

Random Access Compressed Sensing for Energy-Efficient Underwater Sensor Networks

Fatemeh Fazel, Maryam Fazel and Milica Stojanovic

Abstract—Inspired by the theory of compressed sensing and employing random channel access, we propose a distributed energy-efficient sensor network scheme denoted by Random Access Compressed Sensing (RACS). The proposed scheme is suitable for long-term deployment of large underwater networks, in which saving energy and bandwidth is of crucial importance. During each frame, a randomly chosen subset of nodes participate in the sensing process, then share the channel using random access. Due to the nature of random access, packets may collide at the fusion center. To account for the packet loss that occurs due to collisions, the network design employs the concept of sufficient sensing probability. With this probability, sufficiently many data packets – as required for field reconstruction based on compressed sensing – are to be received. The RACS scheme prolongs network life-time while employing a simple and distributed scheme which eliminates the need for scheduling.

Index Terms—Sensor networks, compressed sensing, wireless communications, underwater acoustic networks, random access.

I. INTRODUCTION

UNDERWATER sensor networks are envisioned as consisting of a number of static sensor nodes and/or vehicles that are deployed over a region of interest to monitor a physical phenomenon. Applications of such networks are in oceanographic data collection (e.g., temperature, salinity, zonal and meridional currents), field monitoring and disaster prevention [1] [2]. Wireless acoustic communication is the physical layer technology used in underwater networking. In this paper, we consider a static area network, where sensor nodes are anchored to the bottom of the ocean and deployed for long periods of time. Each sensor node communicates its observations of the field to a central node, referred to as the Fusion Center (FC) and the FC reconstructs the map of the physical field. Bandwidth and battery power are severely limited in underwater networks, and hence energy and bandwidth efficiency are of particular importance.

Exploiting the fact that most natural phenomena are compressible (sparse) in an appropriate basis, we employ compressed sensing to reduce the energy consumption of the network. The theory of compressed sensing establishes that

Manuscript received 3 October 2010; revised 15 February 2011. Research funded in part by ONR grant N00014-09-1-0700, NSF grant 0831728, and NSF CAREER grant ECCS-0847077. Preliminary results of this paper have been presented at the Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing.

F. Fazel and M. Stojanovic are with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, 02115, USA (email: ffazel,millitsa@ece.neu.edu).

M. Fazel is with the Department of Electrical Engineering, University of Washington, Seattle, WA, 98195, USA (email: mfazel@u.washington.edu).

Digital Object Identifier 10.1109/JSAC.2011.110915.

under certain conditions, exact signal recovery is possible with a small number of random measurements [3][4]. Authors in [5] are the first to introduce the application of compressed sensing in networks. In [5][6] and [7] the authors used phase-coherent transmission of randomly-weighted data from sensor nodes to the FC over a dedicated multiple-access channel, to form distributed projections of data onto an appropriate basis at the FC. Note that in this approach sensors need to be perfectly synchronized which is a difficult assumption to maintain in underwater acoustic networks. Reference [8] proposes compressive cooperative spatial mapping using mobile sensors based on a small set of observations. In [9] ultra-low power infrastructure monitoring is achieved by employing data compression and a low-collision MAC protocol. In [10] adaptive compressed sensing is applied to wireless sensor networks. Initially, a random set of readings are observed at the FC. If the accuracy level is not satisfactory a projection vector is obtained and the data is updated. The authors determine the projection vector so as to optimize the information gain per energy expenditure. A number of references, such as [11], [12] and [13] focus on phenomena that are sparse in the spatial domain, e.g., event detection or tracking of multiple targets. In [12] authors consider a decentralized network (without FC), where active nodes exchange measurements locally. The authors formulate sparse recovery as a decentralized consensus optimization problem and show that their iterative algorithm converges to a globally optimal solution. In [13] sensors are tracking the location of an audio source, transmitting their readings to the FC. In this setting, the signals appearing at each sensor are jointly sparse. The authors show that a very small number of measurements can achieve the signal detection goal. Authors in [14] and [15] also consider spatial mapping using mobile sensors (robots), [15] proposes an efficient way to reconstruct natural fields using random-walk-based sampling and compressed sensing. Finally, in [16] capacity bounds of an on-off random multiple access channel are determined by transforming the problem to an equivalent compressed sensing problem and using sparsity detection algorithms.

