A UNIVERSAL MATCHED SOURCE-CHANNEL COMMUNICATION SCHEME FOR WIRELESS SENSOR ENSEMBLES

Waheed Bajwa, Jarvis Haupt, Akbar Sayeed, and Robert Nowak

University of Wisconsin-Madison
Department of Electrical and Computer Engineering
{bajwa, jhaupt}@cae.wisc.edu, {akbar, nowak}@engr.wisc.edu

ABSTRACT

The *essential task* in nearly all applications of sensor networks is to extract relevant information about the sensed data and deliver it to a desired destination. The *overall goal* in the design of sensor networks is to execute this task with least consumption of network resources. In this regard, the relevant metrics of interest are 1) the latency (bandwidth) involved in network data acquisition; and 2) the energy-distortion (E-D) tradeoff: given some desired distortion level D, how much energy E does the sensor network consume in extracting and delivering relevant information up to distortion D at a (usually) distant destination. It is generally recognized that given sufficient prior knowledge about the sensed data, there exist distributed processing and communication schemes that have a very favorable E-D tradeoff in the sense that $D \searrow 0$ as $n \to \infty$ while E grows at most sub-linearly with the number of nodes (n) in the network. However, it is not known whether such schemes exist when little or no prior knowledge about the sensed data is available. In this paper, we present a distributed matched-source channel communication scheme that naturally integrates the operations of processing and communications in a sensor network and is universal in the sense that it provides us with a consistent estimation scheme such that E grows sub-linearly with n even when little prior knowledge about the sensed data is assumed. This universality, however, comes at the price of increased latency (bandwidth) and a less favorable E-D tradeoff and we quantify this price by comparing our scheme to the case when sufficient prior information about the sensed data is available.

1. INTRODUCTION

Sensor networking is an emerging technology that promises an unprecedented ability to monitor and manipulate the physical world via a spatially distributed network of small and inexpensive wireless sensor nodes that have the ability to self-organize into a well-connected network. The *essential task* in nearly all applications of sensor networks is to extract relevant information about the sensed data and deliver it to a desired destination. The *overall goal* in the design of sensor networks is to execute this task with least consumption of network resources. Consequently, a major challenge in sensor networking applications is the development of efficient distributed methods for processing and communication of information from within the network to a given destination. In this regard, the relevant metrics of interest are 1) the latency (bandwidth) involved in network data acquisition; and 2) the energy-distortion (E-D) tradeoff: given some desired distortion level D, how much energy E does the sensor network consume in extracting and delivering relevant information up to distortion D at a (usually) distant destination.

It is generally recognized that given sufficient prior knowledge about the sensed data (e.g., statistical/topological characterization of the sensor network data, homogeneity of the sensor network data etc.), there exist distributed processing and communication schemes that have a very favorable E-D tradeoff in the sense that $D \searrow 0$ as $n \to \infty$ while E grows at most sub-linearly with the number of nodes (n) in the network (see, e.g., [1, 2, 3, 4]). However, it is not known whether such schemes *always* exist when little or no prior knowledge about the sensed data is available (see, e.g., Section 2).

In this paper, we propose a distributed matched-source channel communication scheme, based in part on recent results in wireless communications [1, 2, 5] and compressive sampling theory [6, 7, 8], that is universal in the sense that it provides us with a consistent estimation scheme such that E grows sub-linearly with n without requiring any prior knowledge about the sensed data. Moreover, this scheme naturally integrates the operations of processing and communications, thereby reducing the amount of processing and communications required inside the network and provides us with a system that often acts less like networks and more like coherent ensembles of sensors, thereby reducing the overhead of network-centric functions such as routing etc. The added flexibility and universality of the proposed scheme, however, comes at the price of increased latency (bandwidth) and a less favorable E-D tradeoff and we quantify this price by comparing our scheme to the case when sufficient prior information about the sensed data is available.

