters are outside the scope of this paper.

3. BACKGROUND ON COMPRESSED SENSING

Let $\boldsymbol{x} \in \mathbb{R}^N$ be a real-valued vector. It can be represented in basis $\boldsymbol{\Psi} \in \mathbb{R}^{N \times N}$ as $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\theta}$, where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$ are the trans-

This research was supported by Academy of Finland and Infotech Oulu Doctoral Program.

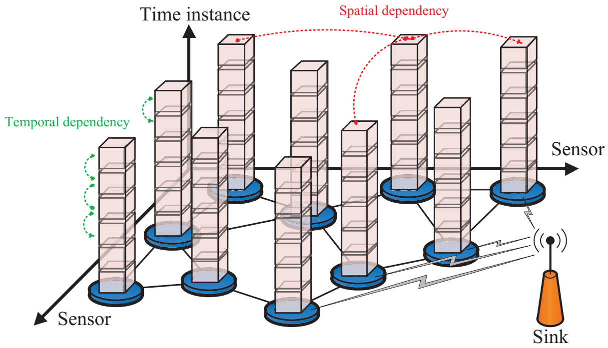


Fig. 1. A wireless sensor network with temporally and spatially correlated data samples with $N = 10$ and $T = 6$.

form domain coefficients. We say that $\boldsymbol{x}$ is K-sparse in $\boldsymbol{\Psi}$ if $\boldsymbol{\theta}$ has K non-zero entries. For natural signals, which typically are not exactly sparse, K-compressibility plays a role similar to K-sparsity in CS recovery performance [21]. Namely, the transform domain coefficients of a compressible signal, when ordered according to their magnitudes, typically decay rapidly according to a power law [2].

By the CS, a signal $\boldsymbol{x}$ that is K-sparse (or K-compressible) in $\boldsymbol{\Psi}$ can be accurately reconstructed from $M < N$ linear (random) measurements $\boldsymbol{y} = \boldsymbol{\Phi x} = \boldsymbol{\Phi\Psi\theta} = \boldsymbol{\Upsilon\theta}$, where $\boldsymbol{y} \in \mathbb{R}^M$ are the measurements, $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ is a measurement matrix and $\boldsymbol{\Upsilon} = \boldsymbol{\Phi\Psi} \in \mathbb{R}^{M\times N}$ is the sensing matrix [6]. The signal can be reconstructed via sparsity-encouraging L_1-minimization as $\hat{\boldsymbol{\theta}} := \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1$ s.t. $\boldsymbol{y} = \boldsymbol{\Phi\Psi\theta}$, resulting in $\hat{\boldsymbol{x}} = \boldsymbol{\Psi}\hat{\boldsymbol{\theta}}$ [21].

Two key features for CS performance are the restricted isometry property (RIP) and mutual coherence [17]. The RIP of $\boldsymbol{\Upsilon}$ ensures the measurements to approximately preserve the Euclidean length of all K-sparse signals, whereas the mutual coherence between $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ determines the number of required measurements for successful recovery [6]. Accordingly, e.g., i.i.d. Gaussian and Bernoulli $\boldsymbol{\Phi}$ are highly incoherent with any $\boldsymbol{\Psi}$, satisfying the RIP (with overwhelming probability) for $M \geq CK\log(N/K)$ with a constant $C > 0$ [6].

4. COMPRESSED ACQUISITION OF MULTI-DIMENSIONAL CORRELATED DATA

Let $\boldsymbol{X} \in \mathbb{R}^{T\times N}$ denote the 2D signal ensemble, which consists of T data readings from each randomly positioned sensor $i \in \mathcal{N}$. A (t, i)th entry of $\boldsymbol{X}$, denoted as x_{ti}, represents tth data sample of sensor $i \in \mathcal{N}$. The structure of $\boldsymbol{X}$ can be written as

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] = [\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^T]^{\mathrm{T}}, \tag{1}$$

where column $\boldsymbol{x}_i \in \mathbb{R}^T$ represents the data samples of sensor $i \in \mathcal{N}$ and row $\boldsymbol{x}^t \in \mathbb{R}^N$ are the WSN data readings at time instance t.