In this work, we consider an underwater sensor network that measures a physical phenomenon for geographical and environmental monitoring purposes. We assume that the physical phenomenon to be studied is compressible (sparse) in the frequency domain. The proposed method, based on compressed sensing and random access, results in a simple and energy-efficient scheme referred to as *Random Access Compressed Sensing* (RACS). The system functions consist of (i) a sampling procedure, during which sensor nodes perform measurements; followed by (ii) a channel access

method, during which measurements are transmitted to the FC; and finally (iii) a reconstruction process, during which sparse recovery algorithms are used to recover the measured field at the FC. In the sampling procedure, inspired by the theory of compressed sensing, we employ random sensing, while for the channel access phase, we propose a simple random access. As in any random access, the data packets of two or more sensors may collide at the FC. The key idea is that random collisions (which are inevitable) do not change the random nature of the observations provided to the FC. Since the FC only needs to receive some, and not all the sensor packets, it can simply disregard the collisions. The FC obtains an incomplete set of measurements (due to both random sensing and losses due to random access) from which it reconstructs the field using compressed sensing techniques. Note, however, that in order to achieve successful reconstruction, a certain minimum number of measurements—as determined by compressed sensing theory—are required at the FC. We thus need to compensate for the collision losses by initially selecting the number of participating sensors to be greater than the minimum number of required packets. We provide an analytical framework for system design based on the sufficient sensing probability. Note that our method is completely distributed, requiring no coordination among nodes. It also requires no downlink feedback from the FC to the sensor nodes.

The paper is organized as follows: In Section II we introduce the network model. In Section III, we outline the system model and introduce the RACS scheme. Section IV provides an analytical model for RACS based on which we propose a network design methodology. In Section V we provide performance assessment of our scheme and compare the energy and bandwidth usage of RACS with that of a conventional network. In Section VI using a real data example, we demonstrate the performance of the RACS scheme. Finally, we provide concluding remarks in Section VII.

Notation: We denote by ℓ_p the p -norm of a vector $x = [x_1, \dots, x_N]^T$ defined by $\|x\|_{\ell_p} = \left(\sum_{i=1}^N |x_i|^p\right)^{1/p}$. If \mathbf{V} is a $k \times l$ matrix, $\text{vec}(\mathbf{V})$ denotes the $kl \times 1$ vector formed by stacking the columns of matrix \mathbf{V} , i.e., $\text{vec}(\mathbf{V}) = [v_{11} \dots v_{l1} \dots v_{1k} \dots v_{lk}]^T$. We denote by $B(N, p)$ the Binomial probability distribution of the number of successes in a sequence of N independent experiments, each of which has a success probability p . Finally, $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of matrices \mathbf{A} and \mathbf{B} .

II. SYSTEM MODEL

Consider a grid network shown in Fig. 1, which consists of $N = IJ$ sensors located on a two-dimensional plane, with J and I sensors in x and y directions, respectively.¹ The sensors are separated by distance d in each direction. Let us define the coverage area A of a network as the total area covered by the sensors, in our grid network $A = Nd^2$. The network is deployed to monitor a physical phenomenon, $u(x, y, t)$, (e.g., temperature, pressure, current, etc.) over a long period of time.

¹Note that the results of this paper can be easily extended to three-dimensional (volume) networks as well.

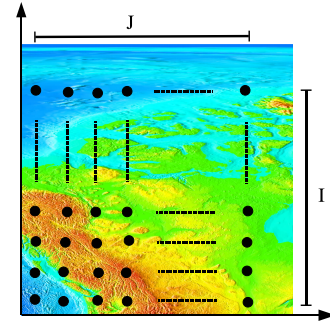


Fig. 1. An area sensor network consisting of $N = IJ$ sensor nodes.