1.1. Problem Formulation

In this section, we formally define the problem considered in the paper. In the following sections, we shall elaborate on the technical details of the proposed scheme. To begin, consider a wireless sensor network with n nodes where each node takes a noisy sample of the form

$$x_j \;=\; x_j^* \,+\, w_j\,, \quad j = 1, \ldots, n \tag{1}$$

and w_j is assumed to be zero mean, independent and identically distributed (i.i.d) Gaussian measurement noise (in space and time) with variance σ_w^2. We can consider this data as a vector $x \in \mathbb{R}^n$ such that $x = x^* + w$, where $x^* \in \mathbb{R}^n$ is the noiseless data vector and $w \sim \mathcal{N}\left(0, \sigma_w^2 \mathbf{I}_n\right)$. We further assume that $|x_j^*| \leq B, j = 1, \ldots, n,$ for some known constant $B > 0$, which is determined by the sensing range of the sensors.

Given x, the goal of the sensor network is to compute a reconstruction $\widehat{x}$ of the noiseless data vector x^* at a distant destination and the reconstruction to have a small latency, L = number

This research was supported in part by the NSF under grants CCR-0310889 and CNS-0519824.

of network-to-destination channel uses, and expected squared error, $D = \mathbb{E}\left[\frac{1}{n}\left\|\widehat{x} - x^*\right\|^2\right]$, while at the same time consuming minimal amount of energy E.

Assumptions: In order to facilitate our analysis, we shall make the following assumptions:

A1: Each sensor is equipped with a single isotropic antenna.

A2: Let d_j, $j = 1, \dots, n$, be the distance between the sensor at location j and the destination. The destination is assumed to be far away from the sensor network so that $d_1 \approx \cdots \approx d_n \approx d$ and therefore, the path losses of all sensors are identical.

A3: The sensors communicate with the destination over a narrowband Additive White Gaussian Noise (AWGN) wireless channel of bandwidth W Hz at some carrier frequency f_c, where $f_c \gg W$, and each channel use is characterized by (real) transmission over a period of $T = 1/2W$ seconds.

A4: x^* lies in an m-dimensional subspace of $\mathbb{R}^n$, where $m \ll n$. That is, let $\Psi \triangleq \{\psi_i\}_{i=1}^n$ be an orthonormal basis of $\mathbb{R}^n$ and denote by $\theta_i = \psi_i^T x^*$ the coefficients of x^* in this new basis. Then, $x^* = \sum_{i=1}^m \theta_i \, \psi_i$ (perhaps after re-labeling the indices i). $\diamondsuit$

While **A1**-**A3** are quite self-explanatory and in line with the real-world scenarios, **A4** requires a few words of explanation. Indeed, in most real-world scenarios, we do not expect x^* to be sparse in any basis of $\mathbb{R}^n$. However, it is well known that data collected at nearby nodes in a dense sensor network is expected to be highly correlated [9] and thus, shall admit a *nearly* sparse representation in a compressing basis. The fact that many real-life signals are compressible is evidenced by the success of familiar compression standards such as JPEG, MPEG and MP3. Therefore, it is quite reasonable to assume x^* to be *compressible* in some basis of $\mathbb{R}^n$. However, to motivate the proposed scheme we shall restrict ourselves in this paper to signals that are completely sparse in some basis ($m \ll n$ non-zero coefficients in some transform domain) and shall make the transition from sparse to compressible signals in a future contribution.

1.2. Distributed Projections of Sensor Network Data

In this section, we develop the basic communication structure of our proposed scheme. At the heart of our approach is an efficient (distributed) method of estimating projections of the noiseless sensor network data onto any normalized vector in $\mathbb{R}^n$ by using only a fixed amount of energy (independent of n). However, before describing this procedure in detail, we shall define the notion of a *Sparsity Map*.

