JOINT SPARSE SIGNAL ENSEMBLE RECONSTRUCTION IN A WSN USING DECENTRALIZED BAYESIAN MATCHING PURSUIT

George Tzagkarakis[1], *Jean-Luc Starck*[1] *and Panagiotis Tsakalides*[2]

[1] CEA/DSM, SEDI-SAP, Service d'Astrophysique, Centre de Saclay, F-91191 Gif-Sur-Yvette cedex, France
[2] Institute of Computer Science (ICS) - Foundation for Research & Technology - Hellas (FORTH), Crete, Greece
Fax: + (33) 169086577, Email: {georgios.tzagkarakis, jstarck}@cea.fr, tsakalid@ics.forth.gr

ABSTRACT

Wireless networks comprised of low-cost sensory devices have been increasingly used in surveillance both at the civilian and military levels. Limited power, processing, and bandwidth resources is a major issue for abandoned sensors, which should be addressed to increase the network's performance and lifetime. In this work, the framework of compressive sensing is exploited, which allows accurate recovery of signals being sparse in some basis using only a small number of random incoherent projections. In particular, a recently introduced Bayesian Matching Pursuit method is modified in a decentralized way to reconstruct a multi-signal ensemble acquired by the nodes of a wireless sensor network, by exploiting a joint sparsity structure among the signals of the ensemble. The proposed approach requires a minimal amount of data transmissions among the sensors and a central node thus preserving the sensors' limited resources. At the same time, it achieves a reconstruction performance comparable to other distributed compressive sensing methods, which require the communication of a whole set of measurements to the central node.

1. INTRODUCTION

In many modern signal processing applications sampling is a key concept largely dominated by the long-term trend of the classical Nyquist's theorem. However, the high-resolution capabilities of modern sensing devices require high sampling rates resulting in ever increasing amounts of data, thus making compression a necessity prior to storage or transmission. Several seminal studies [1, 2] have shown that many natural signals result in a highly sparse representation when they are projected on localized orthonormal bases (*e.g.*, wavelets and sinusoids). Subsequently, compressing a sparse signal reduces in computing its transform coefficients resulting in a small number of large amplitude coefficients.

Compressive sensing (CS) provides a framework for simultaneous sensing and compression [3], enabling a potentially significant reduction in the sampling and computation costs at a sensor with limited capabilities. According to the theory of CS, a signal having a sparse transform representation can be reconstructed from a small set of incoherent random projections.

Let $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ be a matrix whose columns correspond to a transform basis. In terms of signal approximation it has been demonstrated [3, 4] that if a signal $\mathbf{f} \in \mathbb{R}^N$ is L-sparse in basis $\mathbf{\Psi}$ (meaning that the signal is exactly or approximately represented by L elements of this basis), then it can be reconstructed from $M = \mathcal{O}(L \log(N/L))$ non-adaptive linear projections onto a second measurement basis, which is incoherent with the sparsity basis. The general noisy measurement model in the sparsifying transform-domain is written as

$$\mathbf{g} = \mathbf{\Phi}\mathbf{\Psi}^T \mathbf{f} + \eta = \mathbf{\Phi}\mathbf{w} + \eta \; , \tag{1}$$

where $\mathbf{g} \in \mathbb{R}^M$ is the measurement vector, $\mathbf{\Phi} = [\phi_1, \ldots, \phi_M]^T \in \mathbb{R}^{M \times N}$ is the measurement matrix with $\phi_m \in \mathbb{R}^N$ being a random vector with independent and identically distributed (i.i.d.) components, $\mathbf{w}$ denotes the sparse vector of transform coefficients, and η is a noise term with bounded energy and unknown variance. Examples of measurement matrices, which are incoherent with any fixed transform basis with high probability (universality property [3]), are random matrices with i.i.d. Gaussian or Bernoulli entries. Besides, in the following we restrict ourselves to the case of additive white Gaussian noise (AWGN).

By employing the M compressive measurements and given the L-sparsity property in the transform basis, the original signal can be recovered by taking a number of different approaches. The majority of these solve constrained optimization problems, where commonly used techniques are based on convex relaxation [4] and greedy strategies (*e.g.*, Orthogonal Matching Pursuit (OMP) [5]). On the other hand, in recent studies [6, 7, 8] the problem of reconstructing a sparse signal from a highly reduced number of compressed measurements was treated in a Bayesian framework. In particular, given $\mathbf{g}$, and under a prior belief that $\mathbf{w}$ should be sparse, the objective is to formulate a *posterior probability distribution* for $\mathbf{w}$ and then seek for its maximum. This approach improved the accuracy over the previous norm-based CS techniques by providing not only a single signal estimate, but also the associated confidence intervals (error bars) in the approximation of $\mathbf{w}$ (and subsequently of $\mathbf{f}$), which are then used to refine the previous estimate with the goal of reducing the reconstruction uncertainty.