As depicted in Fig. 1, we assume spatio-temporal dependency in sensor signal ensemble $\boldsymbol{X}$. Then, each of its signal dimension has a sparse (or compressible) representation in a proper basis, denoted as $\boldsymbol{\Psi}_{\mathrm{T}} \in \mathbb{R}^{T\times T}$ for temporal and $\boldsymbol{\Psi}_{\mathrm{S}} \in \mathbb{R}^{N\times N}$ for spatial domain, respectively. Thus, each WSN data sample vector $\boldsymbol{x}^t$, $t = 1, \ldots, T$, can be represented as $\boldsymbol{x}^t = \boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\theta}_{\mathrm{S},t}$, where $\boldsymbol{\theta}_{\mathrm{S},t} \in \mathbb{R}^N$ are the $K_{\mathrm{S},t}$-compressible coefficients in the spatial domain. By stacking the coefficients as $\boldsymbol{\Theta}_{\mathrm{S}} = [\boldsymbol{\theta}_{\mathrm{S},1}, \ldots, \boldsymbol{\theta}_{\mathrm{S},T}]$, the spatial transformation of signal $\boldsymbol{X}$ is compactly represented as $\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\Theta}_{\mathrm{S}}$, i.e.,

$$[\boldsymbol{x}^1, \ldots, \boldsymbol{x}^T] = \boldsymbol{\Psi}_{\mathrm{S}}[\boldsymbol{\theta}_{\mathrm{S},1}, \ldots, \boldsymbol{\theta}_{\mathrm{S},T}]. \tag{2}$$

Similarly, each sensor signal $\boldsymbol{x}_i$, $i = 1, \ldots, N$, can be represented as $\boldsymbol{x}_i = \boldsymbol{\Psi}_{\mathrm{T}}\boldsymbol{\theta}_{\mathrm{T},i}$, where $\boldsymbol{\theta}_{\mathrm{T},i} \in \mathbb{R}^T$ are the $K_{\mathrm{T},i}$-compressible coefficients in the temporal domain. By introducing $\boldsymbol{\Theta}_{\mathrm{T}} = [\boldsymbol{\theta}_{\mathrm{T},1}, \ldots, \boldsymbol{\theta}_{\mathrm{T},N}]$, the temporal transformation of signal $\boldsymbol{X}$ reads concisely as $\boldsymbol{X} = \boldsymbol{\Psi}_{\mathrm{T}}\boldsymbol{\Theta}_{\mathrm{T}}$, i.e.,

$$[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N] = \boldsymbol{\Psi}_{\mathrm{T}}[\boldsymbol{\theta}_{\mathrm{T},1}, \ldots, \boldsymbol{\theta}_{\mathrm{T},N}]. \tag{3}$$

Kronecker sparsifying bases can succinctly combine the different correlation patterns present in each signal dimension into a single matrix [17]. Thus, we can merge the transformations (2) and (3) as

$$\begin{aligned} \boldsymbol{x} = \mathrm{vec}(\boldsymbol{X}^{\mathrm{T}}) = \mathrm{vec}(\boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\Theta}_{\mathrm{S}}) &= \mathrm{vec}\Big(\boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\Theta}_{\mathrm{S}}\big(\boldsymbol{\Psi}_{\mathrm{T}}^{\mathrm{T}}\big)^{-1}\boldsymbol{\Psi}_{\mathrm{T}}^{\mathrm{T}}\Big) \\ &= \mathrm{vec}\big(\boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{Z}\boldsymbol{\Psi}_{\mathrm{T}}^{\mathrm{T}}\big) \\ &= (\boldsymbol{\Psi}_{\mathrm{T}}\otimes\boldsymbol{\Psi}_{\mathrm{S}})\mathrm{vec}(\boldsymbol{Z}) \\ &= \boldsymbol{\Psi z}, \end{aligned} \tag{4}$$

where $\boldsymbol{x} = [(\boldsymbol{x}^1)^{\mathrm{T}}, \ldots, (\boldsymbol{x}^T)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{TN}$, $\otimes$ denotes the Kronecker product, $\boldsymbol{\Psi} = (\boldsymbol{\Psi}_{\mathrm{T}}\otimes\boldsymbol{\Psi}_{\mathrm{S}}) \in \mathbb{R}^{TN\times TN}$ is the Kronecker sparsifying basis, and $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{Z}) \in \mathbb{R}^{TN}$ are the K_{J}-compressible coefficients for the joint spatio-temporal transformation.

At each time instance $t = 1, \ldots, T$, we acquire CS measurements $\boldsymbol{v}_t \in \mathbb{R}^{J_t}$, $J_t < N$, of WSN data samples $\boldsymbol{x}^t$ as

$$\boldsymbol{v}_t = \boldsymbol{\Omega}_t\boldsymbol{x}^t,\ t = 1, \ldots, T, \tag{5}$$

where $\boldsymbol{\Omega}_t \in \mathbb{R}^{J_t\times N}$ is the measurement matrix. By block-diagonal measurement matrix $\boldsymbol{\Omega} = \mathrm{diag}\{\boldsymbol{\Omega}_1, \ldots, \boldsymbol{\Omega}_T\} \in \mathbb{R}^{\bar{J}\times NT}$, the measurement process in (5) can be concisely expressed as

$$\boldsymbol{v} = \boldsymbol{\Omega x}, \tag{6}$$

where $\boldsymbol{v} = [\boldsymbol{v}_1^{\mathrm{T}}, \ldots, \boldsymbol{v}_T^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{\bar{J}}$ represents the measurements, and $\bar{J} = \sum_{t=1}^{T} J_t$ denotes the total number of involved measurements.