Definition 1 (The Sparsity Map): Let $q \in \mathbb{R}^n$ and $S_p : \mathbb{R}^n \to \mathrm{P}(\{1, \dots, n\})$, where $\mathrm{P}(X)$ means power set of X. We call S_p the sparsity map of q if $S_p(q) = \{j \in \{1, \dots, n\} : q_j \neq 0\}$ and $|S_p(q)|$ is a counting measure on $S_p(q)$. $\diamondsuit$

Now, let $\varphi \in \mathbb{R}^n$, where $\left\|\varphi\right\|^2 = 1$, and $\upsilon = \sum_{j=1}^n \varphi_j x_j^*$ be the projection of x^* onto φ. Using the notion of sparsity map, let us denote $|S_p(\varphi)| = n_\varphi$. Since $\left\|\varphi\right\|^2 = 1$, we have $|\varphi_j|^2 \approx \left\|\varphi\right\|^2/n_\varphi = 1/n_\varphi \; \forall \, j \in S_p(\varphi)$. Then, given any $E_o > 0$ and x as in (1), the destination can compute an estimate ($\widehat{\upsilon}$) of υ in $E \sim E_o$ amount of energy, such that $\mathbb{E}\left[|\widehat{\upsilon} - \upsilon|^2\right] = \sigma_w^2 + \frac{\sigma_z^2}{E_o}$, by making the sensor network sequentially perform the following steps:

S1: The destination transmits φ_j to the sensor at location j, where $j = 1, \dots, n$. Given the nature of the problem, we can assume the downlink (from the destination to the sensor network) to be error free. Thus, each sensor receives φ_j in an error free manner[1].

S2: The sensor at location j multiplies its measurement x_j with $\left(\sqrt{E_o}\,\varphi_j\right)$ to obtain $y_j = \sqrt{E_o}\,\varphi_j x_j$. Moreover, $\mathbb{E}\left[|y_j|^2\right] \leq \frac{E_o\left(B^2 + \sigma_w^2\right)}{n_\varphi} \preceq \frac{E_o}{n_\varphi}$ if $j \in S_p(x^*) \cap S_p(\varphi)$; $\mathbb{E}\left[|y_j|^2\right] = \frac{E_o\,\sigma_w^2}{n_\varphi}$ if $j \in S_p(x^*)^c \cap S_p(\varphi)$; and $\mathbb{E}\left[|y_j|^2\right] = 0$ if $j \notin S_p(\varphi)$. Thus, $\mathbb{E}\left[|y_j|^2\right] \preceq \frac{E_o}{n_\varphi} \; \forall \, j \in S_p(\varphi)$.

S3: All the sensors coherently transmit their corresponding y_j in an analog fashion over the network-to-destination AWGN channel, effectively transforming it into an AWGN MAC channel, and the received signal at the destination is given by[2]

$$r \;=\; \sum_{j=1}^{n} y_j + z \;=\; \sqrt{E_o}\left(\sum_{j=1}^{n} \varphi_j x_j\right) \;+\; z \qquad (2)$$

$$\;=\; \sqrt{E_o}\,(\upsilon + \widetilde{w}) \;+\; z \qquad (3)$$

where $z \sim \mathcal{N}(0, \sigma_z^2)$ is the channel additive white Gaussian noise and $\widetilde{w} \sim \mathcal{N}\left(0, \sigma_w^2\right)$. $\diamondsuit$

In essence, the combination of **S1**-**S3** corresponds to obtaining at the destination a noisy projection of the data vector x onto φ. Thus, at the end of **S3**, the destination can estimate υ as $\widehat{\upsilon} = r/\sqrt{E_o}$ and the resulting distortion is given by

$$D_\upsilon \;=\; \mathbb{E}\left[|\widehat{\upsilon} - \upsilon|^2\right] \;=\; \sigma_w^2 \;+\; \frac{\sigma_z^2}{E_o} \qquad (4)$$

where the first term in the above expression is due to the measurement noise (unaffected by E_o) and the second term is due to the communication noise that decays as $1/E_o$. Moreover, since a total of n_φ nodes transmitted during this distributed projection, each with energy $\preceq \frac{E_o T}{n_\varphi}$, the total energy consumed in obtaining $\widehat{\upsilon}$ at the destination is given by

$$E \;\preceq\; \left(\frac{E_o T}{n_\varphi}\right) n_\varphi \;\sim\; E_o \qquad (5)$$