Such long-term monitoring is crucial in climate monitoring or environmental surveillance applications.

In frame n , the sensor node located at position (i, j) in the network grid acquires a measurement $u_{ij}(n) = u(x_j, y_i, n)$, where x_j and y_i denote the sensor's position in the 2-dimensional space. The measurements are encoded, along with the sensor's location tag, into a data packet of L bits, which is then modulated and transmitted to the FC. Upon reception, the FC demodulates the signal and extracts the measurement information from which it reconstructs the map of the field. Assuming that the system has bandwidth B and that each sensor transmits at a bit-rate equal to the bandwidth, the packet duration is $T_p = L/B$. Let D_i denote the distance of node i from the FC, where $i \in \{1, \dots, N\}$. The propagation delay of sensor i 's packet is thus given by $\tau_i = \frac{D_i}{c}$, where $c = 1500$ meters/sec is the nominal speed of sound.

In this paper, we consider a frame-based (slotted) transmission, i.e., the FC collects the incoming data packets during a frame of length T . At the end of the frame, the FC reconstructs the field based on the data packets received during that frame. Once the reconstruction is performed, the frame is discarded and FC waits for a new set of data in the next frame. In order to determine a reasonable frame duration T , we consider the correlation properties of the physical process $u(x, y, t)$. Let us define the coherence time T_{coh} of a process as the time-duration over which the process almost de-correlates in time, i.e., the process is slowly varying during T_{coh} . A conventional design choice is thus to obtain a new map of the field $u(x, y, t)$ at least once per T_{coh} .

The map of the process over the entire sensor network is denoted by $\mathbf{U}(n)$

$$\mathbf{U}(n) = \begin{bmatrix} u_{11}(n) & \dots & u_{1J}(n) \\ \vdots & \vdots & \vdots \\ \dots & u_{ij}(n) & \dots \\ \vdots & \vdots & \vdots \\ u_{I1}(n) & \dots & u_{IJ}(n) \end{bmatrix} \quad (1)$$

The data gathering procedure in a network consists of two phases: a) sensing and b) communication. The sensing phase can be (i) deterministic (conventional case), in which case all the sensors sample the physical phenomenon; or (ii) random (compressed), in which case only a random subset of sensor nodes participate in sensing. The nodes that have taken part in sensing now need to access the channel in

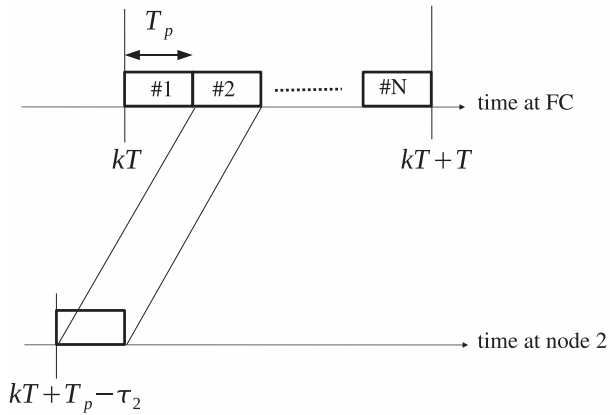


Fig. 2. The scheduling required at each node in the benchmark case of TDMA.

order to communicate their measurements to the FC. Multiple-access schemes are generally divided into two categories: (i) deterministic access methods, e.g., TDMA, FDMA, and CDMA; and (ii) random access methods, e.g., Aloha, CSMA, and CSMA/CD. In what follows we will consider both types of access for use with compressed sensing, while we consider deterministic sensing as a benchmark for comparison.