In order to have an applicable and energy efficient delivery of measurements (5) in the WSN, we will use a measurement process where only a subset of the sensors report their readings to the sink at each time instance [12, 13]. This induces a particular structure for $\boldsymbol{\Omega}_t \in \mathbb{R}^{J_t\times N}$, i.e., it has all the entries zeros except exactly single "1" at each row $j = 1, \ldots, J_t$ and at most single "1" at each column $i = 1, \ldots, N$ [12]. Thus, the projection performs sub-sampling of each $\boldsymbol{x}^t$, as only $J_t < N$ sensors are measured at each $t = 1, \ldots, T$.

Enabled by the compressibility in the spatial domain (2), a conventional CS decoding reconstructs each $\boldsymbol{x}^t$, $t = 1, \ldots, T$, separately from measurements (5) by solving

$$\hat{\boldsymbol{\theta}}_{\mathrm{S},t} := \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \ \text{s.t.}\ \boldsymbol{v}_t = \boldsymbol{\Omega}_t\boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\theta}, \tag{7}$$

resulting in $\hat{\boldsymbol{x}}^t = \boldsymbol{\Psi}_{\mathrm{S}}\hat{\boldsymbol{\theta}}_{\mathrm{S},t}$. However, by exploiting the joint correlation structure (4), the multi-dimensional signal $\boldsymbol{X}$ can be reconstructed from measurements (6) by solving a joint recovery problem

$$\hat{\boldsymbol{z}} := \arg\min_{\boldsymbol{z}} \|\boldsymbol{z}\|_1 \quad \text{s.t.} \quad \boldsymbol{v} = \boldsymbol{\Omega\Psi z}, \tag{8}$$

where we obtain $\hat{\boldsymbol{x}} = \boldsymbol{\Psi}\hat{\boldsymbol{z}} = [(\hat{\boldsymbol{x}}^1)^{\mathrm{T}}, \ldots, (\hat{\boldsymbol{x}}^T)^{\mathrm{T}}]^{\mathrm{T}}$, which then can be reshaped into $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_N] \in \mathbb{R}^{T\times N}$.

5. PROGRESSIVE JOINT SIGNAL RECOVERY

In the previous section, the presented CS acquisition procedure essentially follows the KCS framework [17], as the entire signal $\boldsymbol{X}$ is jointly reconstructed in (8) from measurements (6) collected over T time instances. Since the decoding delay and complexity of (8) is proportional to T, the approach imposes additional restrictions for an application. In order to balance between the decoding delay and complexity, we will derive a novel joint data acquisition and reconstruction method which relies on sliding window mechanism in the signal recovery. The sliding window processing allows not only to instantaneously reconstruct current WSN samples, but also allows to refine the accuracy of the estimates in successive decoding instances by utilizing the prior information available from multiple previous decoding instances.

The following basic notations will be used in the derivation. At time instance $t = 1, \ldots, T$, the data block inside the sliding window of size $W_t > 0$ is denoted as $\boldsymbol{X}_{\mathrm{W}}^t \in \mathbb{R}^{W_t \times N}$, referring to a portion of signal $\boldsymbol{X}$ given in (1). Namely, $\boldsymbol{X}_{\mathrm{W}}^t$ comprises of $\boldsymbol{X}_{\mathrm{W}}^t = [(\boldsymbol{X}_{\mathrm{B}}^t)^{\mathrm{T}}, \boldsymbol{x}^t]^{\mathrm{T}}$, where $\boldsymbol{x}^t$ are the current WSN samples and $\boldsymbol{X}_{\mathrm{B}}^t = [\boldsymbol{x}^{t-D_t}, \ldots, \boldsymbol{x}^{t-1}]^{\mathrm{T}} \in \mathbb{R}^{D_t \times N}$ are the samples of D_t previous time instances, $D_t = W_t - 1$.

Owing to the joint signal compressibility in (4), we can represent each data window $\boldsymbol{X}_{\mathrm{W}}^t$, $t = 1, \ldots, T$, as

$$\boldsymbol{x}_{\mathrm{W}}^t = \mathrm{vec}\big([\boldsymbol{X}_{\mathrm{W}}^t]^{\mathrm{T}}\big) = (\boldsymbol{\Psi}_{\mathrm{T_W}}^t \otimes \boldsymbol{\Psi}_{\mathrm{S}})\boldsymbol{z}_{\mathrm{W}}^t = \boldsymbol{\Psi}_{\mathrm{W}}^t \boldsymbol{z}_{\mathrm{W}}^t, \quad (9)$$

where $\boldsymbol{x}_{\mathrm{W}}^t = [(\boldsymbol{x}^{t-D_t})^{\mathrm{T}}, \ldots, (\boldsymbol{x}^t)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{NW_t}$, $\boldsymbol{\Psi}_{\mathrm{T_W}}^t \in \mathbb{R}^{W_t \times W_t}$ is the sparsifying basis for the temporal domain, $\boldsymbol{\Psi}_{\mathrm{W}}^t \in \mathbb{R}^{NW_t \times NW_t}$ is the Kronecker sparsifying basis and $\boldsymbol{z}_{\mathrm{W}}^t \in \mathbb{R}^{NW_t}$ denotes the K_{W}^t-compressible coefficients for the joint spatio-temporal domain.