From (4), it is clear that one way of reducing the distortion of the projection coefficient $\widehat{\upsilon}$ is to increase E_o. If, however, there are some constraints on the maximum allowable E_o, then the destination can repeat the above procedure over p independent channel uses to obtain $\{\widehat{\upsilon}_k\}_{k=1}^p$ and then calculate $\widehat{\upsilon}$ as $\widehat{\upsilon} = \frac{1}{p}\sum_{k=1}^p \widehat{\upsilon}_k$. For a fixed E_o, this procedure would give us the following latency (L) and E-D_υ relations

$$D_\upsilon \;=\; \sigma_w^2 \;+\; \frac{\sigma_z^2}{p\,E_o} \qquad (6)$$

$$E \;\sim\; (p\,E_o) \qquad (7)$$

$$L \;=\; p \qquad (8)$$

2. MAIN RESULTS

In this section, using the scheme of distributed projections as a basic building block, we shall derive results for latency (L) and E-D tradeoff first under the assumption that the destination has perfect knowledge of the subspace in which x^* lies and then under the assumption that the destination has little or no knowledge of the subspace in which x^* lives.

[1]Note that the φ_j's can also be made available to the sensors using other methods and feedback from the destination to the sensors is not really necessary. See the discussion in Section 3 for more details

[2]Because of **A2**, we can ignore the effect of path loss on the received signal as it would just be a constant uniform attenuation (independent of n)

2.1. Signal Reconstruction: Known Subspace

Let $\Psi \triangleq \{\psi_i\}_{i=1}^n$ be an orthonormal basis of $\mathbb{R}^n$ such that $x^* = \sum_{i=1}^m \theta_i\,\psi_i$ (perhaps after re-labeling the indices i), where each coefficient θ_i is computed as a projection (inner-product) of the form $\theta_i = \psi_i^T x^* = \sum_{j=1}^n \psi_{ij} x_j^*$. Then, under the assumption that the destination has perfect knowledge of the subspace in which x^* lies, we have the following latency (L) and E-D relations.

Theorem 1. *If the destination knows Ψ as well as which elements $\psi_i \in \Psi$ give the sparse representation of x^* then (a) There exists an estimation scheme such that $\forall\ \sigma_z^2 > 0$, $E_o > 0$, and $k\ (= p\,m) \geq m$, where $p \in \mathbb{N}$*

$$D \;=\; \frac{m}{n}\left(\sigma_w^2 + \frac{m\,\sigma_z^2}{k\,E_o}\right) \tag{9}$$

$$E \;\sim\; (k\,E_o) \tag{10}$$

$$L \;=\; k \tag{11}$$

and; (b) if Φ is any other orthonormal basis of $\mathbb{R}^n$ such that $x^* = \sum_{i=1}^m \eta_i\,\phi_i$ and the destination knows Φ as well as which elements $\phi_i \in \Phi$ give the sparse representation of x^* then again the results of (a) hold. Moreover, (c) If $\sigma_w^2 \ll \frac{m\,\sigma_z^2}{k\,E_o}$ then (9) reduces to

$$D \;\approx\; \frac{m}{n}\left(\frac{m\,\sigma_z^2}{k\,E_o}\right) \tag{12}$$

Sketch of Proof: (c) follows trivially from (a). For (a), the destination computes m distributed projections of x onto $\{\psi_i\}_{i=1}^m$ over m independent channel uses. The destination can also repeat this procedure p times, as described in Section 1.2, for each of the m basis elements. Thus, at the end of $k = p\,m$ projections (and channel uses), the destination has access to m projection coefficients $\{\widehat{\theta}_i\}_{i=1}^m$ such that $\mathbb{E}\left[|\widehat{\theta}_i - \theta_i|^2\right] = \sigma_w^2 + \frac{\sigma_z^2}{p\,E_o}$. Therefore, the destination can estimate x^* as $\widehat{x} = \sum_{i=1}^m \widehat{\theta}_i \psi_i$ and the resulting distortion is given by