As the field of mobile computing and communication advances, so does the need of deploying a distributed, ad-hoc wireless sensor network (WSN) consisting of hundreds of sensors for spatiotemporal monitoring of a specific phenomenon in an area (sensor field) of interest. The fact that the sensor data are collected at distinct spatial locations necessitates the requirement for in-network communication and local information processing. This, in conjunction with the limited energy and bandwidth resources, makes the development of methods that extract relevant information about the sensor field, using as fewer measurements as possible, a challenging task.

The theory of *distributed compressive sensing* (DCS) [9] exploits intra- and inter-signal correlations. In a typical DCS setting, the sensors of a WSN measure signals that are each individually sparse in some basis. Each sensor independently compresses its signal by projecting it onto an incoherent basis and then transmits the associated information to a fusion center (FC), where under the right conditions one can reconstruct *jointly* all the signals.

In the present work, we develop a decentralized CS algorithm for reconstructing an ensemble of signals characterized by a joint sparsity structure, which are also corrupted by AWGN, in such a way that we achieve significant savings with respect to the amount of information that must be handled and communicated by each sensor. For this purpose, a recently introduced Bayesian Matching Pursuit algorithm [7] is modified and extended in a DCS framework by employing a suitable tree-structured scheme.

The rest of the paper is organized as follows: in Section 2, the joint sparsity structure of the signal ensemble is described and previous DCS algorithms are reviewed in brief. In Section 3, the proposed decentralized CS estimation of the sparse signal ensemble is analyzed in detail. In Section 4, the performance of the proposed

This work is supported by CS-ORION Marie Curie Industry-Academia Partnerships and Pathways (IAPP) project funded by the European Commission in FP7 (PIAP-GA-2009-251605)

method is compared with the performance of recent DCS methods, while in Section 5 conclusions are drawn along with directions for future work.

2. STATISTICAL SIGNAL ENSEMBLE MODEL UNDER A JOINT SPARSITY STRUCTURE

Due to the equivalence between the time-domain representation of a signal $\mathbf{f}$ and its transform-domain counterpart $\mathbf{w}$, without loss of generality we consider that the noisy CS measurements are acquired in the transform domain.

In this work, we employ the generalized notion of an ensemble of signals being *jointly sparse* in some basis. We consider the following joint sparsity model for a WSN consisting of J sensors:

$$\mathbf{g}_j = \mathbf{\Phi}_j \mathbf{\Psi}^T \mathbf{f}_j \overset{(\mathbf{w}_j = \mathbf{\Psi}^T \mathbf{f}_j)}{=} \mathbf{\Phi}_j \mathbf{w}_j + \eta_j \ , \ j \in \{1, \ldots, J\} \ , \quad (2)$$

where $\mathbf{\Phi}_j$ is the measurement matrix of the j-th sensor and η_j is the associated noise component. Assuming that $\mathbf{\Phi}_j$ is known, the quantities to be estimated, given $\mathbf{g}_j$, are the sparse weight vectors $\mathbf{w}_j$ and the noise variances $\sigma^2_{\eta_j}$.

The joint sparsity stands for the fact that all vectors $\{\mathbf{w}_j\}_{j=1}^J$ are L-sparse with their corresponding L non-zero coefficients placed at the same positions. That is, each $\mathbf{w}_j$ is supported on the same subset $\Upsilon \subset \{1, 2, \ldots, N\}$ with $card(\Upsilon) = L$ (where $card(\cdot)$ denotes the cardinality of a set). Also, in general, each matrix $\mathbf{\Phi}_j$ in (2) may be of different dimension $M_j \times N$ from sensor to sensor, where M_j is the number of CS measurements acquired by the j-th sensor. In the following, we consider for convenience that all sensors receive the same number of measurements $M_j = M$, $j = 1, \ldots, J$.

A real-world scenario well-described by the above model is when the sensors of a WSN acquire the same signal but with phase shifts and attenuations due to the signal propagation. In many cases it is critical to recover each one of the sensed signals $\mathbf{f}_j$, such as in acoustic localization and array processing.