A. Conventional Network (Benchmark)

As a benchmark design, we consider a sensor network with deterministic sensing and deterministic access, i.e., in each frame, all N nodes conduct measurements and transmit their measurement packets to the FC using a deterministic multiple-access scheme. We assume that the conventional network employs standard TDMA. This requires scheduling at each node such that packets from different nodes arrive back-to-back at the FC. Fig. 2 depicts the required scheduling process. The received data at the FC at frame n , denoted by $\mathbf{y}(n)$, is given by

$$\mathbf{y}(n) = \mathbf{u}(n) + \mathbf{z}(n) \quad (2)$$

where, $\mathbf{u}(n) = \text{vec}(\mathbf{U}(n))$ and $\mathbf{z}(n)$ represents the sensing noise which arises due to the limitations in the sensing device. The communication noise translates into bit errors, i.e., it does not appear as an additive term in Eq. (2). In the present analysis we neglect the communication noise. In TDMA, one frame of data contains N packets; therefore, $T = NT_p$. The network needs updated measurements every $T \leq T_{coh}$. The total number of nodes that can be deployed in a conventional network is thus upper-bounded by

$$N \leq \frac{T_{coh}}{T_p} \quad (3)$$

where T_{coh} is the property of the monitored field. Consequently, the coverage area of a conventional network is limited to $A = T_{coh}d^2/T_p$.

III. RANDOM SENSING

Most natural phenomena have a compressible (sparse) representation in the spatial frequency domain, and we therefore assume that the vector of Fourier coefficients of $\mathbf{U}(n)$ is

sparse. Specifically, if $\mathbf{V}(n)$ is the two-dimensional spatial discrete Fourier transform of $\mathbf{U}(n)$, it can be shown that $\mathbf{v}(n) = (\mathbf{W}_J \otimes \mathbf{W}_I)\mathbf{u}(n)$, where $\mathbf{v}(n) = \text{vec}(\mathbf{V}(n))$ and \mathbf{W}_I is the matrix of discrete Fourier transform coefficients ($\mathbf{W}_I[m, k] = e^{-j2\pi mk/I}$). Thus, in our case, the Fourier representation $\mathbf{v}(n)$ is assumed to be sparse. Note that a *sparse* signal is a signal that can be represented by a small number of non-zero coefficients, compared to the dimension of the signal. As an example, Fig. 3 shows the zonal currents recorded at the Southern California bight, and their corresponding discrete Fourier transform. One can show that almost 99% of the energy of the signal is contained in $S = 13$ Fourier coefficients. Based on the theory of compressed sensing, if a signal has a sparse representation in some domain, it can be recovered from a small subset of random measurements [3], [4]. Thus taking into account the sparsity of natural phenomena, we can reduce the number of measurements required for field recovery from N to some $M < N$.

Let us assume that all the nodes know the beginning time of a frame at the FC. At frame n , a subset of sensors is selected at random to conduct measurements. By randomly selecting sensors, we perform the compression directly in the spatial domain. If we denote by $\mathbf{y}(n)$ the observations of a random subset of M sensors, the received data vector at the FC can be expressed as

$$\mathbf{y}(n) = \mathbf{R}(n)\mathbf{u}(n) + \mathbf{z}(n) \quad (4)$$

where $\mathbf{R}(n)$ is an $M \times N$ random selection matrix for frame n , consisting of M rows of the identity matrix selected uniformly at random. Noting that $\mathbf{u}(n) = \mathbf{\Psi}\mathbf{v}(n)$, where $\mathbf{\Psi} = (\mathbf{W}_J \otimes \mathbf{W}_I)^{-1}$ is the Inverse Discrete Fourier Transform (IDFT) matrix, Eq. (4) can be re-written in terms of the sparse vector $\mathbf{v}(n)$ as

$$\mathbf{y}(n) = \mathbf{R}(n)\mathbf{\Psi}\mathbf{v}(n) + \mathbf{z}(n) \quad (5)$$