$$D \;=\; \mathbb{E}\left[\frac{1}{n}\left\|\widehat{x} - x^*\right\|^2\right] = \frac{1}{n}\sum_{i=1}^m \mathbb{E}\left[|\widehat{\theta}_i - \theta_i|^2\right] \tag{13}$$

$$\;=\; \frac{m}{n}\left(\sigma_w^2 + \frac{\sigma_z^2}{p\,E_o}\right) = \frac{m}{n}\left(\sigma_w^2 + \frac{m\,\sigma_z^2}{k\,E_o}\right) \tag{14}$$

Moreover, $E \sim (k\,E_o)$ and $L = k$ follow trivially from Section 1.2 and the fact that we are computing a total of k projections.

For (b), let $\Phi \triangleq \{\phi_i\}_{i=1}^n$ be any other orthonormal basis of $\mathbb{R}^n$ (known to the destination) such that $x^* = \sum_{i=1}^m \eta_i \phi_i$ (perhaps after re-labeling the indices i), where each coefficient $\eta_{(i)}$ is computed as an inner product of the form $\eta_i = \phi_i^T x^* = \sum_{j=1}^n \phi_{ij} x_j^*$. Then, if the destination wants to reconstruct x^* by projecting x onto $\{\phi_i\}_{i=1}^m$ using the above procedure, it is easy to see that the above results would still hold. ◊

When there is significant measurement noise, it is obvious from (9) that the distortion scaling is limited by the measurement noise term in D i.e., $\left(\frac{m}{n}\right)\sigma_w^2 \leq D$. In that case, a stronger result can be obtained as stated in the following Corollary.

Corollary 1. *If $\sigma_w^2 > 0$ and $\sigma_w^2 \sim \sigma_z^2$, then (9) in Theorem 1 reduces to $D \sim \left(\frac{m}{n}\right)$ and it is possible to achieve this distortion scaling by using $E \sim m$ and $L \sim m$.*

Proof: Put $(kE_o) = m$ in Theorem 1 and the result follows. ◊

An important implication of Theorem 1 is that all m-dimensional orthonormal bases that span the subspace in which x^* lies are equivalent in terms of the latency (L) and E-D tradeoff and thus, for the purposes of reconstruction of x^*, using any one of these bases is as good as using any other basis. Generally speaking, however, even if the destination knows the basis of $\mathbb{R}^n$ in which x^* is sparse, it is highly unlikely that it will know ahead of time which m of the basis elements give the sparse representation of x^* and this is where the *universality* of the following scheme comes into play.

2.2. Signal Reconstruction: Unknown Subspace

Let us now assume that the destination has little or no knowledge of the subspace in which x^* lives. As mentioned in Section 2.1, this includes the scenario where the destination knows the basis of $\mathbb{R}^n$ in which x^* is sparse but does not know which of the m elements of that basis to use. In that case, the destination employs state-of-the-art compression techniques based on random projections of the data to efficiently summarize the information in x, resulting in the following latency (L) and E-D relations.

Theorem 2. *If the subspace of $\mathbb{R}^n$ in which x^* lies is not known to the destination then there exists a constant $C_1 > 0$ and an estimation scheme such that $\forall\ E_o > 0$, and $k \in \mathbb{N}$ such that $(m \log n) < k \leq (n \log n)$*

$$D \;\leq\; C_1\left(\frac{m\,\log n}{k}\right) \tag{15}$$

$$E \;\sim\; (k\,E_o) \tag{16}$$

$$L \;=\; k \tag{17}$$

Moreover, (b) In the case of little or no measurement noise i.e., $\sigma_w^2 = 0$, similar results hold with slightly different constant C_1.