According to (6), the CS measurements associated with $\boldsymbol{X}^t$, i.e., $\boldsymbol{v}_{\mathrm{W}}^t = [\boldsymbol{v}_{t-D_t}^{\mathrm{T}}, \ldots, \boldsymbol{v}_t^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{\bar{J}_t}$, correspond to a portion of $\boldsymbol{v}$ as

$$\boldsymbol{v}_{\mathrm{W}}^t = \boldsymbol{\Omega}_{\mathrm{W}}^t \boldsymbol{x}_{\mathrm{W}}^t, \; t = 1, \ldots, T \quad (10)$$

where $\boldsymbol{\Omega}_{\mathrm{W}}^t = \mathrm{diag}\{\boldsymbol{\Omega}_{t-D_t}, \ldots, \boldsymbol{\Omega}_t\} \in \mathbb{R}^{\bar{J}_t \times NW_t}$ is the block-diagonal measurement matrix and $\bar{J}_t = \sum_{\tau=t-D_t}^{t} J_\tau$ denotes the total number of involved measurements. Similarly to (8), each $\boldsymbol{X}_{\mathrm{W}}^t$, $t = 1, \ldots, T$, can be reconstructed from measurements (10) as

$$\hat{\boldsymbol{z}}_{\mathrm{W}}^t := \arg\min_{\boldsymbol{z}} \|\boldsymbol{z}\|_1 \quad \text{s.t.} \quad \boldsymbol{v}_{\mathrm{W}}^t = \boldsymbol{\Omega}_{\mathrm{W}}^t \boldsymbol{\Psi}_{\mathrm{W}}^t \boldsymbol{z}, \quad (11)$$

resulting in $\hat{\boldsymbol{x}}_{\mathrm{W}}^t = \boldsymbol{\Psi}_{\mathrm{W}}^t \hat{\boldsymbol{z}}_{\mathrm{W}}^t$, which then can be reshaped into matrix form of $\hat{\boldsymbol{X}}_{\mathrm{W}}^t = [\hat{\boldsymbol{x}}^{t-D_t}, \ldots, \hat{\boldsymbol{x}}^t]^{\mathrm{T}}$.

Depending on the chosen W_t, each $\boldsymbol{x}^t$ may be included multiple times in $\boldsymbol{X}_D^t$ associated with different time instances $t = 1, \ldots, T$. As a result, $\boldsymbol{x}^t$ will be reconstructed several times in (11). Therefore, the resulting signal estimates can be used as prior information in the subsequent joint signal recovery instances of (11). Next, we will propose a novel method where the decoder utilizes the prior information to improve the reconstruction accuracy for the signal ensemble $\boldsymbol{X}$.

For notational convenience, but without loss of generality, let us fix $W_t = W$, $t \geq W$, when $D_t = D$. Let us assume that the decoder has reconstructed estimates for the first W WSN data samples, e.g., via the associated CS recovery (11). Then, at time instance $t > W$, the already reconstructed estimates for D previous WSN data samples are stored in decoder buffer $\boldsymbol{b}^t \in \mathbb{R}^{DN}$ as

$$\boldsymbol{b}^t = \hat{\boldsymbol{x}}_{\mathrm{B},(t-1)}^t = \big[\big(\hat{\boldsymbol{x}}_{(t-1)}^{t-D}\big)^{\mathrm{T}}, \ldots, \big(\hat{\boldsymbol{x}}_{(t-1)}^{t-1}\big)^{\mathrm{T}}\big]^{\mathrm{T}} \quad (12)$$

where $\hat{\boldsymbol{x}}_{(t-1)}^{t-d} \in \mathbb{R}^N$ is the estimate of $\boldsymbol{x}^{t-d}$ obtained at the decoding instance $t-1$, $d = 1, \ldots, D$. Hence, the last N entries of $\boldsymbol{b}^t$ correspond to the estimates of the most recent WSN data, whereas the first N entries contain the estimates of the most out-dated ones.