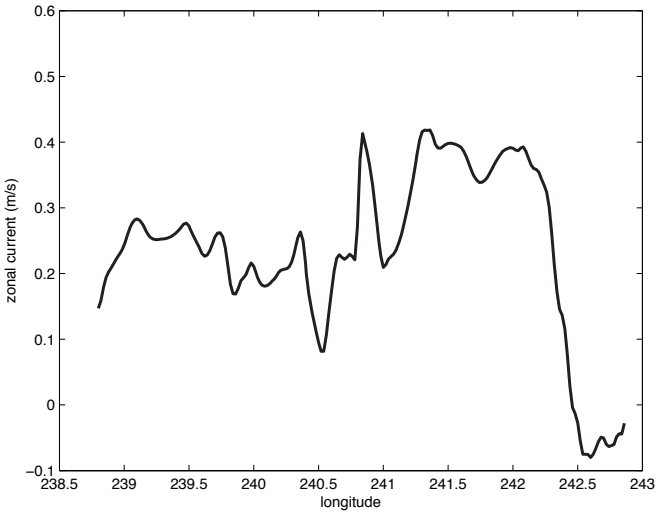
The IDFT matrix $\mathbf{\Psi}$ is referred to as the representation basis, which is the basis over which $\mathbf{u}(n)$ has a sparse representation.

To reconstruct the field at the end of the frame n , the FC first tries to recover the vector $\mathbf{v}(n)$ as accurately as possible, then uses it to construct the map $\mathbf{U}(n)$. Given the observations $\mathbf{y}(n)$, the random selection pattern $\mathbf{R}(n)$ and the sparsity basis $\mathbf{\Psi}$, and in the absence of sensing noise $\mathbf{z}(n)$ – which is the case we will be focusing on – reconstruction can be performed by solving the following minimization problem:

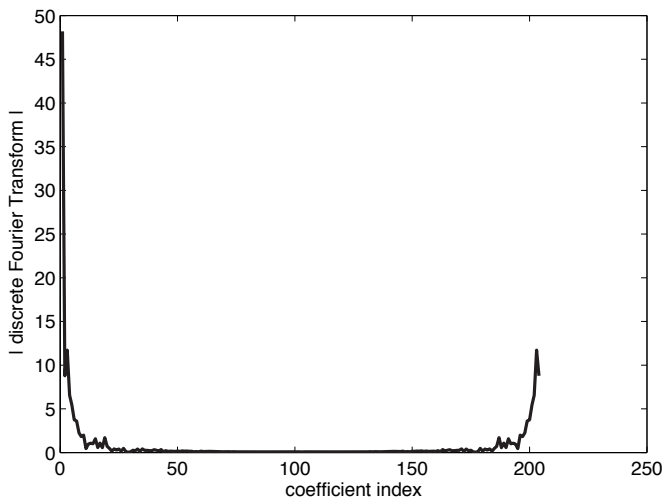
$$\text{minimize}_{\tilde{\mathbf{v}}(n)} \|\tilde{\mathbf{v}}(n)\|_{\ell_1} \quad \text{subject to} \quad \mathbf{R}(n)\mathbf{\Psi}\tilde{\mathbf{v}}(n) = \mathbf{y}(n). \quad (6)$$

The theory of compressed sensing (specifically, [17]) states that as long as the number of observations, picked uniformly at random, is greater than $N_s = CS \log N$, then with very high probability the solution to the convex optimization problem (6) is unique and is equal to $\mathbf{v}(n)$. Here C is a constant that is independent of N and S (see [17] for the details).

We thus conclude that in our wireless network setting, it suffices to ensure that the FC collects at least N_s packets picked uniformly at random from different sensors to guarantee accurate reconstruction of the field with very high probability.



(a) original field.



(b) amplitude of the Fourier transform of the field.

Fig. 3. (a) Zonal current (m/s) at a latitude of 32.5° , plotted versus the longitude [$238.5^\circ, 243^\circ$]; and (b) the amplitude of the corresponding spatial Fourier transform. Almost 99% of the energy of the signal of size $N = 204$ is contained in $S = 13$ Fourier coefficients. This data is accessible at <http://lourocean.jpl.nasa.gov>.