Sketch of Proof: The destination generates k length-n random vectors $\{\phi_i\}_{i=1}^k$ such that the components ϕ_{ij}, $j = 1, \ldots, n$, of ϕ_i are i.i.d random variables (independent of w_j) which take the values $\pm 1/\sqrt{n}$ with equal probability. Thus, $\mathbb{E}\left[\phi_{ij}\right] = 0$ and $\mathbb{E}\left[\phi_{ij}^2\right] = 1/n$. The destination now computes k distributed (random) projections of x onto $\{\phi_i\}_{i=1}^k$ over k independent channel uses. Thus, at the end of k projections, the destination has access to the k noisy random projections ($\{\eta_i\}_{i=1}^k : \eta_i = \phi_i^T x^* + \phi_i^T w + \widetilde{z}_i$, where $\widetilde{z}_i \sim \mathcal{N}\left(0, \sigma_z^2/E_o\right)$) of noisy data that lies in an m-dimensional subspace. And since the destination has access to the original random vectors $\{\phi_i\}_{i=1}^k$, it is easy to see from the theory developed by Haupt and Nowak in [8] that x^* can be easily reconstructed from $\{\eta_i\}_{i=1}^k$ such that the resulting distortion behaves like

$$D \;=\; \mathbb{E}\left[\frac{1}{n}\left\|\widehat{x} - x^*\right\|^2\right] \tag{18}$$

$$\;\leq\; C_1\left(\frac{m\log n}{k}\right), \tag{19}$$

where $C_1 > 0$ is a constant. Rather than reworking the proof of this statment, we refer the reader to [8] for further details. Similarly, (b) follows from Corollary 2 in [8]. Moreover, $E \sim (k\,E_o)$ and $L = k$ follow trivially from Section 1.2 and the fact that we are computing a total of k projections. ◊

As a motivation for Theorem 2, consider the following simple example. Suppose x^* is a spatially non-sparse vector of length n ($S_p(x^*) = n$) with only one non-zero coefficient of amplitude $\sqrt{n}$

in some transform basis $\Psi \triangleq \{\psi_i\}_{i=1}^n$ such that $\|x^*\|^2/n = 1$. This is an example of the case where we know the basis in which x^* is sparse but do not know which elements of the basis to use. One naive approach to this problem is to require each sensor to digitally transmit its measurement to the destination, where the reconstruction is then performed. Alternatively, all the sensors might collaboratively process their measurements to reconstruct x^* in-network and then transmit the result to the destination. Both approaches, however, while providing us with consistent estimates, would require at least $E \sim n$ and $L = n$.

Another approach to this problem could be random transform point sampling where the destination computes a distributed projection of the data onto ψ_i and i is selected uniformly at random from the set $\{1, \ldots, n\}$. Ignoring the distortion due to the measurement noise, the squared reconstruction error is 0 if the spike in Ψ domain corresponds to ψ_i and 1 otherwise and the probability of not finding the spike in k trials is $\left(1 - \frac{1}{n}\right)^k$, giving an average squared error of $\left(1 - \frac{1}{n}\right)^k .1 + (k/n).0 = \left(1 - \frac{1}{n}\right)^k$. If n is large, we can approximate this by $D = \left(1 - \frac{1}{n}\right)^k \approx e^{-k/n}$. Therefore, for any $k < n$, we have $D \to 1$ as $n \to \infty$ while E and $L \sim k$, and for $k = n$, we have $D = e^{-1}$ while E and $L \sim n$ (grow linearly with n). However, Theorem 2 guarantees us a consistent estimator even in this situation ($m = 1$) by taking $k = n^\alpha$ $(0 < \alpha < 1)$, resulting in $D \preceq \left(\frac{\log n}{n^\alpha}\right) \preceq \left(n^{-\alpha}\right)$, while E and $L \sim n^\alpha$ (grow sub-linearly with n).

2.3. Cost of Universality

It is important to realize that the added flexibility and universality of the scheme proposed in Theorem 2 comes at the price of increased latency (L) and a less favorable E-D tradeoff. For example, an immediate consequence of Theorem 2 is that using this scheme, the destination needs to expend at least $E \sim (m \log n)$ amount of energy and would incur a latency of at least $L \sim (m \log n)$ for *barely* consistent estimator of x^*, whereas if one had knowledge of the subspace in which x^* lied then, assuming $\sigma_w^2 \sim \sigma_z^2$ (Corollary 1), one would only require $E \sim m$ and get $D \sim \left(\frac{m}{n}\right)$.