As the recovery problem (11) aims of jointly reconstructing the data window $\boldsymbol{X}_{\mathrm{W}}^t = [(\boldsymbol{X}_{\mathrm{B}}^t)^{\mathrm{T}}, \boldsymbol{x}^t]^{\mathrm{T}}$, the problem can be reformulated to exploit the already achieved estimates $\hat{\boldsymbol{x}}_{\mathrm{B},(t-1)}^t$ stored in (12). Thus, instead of (11), the decoder solves at each $t > W$ a modified joint recovery problem to obtain $\hat{\boldsymbol{z}}_{\mathrm{W}}^t \in \mathbb{R}^{NW}$, i.e.,

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{z}\|_1 + \epsilon_{\mathrm{B}} \big\|\boldsymbol{\Psi}_{\mathrm{B}}\boldsymbol{z} - \boldsymbol{b}^t\big\|_2 \\ \text{subject to} & \boldsymbol{v}_{\mathrm{W}}^t = \boldsymbol{\Omega}_{\mathrm{W}}^t \boldsymbol{\Psi}_{\mathrm{W}} \boldsymbol{z}, \end{array} \quad (13)$$

where $\boldsymbol{z} = [z_1, \ldots, z_{NW}]^{\mathrm{T}}$ are the optimization variables, $\epsilon_{\mathrm{B}} \geq 0$ is a weighting parameter and matrix $\boldsymbol{\Psi}_{\mathrm{B}} \in \mathbb{R}^{ND \times NW}$ consists of the first ND rows of sparsifying basis $\boldsymbol{\Psi}_{\mathrm{W}}$ given in (9). As a result, we obtain $\hat{\boldsymbol{x}}_{\mathrm{W},(t)}^t = \boldsymbol{\Psi}_{\mathrm{W}} \hat{\boldsymbol{z}}_{\mathrm{W}}^t$, which can be reshaped into $\hat{\boldsymbol{X}}_{\mathrm{W},(t)}^t$.

Evidently, the regularization term is what differentiates problem (13) from (11). With the variables obtained in (13), the term reads as $\epsilon_{\mathrm{B}} \big\|\boldsymbol{\Psi}_{\mathrm{B}} \hat{\boldsymbol{z}}_{\mathrm{W}}^t - \boldsymbol{b}^t\big\|_2$, i.e., $\epsilon_{\mathrm{B}} \big\|\mathrm{vec}\big(\hat{\boldsymbol{X}}_{\mathrm{B},(t)}^t - \hat{\boldsymbol{X}}_{\mathrm{B},(t-1)}^t\big)\big\|_2$. Thus, the regularization induces additional penalty by the deviation between the estimates of $\boldsymbol{X}_{\mathrm{B}}^t$ obtained at consecutive decoding instances, with respect to L_2-norm. The emphasis between this deviation and the L_1-norm term, encouraging sparse or compressible variables $\boldsymbol{z}$, is controlled by parameter ϵ_{B}. Note that in general for $\epsilon_{\mathrm{B}} > 0$, problems (13) and (11) have different solutions for $t > W + 1$.

The joint compressed data acquisition and progressive signal recovery method is summarized in Algorithm 1. It is worth noting that the method can instantaneously reconstruct the WSN data samples, beneficial for delay-stringent applications. If certain amount of delay is allowed, these initial estimates can be progressively refined during the D successive decoding instances. As the outcome, the process keeps gradually rebuilding the final estimate based on the most recent estimates, corresponding to Step IV in Algorithm 1. Hence, by adjusting the sliding window size $W \in [1, T]$, the proposed method can not only regulate the decoding delay of refined estimates, but also allows to control the decoding complexity of problem (13).

6. NUMERICAL RESULTS

We considered WSNs with $N = 16$ and $N = 36$ sensors with the sensing period of $T = 60$ data samples. The sensors were deployed in an observation field of size $100\sqrt{N} \times 100\sqrt{N}$ units as follows: firstly, the field was divided into a $\sqrt{N} \times \sqrt{N}$-grid of square areas, and then, each of these square areas of size 100×100 units was randomly deployed one sensor according to the uniform distribution.

For $N = 16$ and $N = 36$, the monitoring fields comprised of $S = 2$ and $S = 4$ randomly located independent sources, respectively. The sources contributed additively to each sensor reading as $x_{ti} = \sum_{s \in \mathcal{S}} f_{is}\beta_{ts}$, where f_{is} is a time-invariant influence of source $s \in \mathcal{S}$ on sensor $i \in \mathcal{N}$ and β_{ts} is the source magnitude. The spatial dependency in each $\boldsymbol{x}^t$, $t = 1, \ldots, T$, was created by assuming power exponential correlation for the influence functions as