A. Centralized Random Sensing / Deterministic Access (R/D)

We focus on the centralized selection to illustrate the random sensing concept before moving on to the distributed selection in the next section. In this scheme, the FC picks a random subset of M sensors for sampling and broadcasts the selected set of nodes in each frame. In order to obtain perfect reconstruction, it has to be that $M \geq N_s$. The selected nodes then sample the physical process $u(x, y, t)$ and send their measurements back to the FC using a multiple-access method of choice. Since the FC broadcasts the selected subset, all sensors learn when a frame will start, which nodes will be transmitting and their transmission order. Therefore, the network can simply use deterministic access (TDMA) with M slots as shown in Fig. 4. All transmitting nodes organize their transmissions such that they are received at the FC in

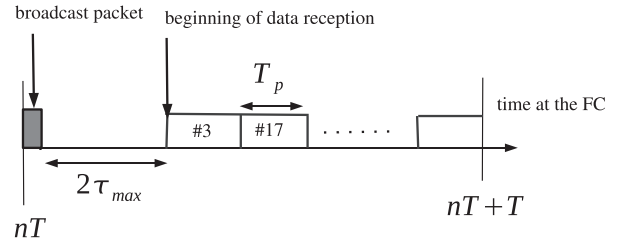


Fig. 4. The frame structure in the R/D scheme. The FC broadcasts the selected subset, the nodes then schedule their transmissions.

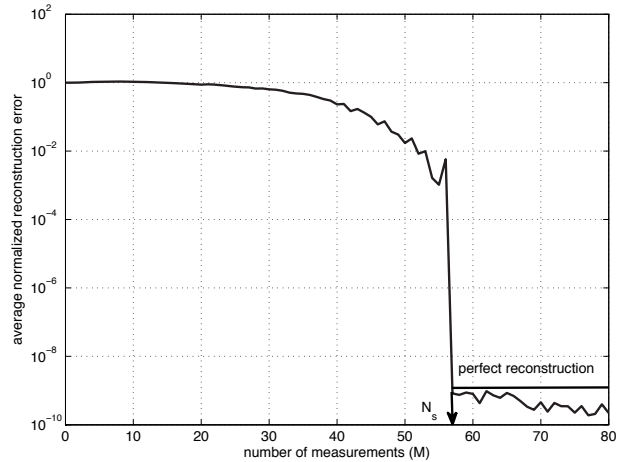


Fig. 5. For a network of size $N = 1000$ and $N_{sim} = 100$ randomly generated signals with sparsity $S = 10$, the average normalized reconstruction error is plotted versus the number of measurements M . The required number of measurements to obtain perfect reconstruction is $N_s \approx 57$ as shown in the figure.

the requested order. Thus, a frame of duration T consists of the round-trip broadcast time followed by M packets of data, i.e., $T = 2\tau_{max} + MT_p$, where $\tau_{max} = \frac{1}{c} \max_{i \in \{1, \dots, N\}} D_i$ denotes the longest propagation delay in the network.

The required number of observations $N_s = CS \log N$ depends on the value of the constant C , a theoretical upper-bound for which is offered in [17]. However, one can find N_s empirically as the number of measurements for which the reconstruction error is negligible. The empirical value of N_s is typically much smaller than the one obtained using the theoretical bounds. Here, we illustrate finding N_s in our setting for the following example set of system parameters: $I = 50$, $J = 20$, and $S = 10$. We study the recovery of these signals from different numbers of random measurements in a noise-free setting. Fig. 5 shows the average reconstruction error plotted versus the number of measurements. As seen in the figure, for $M \geq 57$ full recovery is attained. Hence, for the given system parameters, a reasonable choice for N_s is determined to be $N_s = 57$.