In particular, if $\sigma_w^2 \ll \frac{m\,\sigma_z^2}{k\,E_o}$, then Theorem 2 and (12) in Theorem 1 reveal that:

1) For a fixed projection energy budget E_o and total energy budget E, the distortion incurred without knowledge of the signal subspace is about a factor of n/m times larger than if one does know the subspace.

2) For a fixed distortion D, the total energy budget E and the latency L in data acquisition without knowledge of the signal subspace is about a factor of n/m times more than if one does know the subspace.

Similarly from Corollay 1, if $\sigma_w^2 > 0$ and $\sigma_w^2 \sim \sigma_z^2$, then for the distortion of the universal scheme to be equivalent to the distortion of the known subspace case ($D \sim m/n$), the destination must expend $E \sim n$ energy and incur a latency of $L \sim n$; again a factor of about n/m times more than if one does know the subspace.

3. DISCUSSION AND EXTENSIONS

In this paper, we have described and analyzed a universal matched-source channel communication scheme for reconstruction of sensor network data at a distant destination. Our scheme is universal in the sense that it provides us with a consistent estimation scheme such that E grows sub-linearly with n without requiring any prior knowledge about the sensed data. Moreover, this scheme naturally integrates the operations of processing and communications, thereby reducing the amount of processing and communications required inside the network and provides us with a system that often acts less like networks and more like coherent ensembles of sensors, thereby reducing the overhead of network-centric functions such as routing etc. The universality of our proposed scheme, however, comes at the price of increased latency (L) and a less favorable E-D tradeoff by a factor of about n/m, which is a direct consequence of not having sufficient prior knowledge about sensed data, forcing us to probe the entire n-dimensional space instead of focusing our energy on the m-dimensional subspace in which x^* lives.

At the heart of our approach is an efficient (distributed) method of estimating projections of the noiseless sensor network data onto any normalized vector in $\mathbb{R}^n$ by using only a fixed amount of energy (independent of n). Depending upon the structure of the normalized vector, this approach may require the destination to be able to address each sensor individually. Pre-storage of individual elements of the normalized vector in each sensor node is another option which might not be always feasible because of node failures, changes in the structure of sensed data etc. If, however, the sensor network employs the universal scheme based on random projections then the information can be efficiently generated by each sensor by using the seed of a pseudo-random generator and the addresses of the nodes in order to draw the elements of the random vectors $\{\phi_i\}_{i=1}^k$. Similarly, the destination can easily reconstruct the vectors $\{\phi_i\}$ given the seed values and the number of nodes in the network.

An important consequence of our proposed scheme is that it requires phase synchronization among n nodes during each projection – something that might not always be feasible. An interesting extension of our system involves applying this scheme to disjoint subsets of x and reconstructing x^* from that. Our other future work includes extensions to compressible signals and studying the effect of imperfect node synchronization on the proposed scheme.

4. REFERENCES

[1] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. IPSN'03*, Apr. 2003, pp. 162–177.

[2] M. Gastpar and M. Vetterli, "Power-bandwidth-distortion scaling laws for sensor networks," in *Proc. IPSN'04*, Apr. 2004, pp. 320–329.

[3] W. U. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *Proc. IPSN'05*, Apr. 2005, pp. 332–339.

[4] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 745–754, Apr. 2005.

[5] K. Liu and A. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," in *Proc. 42nd Allerton Conf. on Commun., Control and Comp.*, Oct. 2004.

[6] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," Submitted, Jun. 2004.

[7] D. L. Donoho, "Compressed sensing," Manuscript, Sep. 2004.

[8] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," Submitted to IEEE Trans. Info. Th., Mar. 2005.

[9] S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Sig. Proc. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.