Algorithm 1 Compressed Data Acquisition & Progressive Signal Recovery Method

for $t = 1, \ldots, T$ **do**
- **I. CS Measurements**
  - Deliver CS measurements of $\boldsymbol{x}^t$ in (5) to the sink.
- **II. Progressive Joint Signal Recovery** – given W, ϵ_{B}
  - **if** $t = W$: Solve (11) $\rightarrow \hat{\boldsymbol{X}}_{\mathrm{W},(t)}^t = [\hat{\boldsymbol{x}}_{(t)}^1, \ldots, \hat{\boldsymbol{x}}_{(t)}^W]^{\mathrm{T}}$.
  - **if** $t > W$: Solve (13) $\rightarrow \hat{\boldsymbol{X}}_{\mathrm{W},(t)}^t = [\hat{\boldsymbol{x}}_{(t)}^{t-D}, \ldots, \hat{\boldsymbol{x}}_{(t)}^t]^{\mathrm{T}}$.
- **III. Decoder Buffer Update**
  - Set (12) as $\boldsymbol{b}^{t+1} = \hat{\boldsymbol{x}}_{\mathrm{B},(t)}^{t+1} = [(\hat{\boldsymbol{x}}_{(t)}^{t-D+1})^{\mathrm{T}}, \ldots, (\hat{\boldsymbol{x}}_{(t)}^t)^{\mathrm{T}}]^{\mathrm{T}}$.
- **IV. Signal Estimate Construction**
  - Set $\hat{\boldsymbol{X}}_{\mathrm{W}}^t \leftarrow \hat{\boldsymbol{X}}_{\mathrm{W},(t)}^t$.

end for

$f_{is} = \exp\{-(d_{is}/\alpha_1)^{\alpha_2}\}$ [22], where d_{is} is the distance between sensor $i \in \mathcal{N}$ and source $s \in \mathcal{S}$, and α_1 and α_2 are correlation parameters, which were set as $\alpha_1 = 1000$ and $\alpha_2 = 2$. The temporal correlation in each $\boldsymbol{x}_i$, $i \in \mathcal{N}$, was generated by the sequences of slowly varying source magnitudes $\{\beta_{1s}, \beta_{2s}, \ldots, \beta_{Ts}\}$, $s \in \mathcal{S}$. Fig. 2 depicts a source sequence for one simulated data field.

We will evaluate the performance of our proposed progressive CS signal recovery method (Prog-CS) summarized in Algorithm 1 in terms of CS reconstruction error, measured as $\|\hat{\bar{\boldsymbol{x}}} - \bar{\boldsymbol{x}}\|_2/\|\bar{\boldsymbol{x}}\|_2$. For benchmarking, we considered two CS methods involving contrary recovery procedures, i.e., 1) Kronecker CS (Kron-CS) which recovers $\boldsymbol{X}$ at once by (8) with measurements (6), incurring a decoding delay proportional to T, and 2) Spatial CS (Spat-CS) which instantaneously produces $\hat{\boldsymbol{x}}^t$ at each $t = 1, \ldots, T$ from (7) with measurements (5), yet completely ignoring the temporal domain compressibility. Since all three methods end up delivering the same measurements (6), given the WSN, their data transportation costs are equal. Thus, they can be compared by the number of measurements required to obtain certain CS recovery error.

Due to the spatial distribution of the sensors, we applied 2D transformation to sparsify the data in the spatial domain. Namely, each $\boldsymbol{x}^t$, $t = 1, \ldots, T$, was reorganized into matrix $\tilde{\boldsymbol{X}}^t \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$, whose (i_1, i_2)th entry refers to a reading of the sensor located at (i_1, i_2)th square of the $\sqrt{N} \times \sqrt{N}$-grid, $i_1, i_2 = 1, \ldots, \sqrt{N}$. By (4), $\tilde{\boldsymbol{X}}^t = \boldsymbol{\Psi}_{\mathrm{S}_1}\tilde{\boldsymbol{\Theta}}_{S,t}\boldsymbol{\Psi}_{\mathrm{S}_2}^{\mathrm{T}}$, and thus, $\boldsymbol{x}^t = \mathrm{vec}(\tilde{\boldsymbol{X}}^t) = \boldsymbol{\Psi}_{\mathrm{S}}\boldsymbol{\theta}_{S,t}$, where $\boldsymbol{\Psi}_{\mathrm{S}_1} \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$ and $\boldsymbol{\Psi}_{\mathrm{S}_2} \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$ are the bases for the first and second spatial dimension, respectively, $\boldsymbol{\Psi}_{\mathrm{S}} = (\boldsymbol{\Psi}_{\mathrm{S}_2} \otimes \boldsymbol{\Psi}_{\mathrm{S}_1})$ is the Kronecker sparsifying basis, and $\boldsymbol{\theta}_{S,t} \in \mathbb{R}^N$ are the transform domain coefficients.