Working in a Bayesian framework, the sparse vector $\mathbf{w}_j$, as well as the noise component η_j in (2) are random vectors. The use of a Gaussian mixture as a sparsity-enforcing prior for the unknown vector $\mathbf{w}$ is a common choice, due to its convenience in the subsequent statistical inference yielding closed form expressions and computationally efficient implementations. In the following derivations, the noise component is drawn from a multivariate Gaussian, $\eta_j \sim \mathcal{N}(\mathbf{0}, \sigma^2_\eta \mathbf{I})$ (we assume that all components have the same variance σ^2_η), while the sparsity of $\mathbf{w}_j$ is modelled by assuming that its components $\{\mathbf{w}^n_j\}_{n=1}^N$ are i.i.d. realizations of a Gaussian mixture, which is determined by a discrete random vector $\tau = [\tau_1, \ldots, \tau_N]^T$ of mixture parameters. For simplicity we assume that $\tau \sim$Bernoulli(λ_1), that is, $\Pr\{\tau_n = 1\} = \lambda_1$ and $\Pr\{\tau_n = 0\} = \lambda_0 = 1 - \lambda_1$, or equivalently $\lambda_{\tau_n} \in \{\lambda_0, \lambda_1\}$. Thus, we have the diagonal covariance matrices $\mathbf{\Sigma}_j(\tau_j) = diag(\sigma^2_{\tau_1}, \ldots, \sigma^2_{\tau_N})$, with $\sigma^2_{\tau_n} \neq 0$ or $\sigma^2_{\tau_n} = 0$ depending on whether the n-th component of $\mathbf{w}_j$ is significant and activated in the mixture or not, respectively. Similarly, the mean vectors μ_j depend on the same mixture vector τ_j and will be written as $\mu_j(\tau_j)$ to denote this dependence. In general, the components of each Gaussian mixture may be chosen from a set of Ω Gaussians: $\mu_{\tau_n} \in \{\mu_\omega\}_{\omega=1}^\Omega$, $\sigma^2_{\tau_n} \in \{\sigma^2_\omega\}_{\omega=1}^\Omega$. In the present study, we consider for simplicity that $\sigma^2_{\tau_n} \in \{\sigma^2_0, \sigma^2_1\}$, where $\sigma^2_0 = 0$ to enforce sparsity, and $\mu_{\tau_n} \in \{0, \mu_1\}$.

3. DECENTRALIZED FBMP FOR JOINT RECONSTRUCTION OF A SIGNAL ENSEMBLE

In this section, we introduce a decentralized extension of the recently introduced FBMP algorithm [7] for reconstructing an ensemble of sparse vectors $\{\mathbf{w}_j\}_{j=1}^J$ from the corresponding CS measurements $\{\mathbf{g}_j\}_{j=1}^J$ acquired by the J sensors of a WSN. This process is reduced in finding for each sensor j the sparse set of the most probable basis configurations (columns of $\mathbf{\Phi}_j$) associated with the activated mixture components. Then, their corresponding posterior probabilities are employed to obtain a Minimum Mean Squared Error (MMSE) estimate of the sparse vector $\mathbf{w}_j$.

In recent works [9, 10, 11, 12], the problem of reconstructing a signal ensemble with a joint sparsity structure was treated by introducing variants of greedy pursuit CS algorithms or by solving convex cone problems, which seek to identify the elements of the sparse support in an iterative way. The joint signal recovery is performed by solving an optimization problem of the form,

$$\min_{\mathbf{W}} \|\mathbf{W}\|_{1,q} \quad s.t. \quad \|\mathbf{\Phi W} - \mathbf{G}\|_F \leq \varepsilon \ , \quad (3)$$

where the mixed norm $\|\mathbf{W}\|_{1,q}$ is defined by the sum of the ℓ_q-norms of the rows of $\mathbf{W}$, $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_J]$ and $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_J]$ denote the overall matrices of the sparse and measurements vectors, respectively, $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ is a measurement matrix, which is common for all sensors, and $\|\mathbf{A}\|_F$ is the Frobenius norm of a matrix $\mathbf{A}$. In the subsequent evaluations, comparisons with the mixed $(1,2)$-norm will be performed by setting $q = 2$. Notice also that, although the subsequent derivations of the proposed algorithm consider the general case of a distinct measurement matrix $\mathbf{\Phi}_j$ for each sensor j, however for the comparison with previous methods in Section 4 we will consider the simplified case of a single measurement matrix $\mathbf{\Phi}$ for all sensors.

Returning to the proposed decentralized FBMP algorithm (hereafter called DCS-FBMP), a posterior probability is formed for a given mixture vector τ_j at each sensor j, which is obtained by applying Bayes' rule:

$$p(\tau_j|\mathbf{g}_j) = \frac{p(\mathbf{g}_j|\tau_j)p(\tau_j)}{\sum_{\tau \in \mathscr{T}} p(\mathbf{g}_j|\tau)p(\tau)} \ , \quad (4)$$

where $\mathscr{T} = \{0,1\}^N$ contains the 2^N possible basis configurations. In the ideal case, a single vector τ^* would be associated with the sparsest basis configuration that best describes the data (maximum posterior). However, in practice there may be additional configurations, that is, vectors τ, which also describe the data $\mathbf{g}_j$ yielding significant values of the posterior. Let $\mathscr{T}_{j,s}$ be the subset of $\mathscr{T}$ containing the vectors τ with the most *significant* posterior probabilities for the j-th sensor. We expect that the size of $\mathscr{T}_{j,s}$ will be much smaller than the size of $\mathscr{T}$, and thus the posterior probabilities $\{p(\tau_j|\mathbf{g}_j)\}_{\tau_j \in \mathscr{T}_{j,s}}$ can be estimated from $\{p(\mathbf{g}_j|\tau_j)p(\tau_j)\}_{\tau_j \in \mathscr{T}_{j,s}}$, since the denominator in (4) is fixed for the j-th sensor. By employing (2) under the Gaussian assumption for $\mathbf{w}_j$ and the Bernoulli prior for τ_j, the following mixture selection metric is defined as a maximum a posteriori (log-MAP) criterion to decide whether to include a given τ_j in $\mathscr{T}_{j,s}$, or not:

$$\rho(\tau_j, \mathbf{g}_j) = \ln[p(\mathbf{g}_j|\tau_j)p(\tau_j)] = -\frac{1}{2}\Big[\ln|\mathbf{R}_j(\tau_j)| + M \ln 2\pi + \big(\mathbf{g}_j - \mathbf{\Phi}_j \mu_j(\tau_j)\big)^T \mathbf{R}_j(\tau_j)^{-1} \big(\mathbf{g}_j - \mathbf{\Phi}_j \mu_j(\tau_j)\big)\Big] + \sum_{n=1}^N \ln \lambda_{\tau_n} \ , \quad (5)$$

with $$\mathbf{R}_j(\tau_j) = \mathbf{\Phi}_j \mathbf{\Sigma}_j(\tau_j) \mathbf{\Phi}_j^T + \sigma^2_\eta \mathbf{I}_{M \times M} \ , \ j = 1, \ldots, J \ . \quad (6)$$

3.1 MMSE estimate of $\mathbf{w}_j$

A computationally feasible approximation of the MMSE estimate of $\mathbf{w}_j$ using only the most significant posterior probabilities of the j-th sensor, that is, the discrete vectors τ_j corresponding to the largest values of the selection metric $\rho(\tau_j, \mathbf{g}_j)$, is given by:

$$\hat{\mathbf{w}}_{j,\mathrm{MMSE}} \triangleq \sum_{\tau \in \mathscr{T}_{j,s}} p(\tau|\mathbf{g}_j) \mathbb{E}\{\mathbf{w}_j|\mathbf{g}_j, \tau\} \ , \quad (7)$$

$$\mathbb{E}\{\mathbf{w}_j|\mathbf{g}_j, \tau\} = \mu_j(\tau) + \mathbf{\Sigma}_j(\tau) \mathbf{\Phi}_j^T \mathbf{R}_j(\tau)^{-1} (\mathbf{g}_j - \mathbf{\Phi}_j \mu_j(\tau)) \ . \quad (8)$$

3.2 Incremental basis selection via a tree-structure

At the core of the proposed DCS-FBMP approach is the use of an incremental tree-structured procedure, which is employed in several machine learning tasks for selecting the most significant basis configurations for the j-th sensor. We emphasize in advance that this process is only sub-optimal, but efficient enough motivating its use in a WSN scenario with the limitations mentioned before.

More specifically, a distinct tree is maintained and updated at each sensor j, with its root consisting of the zero vector $\tau_j^0 = \mathbf{0}$. Then, from the nodes of the l-th level of the j-th tree the set $\mathscr{T}_j^l$ is formed, which contains the binary vectors τ_j generated from the vectors of the corresponding parent nodes of the previous $(l-1)$-th level by "activating" the remaining $N-(l-1)$ zero mixture components one-by-one. Next, the values of the selection metric $\rho(\tau_j, \mathbf{g}_j)$ are computed for these new mixture vectors and those with the K largest values are stored in $\mathscr{T}_{j,s}^l$. The process is repeated until: i) $\mathscr{T}_{j,s}^{l_{j,max}}$ is estimated, where $l_{j,max}$ is the maximum number of tree levels for the j-th sensor chosen such that $\Pr\{\|\tau_j\|_0 > l_{j,max}\}$ is sufficiently small ($\|\tau_j\|_0$ is equal to the number of non-zero components of τ_j), or ii) the value of the selection metric $\rho(\tau_j, \mathbf{g}_j)$ for at least one mixture vector exceeds a predetermined threshold.

When moving from one level of the tree to the next the values of $\rho(\tau_j, \mathbf{g}_j)$ must be updated as a result from the activation of a single component at a time. For this purpose, let $\tau_j(q)$ denote the mixture vector which is identical to τ_j except for the q-th component, which is "activated" ($=1$) in $\tau_j(q)$ (for convenience we will write $\tau_{j,q}$), while it is "inactive" ($=0$) in τ_j. We are interested in computing the inter-level differences

$$\Delta_j^q(\tau_j) = \rho(\tau_{j,q}, \mathbf{g}_j) - \rho(\tau_j, \mathbf{g}_j) \ , \tag{9}$$

which are then used to decide which mixture components will be activated. More specifically, the set $\mathscr{T}_{j,s}^l$ at the l-th tree-level of the j-th sensor is formed by keeping the K binary vectors of $\mathscr{T}_j^l$ corresponding to the K largest values of $\Delta_j^q(\tau_j)$. Notice that, since $\rho(\tau_j, \mathbf{g}_j)$ is fixed from the previous iteration, the maximization of $\Delta_j^q(\tau_j)$ at the l-th level is exactly equivalent with the maximization of $\rho(\tau_{j,q}, \mathbf{g}_j)$. The key quantities to be updated are the inverse of $\mathbf{R}_j(\tau_j)$ and its determinant when the q-th component is activated. These updates are obtained by applying the matrix inversion lemma and the determinant identity resulting in the following simple incremental expressions:

$$\mathbf{R}_j(\tau_{j,q}) = \mathbf{R}_j(\tau_j) + \sigma_{\tau_q}^2 \phi_{j,q}\phi_{j,q}^T \ , \ \ |\mathbf{R}_j(\tau_{j,q})| = \frac{\sigma_{\tau_q}^2}{\gamma_{j,q}} |\mathbf{R}_j(\tau_j)| \ ,$$

$$\mathbf{R}_j(\tau_{j,q})^{-1} = \mathbf{R}_j(\tau_j)^{-1} - \gamma_{j,q}\mathbf{v}_{j,q}\mathbf{v}_{j,q}^T \ ,$$

where $\gamma_{j,q} = \sigma_{\tau_q}^2(1+\sigma_{\tau_q}^2\phi_{j,q}^T\mathbf{v}_{j,q})^{-1}$ and $\mathbf{v}_{j,q} = \mathbf{R}_j(\tau_j)^{-1}\phi_{j,q}$, with $\phi_{j,q}$ denoting the q-th row of $\mathbf{\Phi}_j$. Notice also that the updated mean vector $\mu_j(\tau_{j,q})$ is the same as $\mu_j(\tau_j)$ except for a change of its q-th component from $\mu_q = \mu_0 = 0$ to $\mu_q = \mu_1$.

Regarding the algorithm's termination, we use the second from the two rules defined before, that is, if at least one of the K selection metric values $\rho(\tau_j, \mathbf{g}_j)$ at the $l_{j,m}$-level search ($l_{j,m} \leq l_{j,max}$) exceeds a predetermined threshold ρ_{Th}, the algorithm returns the sub-optimal set $\mathscr{T}_{j,s}^{l_{j,m}}$ with the most significant mixture vectors, along with the associated values of the posterior distribution obtained by taking the exponent of (5). If this criterion is not satisfied, a new search starts and proceeds by ignoring all the previously explored nodes of the tree. The algorithm continues until the thresholding criterion is satisfied, or until a maximum number of searches, S_{max}, is reached.

3.3 WSN topology for joint reconstruction

The implementation of a typical DCS method requires all the J sensors of the WSN to transmit their corresponding CS measurements $\{\mathbf{g}_j\}_{j=1}^J$ to a central node (FC), where the joint reconstruction of $\mathbf{w}_j$, $j=1,\ldots,J$, takes place. In this case, the cost that dominates the whole process is the communication cost spent by each sensor to transmit its CS measurements to the FC. Besides, it is usually assumed that the FC is placed far enough from the sensor field, such that the transmission energy spent by each sensor is approximately equal for all sensors.

Despite the already reduced number of CS measurements, when compared with the original signal dimensionality, the proposed DCS-FBMP scheme achieves a further reduction of the amount of information to be communicated by the sensors of the WSN. Let $\mathscr{A}_j^l$ denote the set of indices of the "active" (non-zero) components of a mixture vector corresponding to the l-th level of the tree structure of sensor j. That is, $\mathscr{A}_j^l = \{n \in \{1,2,\ldots,N\} | \tau_j^l(n) \neq 0\}$, $l=1,\ldots,l_{j,m}$ ($l_{j,m} \leq l_{j,max}$), $j=1,\ldots,J$, where τ_j^l is the mixture vector obtained at the l-th level. As it was described above, when moving from one level of the tree to the next, τ_j^l is updated by activating a single mixture component. Consequently, if we assume that the q-th component is activated, the set of "active" indices $\mathscr{A}_j^l$ is updated as: $\mathscr{A}_j^{l+1} = \mathscr{A}_j^l \cup \{q\}$, $q \in \{1,2,\ldots,N\} \setminus \mathscr{A}_j^l$. Due to the joint sparsity structure described in Section 2, in the ideal case the decentralized scheme should result in a single set $\mathscr{A}$ for all sensors in the WSN, that is, $\mathscr{A}_j^{l_{j,m}} \equiv \mathscr{A}$, $j=1,\ldots,J$.

Thus, the proposed DCS-FBMP algorithm is carried out in two alternating phases:

- **Sparse support estimate:** the first phase consists of estimating the joint sparse support by exchanging a minimal amount of information, where only the current optimal index q_j needs to be transmitted by each sensor.
- **Amplitude estimate:** in the second phase, having fixed the support obtained during the first phase, each sensor estimates individually the corresponding sparse vector $\mathbf{w}_j$.

The "message passing" during the first phase can be implemented by either a star-shaped topology, where all sensors communicate only with a FC, or a ring-shaped topology, where no FC is required and the sensor j passes its estimate for q_j to the next sensor. The disadvantage of the ring-shaped topology is the increased estimation time, since a complete passing from all sensors is necessary to decide whether a mixture component should be activated or not. On the other hand, in a star-shaped topology the process evolves in parallel. However, *synchronization* issues arise in this case, like for instance, how can we resolve the problem of a different number of tree levels from sensor to sensor, or the difference in estimation times, since a single slow sensor could affect the whole network. These rather network-oriented topics will be the subject of a separate study.

In the present implementation of DCS-FBMP we adopt the star-shaped topology by assuming the ideal case of perfect synchronization among the sensors. Thus, at the $(l+1)$-th iteration, the j-th sensor estimates the optimal activated index q_j that updates its corresponding mixture vector, $\tau_j^{(l+1)}(q_j)$. Then, each sensor transmits only its index to the FC, where the optimal index q_{opt}^{l+1} for the $(l+1)$-th iteration (tree level) is the index that appears with the highest frequency in the set $\{q_j\}_{j=1}^J$. Afterwards, the FC broadcasts to the WSN the value of q_{opt}^{l+1}, and those sensors for which $q_j \neq q_{opt}^{l+1}$ re-estimate their corresponding set $\mathscr{T}_{j,s}^{l+1}$ using q_{opt}^{l+1} instead of q_j, and accordingly the MMSE estimate of the sparse vector $\mathbf{w}_j$ (using (6)-(8)).

4. EXPERIMENTAL RESULTS

In this section, we compare the reconstruction performance of the proposed DCS-FBMP scheme with the performance obtained by reconstructing the ensemble jointly via the solution of (3) using SPGL [11] (J-SPGL) and TFOCS [12] (J-TFOCS), where each sensor transmits the whole CS measurement vector $\mathbf{g}_j$ to the FC for joint reconstruction of $\mathbf{w}_j$, $j = 1, \ldots, J$.

The experimental setup consists of $J \in \{10 : 5 : 30\}$ sensors distributed uniformly in the square $[0,1] \times [0,1]$. For each sensor a measurement vector $\mathbf{g}_j$ is generated according to (2), where the sparse vector $\mathbf{w}_j$ is drawn from a multivariate Gaussian distribution with $N = 512$ components and L non-zero spikes, whose locations are chosen at random (but then they are fixed for all sensors). We set the sparsity level to be equal to $L = \lceil \lambda_1 \cdot N \rceil$, with $\lambda_1 = 0.05$. The measurement noise is drawn from a zero-mean multivariate Gaussian with variance σ_η^2. In addition, the entries of the measurement matrix $\mathbf{\Phi}$, which is the same for all sensors, are i.i.d. Gaussian samples and with its columns being normalized to unit ℓ_2 norm. The values of the mixing parameters μ_1 and σ_1^2 are set equal to 0.5 and 1.5, respectively, while the maximum number of tree searches is set to $S_{max} = 10$. As mentioned in Section 3.2, the thresholding criterion is used to terminate the tree-searches, where the value of the threshold ρ_{Th} is set equal to the mean of the mixture selection metric, $\mathbb{E}\{\rho(\tau, \mathbf{g})\}$, which is given by the following expression after some algebraic manipulation of (5),

$$\rho_{Th} = -\frac{M}{2}\Big(1 + \ln(2\pi\sigma_\eta^2)\Big) - \lambda_1 \frac{N}{2} \ln\Big(\frac{\sigma_1^2}{\sigma_\eta^2} + 1\Big) + N\Big(\ln\lambda_0 + \lambda_1 \ln\frac{\lambda_1}{\lambda_0}\Big)\,. \tag{10}$$

The reconstruction performance is evaluated in terms of the reconstruction error, as well as the accuracy in estimating the true sparse support. The normalized mean-squared error of the MMSE estimated sparse vectors, $\hat{\mathbf{w}}_j$, is employed as a measure of the reconstruction quality, defined by

$$\text{NMSE}_{\text{MMSE}} = \frac{1}{J} \sum_{j=1}^{J} \frac{\|\hat{\mathbf{w}}_{\text{j,MMSE}} - \mathbf{w}_j\|_2^2}{\|\mathbf{w}_j\|_2^2}\,,$$

where $\hat{\mathbf{w}}_{\text{j,MMSE}}$ is the MMSE estimate of $\mathbf{w}_j$, given by (7) using M CS measurements. All of the following results correspond to the average performance over 100 Monte-Carlo runs.

We start by testing the effect of the number of sensors (J) and CS measurements (M) on the reconstruction performance. For this purpose, the value of M varies as a portion of the original signal dimension, $M = rN$, with $r \in [0.10, 0.50]$, $N = 512$. Fig. 1 shows the average MMSE for the proposed algorithm along with the other two methods, as a function of J and M. As it can be seen, DCS-FBMP presents an increased robustness with respect to a reduced amount of measurements resulting in an improved performance compared to the J-SPGL and J-TFOCS-based reconstruction methods. The simple "message passing" process for the recovery of the sparse support during the first phase of DCS-FBMP maintains an better reconstruction quality for the whole range of J. This means that DCS-FBMP could stand up efficiently against potential node failures that may appear in a WSN decreasing the number of sensors.