Fig. 3 depicts the average CS recovery error versus different numbers of measurements for Prog-CS, Kron-CS and Spat-CS with $N = 16$ and $N = 36$. For each method, we set $J = J_t$, $\forall t = 1, \ldots, T$, and used the inverse of a DCT-matrix as each basis $\boldsymbol{\Psi}_{\mathrm{S}_1}$, $\boldsymbol{\Psi}_{\mathrm{S}_2}$ and $\boldsymbol{\Psi}_{\mathrm{T}}$. For $N = 16$, Prog-CS was run with sliding window sizes $W = [5, 15, 25]$ and fixed regularization parameters $\epsilon_B = [1, 3, 4]$, respectively. For $N = 36$, we used $W = [5, 15, 20]$ and $\epsilon_{\mathrm{B}} = [2, 3, 5]$, respectively. For $W = 5$, the figure also demonstrates Prog-CS with $\epsilon_{\mathrm{B}} = 0$, i.e., without memory for previous estimates. Additionally, to illustrate the effect of estimate refinement, we depict Prog-CS for $W = 5$ with respect to the first, instantaneously obtained estimates (Prog-$\mathrm{CS}_{\mathrm{first}}$), and the last estimates after decoded W times (Prog-$\mathrm{CS}_{\mathrm{last}}$), respectively. All the involved convex optimization problems were solved with l_1−MAGIC [23] or CVX [24] running in Matlab.

Fig. 3 shows that by exploiting joint spatio-temporal correlation in Prog-CS and Kron-CS, the number of required CS measurements can be significantly reduced as compared to Spat-CS, which ignores the temporal correlation. It can be observed that by incorporating the prior information on previous estimates in the decoding, the recovery error noticeably decreases even for Prog-$\mathrm{CS}_{\mathrm{first}}$. Obviously, the performance increase becomes more evident with respect to the refined estimates: for $N = 16$, the reconstruction accuracy of Prog-$\mathrm{CS}_{\mathrm{last}}$ with $W = 15$ matches that of Kron-CS, and by increasing the window to $W = 25$, Prog-$\mathrm{CS}_{\mathrm{last}}$ can even slightly outperform Kron-CS. Similarly for $N = 36$, Prog-$\mathrm{CS}_{\mathrm{last}}$ with $W = 20$ yielded as accurate estimates as Kron-CS. Hence, our proposed method was capable to achieve the same CS recovery performance as state of the art Kron-CS, yet by reducing the decoding delay and complexity with factor $W/T = 1/4$ and $1/3$, respectively.

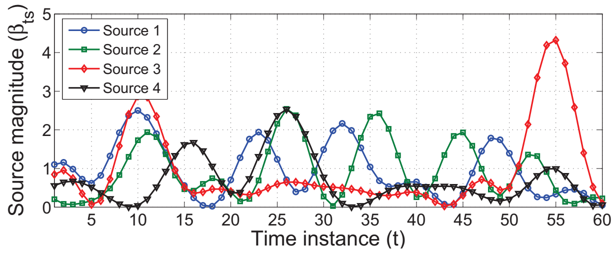
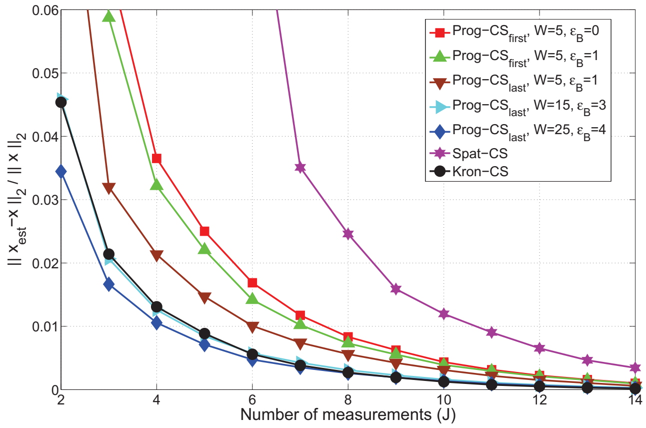
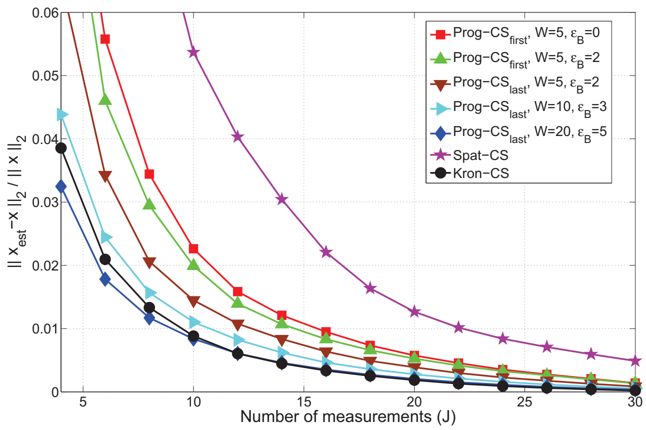


Fig. 2. An instance of source magnitude evolution with $S = 4$.