On the other hand, Fig. 2 shows the average percentage of exact recovery of the true sparse support as a function of J and M. By construction, the solution of DCS-FBMP is exactly sparse and thus an exact reconstruction is achieved if the supports of the estimated and the true vectors coincide. This is not the case for the J-SPGL and J-TFOCS-based solutions, for which an exact reconstruction is defined if the support of the L largest-amplitude components of the reconstructed vector coincides with the original support. As before, the proposed DCS-FBMP method is able to recover perfectly the original support, without being affected by a decreased number of CS measurements or sensors. On the contrary, the capability of J-SPGL and J-TFOCS in recovering the true support decreases as the number of CS measurements decreases.

As a second evaluation, the reconstruction performance of the three methods is compared in the case of noisy measurements. For this purpose, each measurement vector $\mathbf{g}_j$, $j = 1, \ldots, J$, is corrupted by AWGN with the input signal-to-noise ratio (SNR) varying from 10 to 30 dB. The noise variance is determined by solving the equation $\text{SNR} = \frac{\sigma_1^2 \lambda_1 N}{\sigma_\eta^2 M}$, with respect to σ_η^2. The following results correspond to $M = 0.25 \cdot N$ noisy CS measurements.

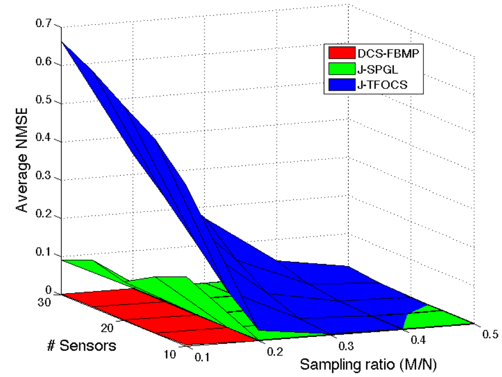

Figure 1: Average MMSE for DCS-FBMP, J-SPGL, and J-TFOCS as a function of J, M ($L = 26$).

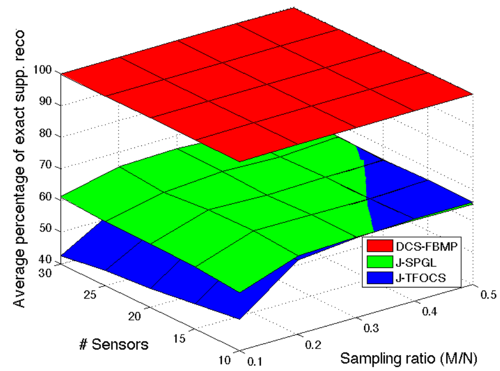

Figure 2: Average percentage of exact support recovery for DCS-FBMP, J-SPGL, and J-TFOCS as a function of J, M ($L = 26$).

Fig. 3 shows the average MMSE for DCS-FBMP, J-SPGL and J-TFOCS as a function of the input SNR and the number of sensors. We observe that the proposed method results again in a superior reconstruction quality for the whole range of (SNR, J) pairs, in contrast to J-SPGL and J-TFOCS, which are affected by low SNR values in conjunction with an increased number of sensors (this is more prominent for J-TFOCS).

Fig. 4 presents the average percentage of exact recovery of the true sparse support as a function of SNR and J, where the exact recovery is defined as in the noiseless case. As before, DCS-FBMP achieves again a perfect recovery of the original support irrespectively from the number of sensors or the noise power. Regarding the other two methods, the estimation accuracy decreases with a decrease of the SNR, with both J-SPGL and J-TFOCS achieving almost the same performance.

As a last experiment, we evaluate the average reconstruction performance as a function of the degree of sparsity (L), expressed as a percentage of the original signal dimensionality, and the number of sensors (J). Fig. 5 shows that the DCS-FBMP approach outperforms J-SPGL and J-TFOCS in terms of the reconstruction quality, while Fig. 6 reveals a significant improvement of DCS-FBMP in recovering the true sparse support, in contrast to the other two methods, whose discriminative capability decreases as the number of sensors and the sparsity level increase.

5. CONCLUSIONS AND FUTURE WORK

In this work, we introduced a decentralized extension of a recent FBMP method for reconstructing a signal ensemble with a joint sparsity structure at the nodes of a WSN requiring a mini-

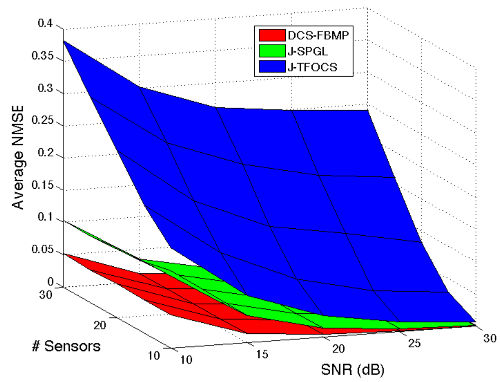

Figure 3: Average MMSE in the noisy case for DCS-FBMP, J-SPGL, and J-TFOCS as a function of SNR, J ($L = 26$).