(a) $S = 2$ and $N = 16$



(b) $S = 4$ and $N = 36$

Fig. 3. CS recovery performance of Prog-CS, Kron-CS and Spat-CS.

7. CONCLUSIONS

We addressed a framework of joint compressed acquisition and progressive reconstruction of multi-dimensional correlated signals in multi-hop WSNs. We proposed a novel method based on sliding window mechanism, where the decoder can instantaneously recover WSN data samples from the collected CS measurements by exploiting the joint spatio-temporal data correlation via Kronecker sparsifying bases. Additionally, by using already reconstructed estimates as prior information, the joint decoding process can progressively improve the reconstruction accuracy of the signal ensemble. The numerical examples illustrated that the proposed method can result in the same CS recovery performance as the state of the art methods, yet with notably reduced decoding delay and complexity.

8. REFERENCES

[1] E.J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] E.J. Candes and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[3] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] J. Haupt and R. Nowak, “Signal reconstruction from noisy random projections,” *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4036–4048, Sept. 2006.

[5] R.G. Baraniuk, “Compressive sensing [lecture notes],” *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118 –121, July 2007.

[6] E.J Candes and M.B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[7] M.F. Duarte, M.B. Wakin, D. Baron, and R.G. Baraniuk, “Universal distributed sensing via random projections,” in *Proc. IEEE Int. Symp. on Inform. Proc. in Sensor Networks*, New York, NY, USA, 2006, pp. 177–185.

[8] J. Haupt, W.U. Bajwa, M. Rabbat, and R. Nowak, “Compressed sensing for networked data: A different approach to decentralized compression,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 92–101, Mar. 2008.

[9] C. Luo, F. Wu, J. Sun, and C.W. Chen, “Compressive data gathering for large-scale wireless sensor networks,” in *Proceedings of the 15th annual international conference on Mobile computing and networking*, Beijing, China, Sept. 20-25 2009, pp. 145–156.

[10] C. Luo, F. Wu, J. Sun, and C.W. Chen, “Efficient measurement generation and pervasive sparsity for compressive data gathering,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3728–3738, Dec. 2010.

[11] L. Xiang, J. Luo, and A. Vasilakos, “Compressed data aggregation for energy efficient wireless sensor networks,” in *Proc. IEEE Int. Conf. Sens., Commun. and Networking*, Salt Lake City, USA, June 27-30 2011, pp. 46–54.

[12] G. Quer, R. Masiero, G. Pillonetto, M. Rossi, and M. Zorzi, “Sensing, compression, and recovery for WSNs: Sparse signal modeling and monitoring framework,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3447–3461, Oct. 2012.

[13] M. Leinonen, M. Codreanu, and M. Juntti, “Distributed correlated data gathering in wireless sensor networks via compressed sensing,” in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, Pacific Grove, CA, Nov.3–6 2013.

[14] C. Luo, J. Sun, and F. Wu, “Compressive network coding for approximate sensor data gathering,” in *Proc. IEEE Global Telecommun. Conf.*, Houston, Texas USA, Dec. 5-9 2011, pp. 1–6.

[15] S. Feizi and M. Médard, “A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing,” *CoRR*, vol. abs/1110.0428, available at http://arxiv.org/abs/1110.0428.

[16] D. Baron, M.B. Wakin, M.F. Duarte, S. Sarvotham, and R.G. Baraniuk, “Distributed compressed sensing,” Tech. Rep., Nov. 2005.

[17] M.F. Duarte and R.G. Baraniuk, “Kronecker compressive sensing,” *IEEE Trans. Image Processing*, vol. 21, no. 2, pp. 494–504, 2012.

[18] N. Vaswani and L. Wei, “Modified-CS: Modifying compressive sensing for problems with partially known support,” *IEEE Trans. Signal Processing*, vol. 58, no. 9, pp. 4595–4607, Sept. 2010.

[19] W. Lu and N. Vaswani, “Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge,” *IEEE Trans. Signal Processing*, vol. 60, no. 1, pp. 182–196, Jan. 2012.

[20] M. S. Asif and J. Romberg, “Sparse recovery of streaming signals using l1-homotopy,” *IEEE Trans. Signal Processing, submitted*, 2013, available at http://arxiv.org/abs/1306.3331.

[21] E.J. Candes, “Compressive sampling,” in *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, Aug. 22–30 2006, vol. 3, pp. 1433–1452.

[22] J.O. Berger, V. De Oliveira, and B. Sanso, “Objective Bayesian analysis of spatially correlated data,” *Journal of the American Statistical Association*, vol. 96, pp. 1361–1374, Dec. 2000.

[23] E. Candes and J. Romberg, “l_1-MAGIC: Recovery of sparse signals via convex programming,” `http://users.ece.gatech.edu/~justin/l1magic/`, Oct. 2005.

[24] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 1.21,” `http://cvxr.com/cvx`, Apr. 2011.