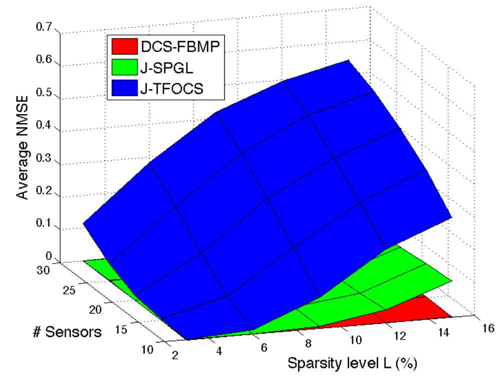

Figure 5: Average MMSE for DCS-FBMP, J-SPGL, and J-TFOCS as a function of L, J ($M = 0.25 \cdot N$).

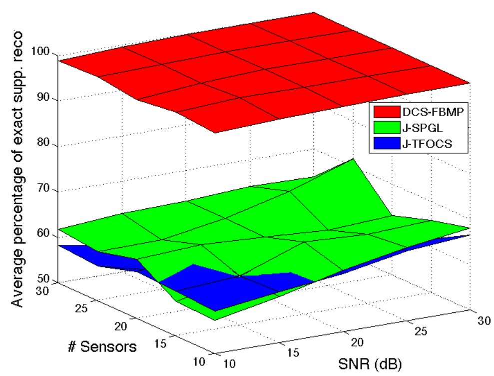

Figure 4: Average percentage of exact support recovery in the noisy case for DCS-FBMP, J-SPGL, and J-TFOCS as a function of SNR, J ($L = 26$).

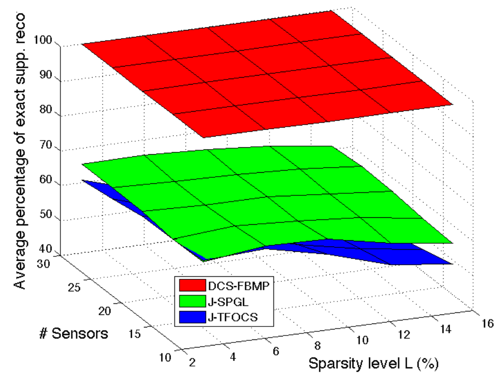

Figure 6: Average percentage of exact support recovery for DCS-FBMP, J-SPGL, and J-TFOCS as a function of L, J ($M = 0.25 \cdot N$).

mal amount of transmitted information. The experimental results revealed a superior performance compared with previous DCS algorithms, where the reconstruction is based on the transmission of the full set of CS measurements to a FC. More importantly, for a WSN scenario, the proposed DCS-FBMP approach is robust to a reduction in the number of CS measurements or to node failures.

In the present work, we assumed that the components of each mixture vector τ_j are chosen from two distributions ("inactive", "active"). Besides, the parameters of these distributions are predetermined and kept fixed during the reconstruction process. As a future work, we are interested in modifying the proposed model so as to employ a larger set of candidate mixture distributions expecting that it will increase the approximation accuracy of the underlying Gaussian mixture model. Moreover, the star-shaped WSN topology used here requires appropriate synchronization among the nodes. Such issues, along with the study of hierarchical schemes based on inter-sensor communications in small local clusters are also of importance in a WSN scenario and should be treated in a thorough separate study.

REFERENCES

[1] S. Mallat, "A Wavelet Tour of Signal Processing", 2nd ed., New York: Academic Press, 1998.

[2] I. Daubechies, "Ten lectures on wavelets", SIAM, 1992.

[3] D. Donoho, "Compressed sensing", *IEEE Trans. Inf. Th.*, Vol. 52 (4), pp. 1289–1306, Apr. 2006.

[4] E. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Inf. Th.*, Vol. 52, pp. 489–509, Feb. 2006.

[5] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Inf. Th.*, Vol. 53, pp. 4655–466, Dec. 2007.

[6] S. Ji, Ya Xue, and L. Carin, "Bayesian compressive sensing", *IEEE Trans. Sig. Proc.*, Vol. 56 (6), pp. 2346–2356, June 2008.

[7] P. Schniter *et al.*, "Fast Bayesian Matching Pursuit", *Proc. Workshop on Inf. Th. & Appl.*, La Jolla, CA, Jan. 2008.

[8] S. Babacan *et al.*, "Fast Bayesian compressive sensing using Laplace priors," *Proc. IEEE Int. Conf. on Acoust., Speech & Sig. Proc. (ICASSP'09)*, Taipei, Taiwan, April 2009.

[9] D. Baron *et al.*, "Distributed compressed sensing", [On-line: `www.ee.technion.ac.il/people/drorb/pdf/DCS012009.pdf`].

[10] M. Duarte *et al.*, "Distributed compressed sensing of jointly sparse signals", *in Proc. 39th Asilomar Conf. Signals, Sys., Comput.*, Nov. 2005.

[11] E. van den Berg and M. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, [Matlab code: `http://www.cs.ubc.ca/labs/scl/spgl1`].

[12] S. Becker, E. Candès, and M. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Tech. Report*, Stanford Univ., Sep. 2010 [Matlab code: `http://tfocs.stanford.edu/code`].