B. Distributed Random Sensing / Random Access (R/R)

As discussed in Section III-A, centralized random sensing requires scheduling among sensors by downlink transmission from the FC. In order to eliminate the need for downlink transmissions at each frame, we decentralize the process of

selecting a random subset of nodes. This can be done by equipping the sensors with independent, identically distributed Bernoulli random generators, i.e., by having each sensor toss an independent coin. At the beginning of a frame, each node determines whether it will participate in the sensing process, which occurs with some probability p . The total number of sensors selected for sampling in a frame, M , is now a random variable with a Binomial distribution, $M \sim B(N, p)$.

The principle of distributed sensing is thus very similar to that of the centralized sensing from the viewpoint of providing a random subset of observations. Its advantage is in the fact that it eliminates the need for duplexing, i.e., no downlink transmission is required from the FC. However, deterministic access is no longer applicable, because a node has no knowledge of the other nodes that transmit, and hence cannot schedule its transmission.² If we couple random sensor selection with random channel access this problem is eliminated. Furthermore, employing random access eliminates the overhead broadcast time and the next data frame starts immediately.

In random access, each sensor i picks a random transmission delay θ_i uniformly in $[0, T - T_p]$. In this scheme there is possibility of collision. A collision is said to have occurred if packets from different sensors overlap in time at the FC. The key idea in RACS is to let the FC simply discard the colliding packets. This approach is motivated by the compressed sensing theory and the fact that the FC does not care *which* specific sensors are selected, as long as (i) the selected subset is chosen uniformly at random, and (ii) there are sufficiently many collision-free packets received to allow for the reconstruction of the field. Therefore, in a RACS scheme, once a collision is detected the FC simply discards the colliding packets and reconstructs the field using the rest. Note that the random reduction matrix $\mathbf{R}(n)$ in Eq. (5) now includes both the effects of random selection and of random collisions.

The proposed frame-based RACS is summarized below:

- Step 1. At the beginning of a frame, sensor node i tosses a coin to determine whether it participates in sensing (with probability p) or stays inactive (with probability $1 - p$) during that frame.
- Step 2. If node i is selected for sensing, it measures the physical quantity of interest and encodes it into a packet of L bits. The sensor's location is also included in the packet.
- Step 3. Node i picks a uniformly-distributed delay θ_i for the transmission of its packet.
- Step 4. FC collects the packets received during one frame. If a collision is detected, FC discards the colliding packets.
- Step 5. At the end of the frame, FC uses the correctly received packets to reconstruct the data using ℓ_1 minimization (or other sparse recovery methods [18]). We assume that packets which do not collide are correctly received.

Let K denote the number of correctly received packets at

²One could in principle reserve N slots, but since only a subset of sensors transmit such a scheme would be wasteful.

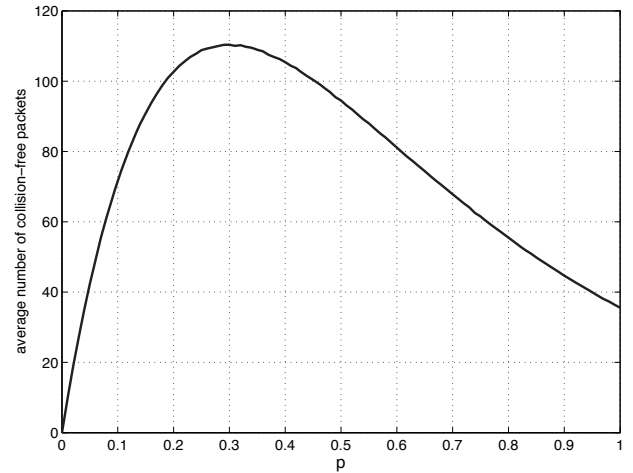


Fig. 6. Average number of collision-free packets \bar{K} versus p ; system parameters are $N = 1000$, $T = 120$ s and $T_p = 0.2$ s.

the FC during one frame. Fig. 6 shows the average number of collision-free received packets \bar{K} versus the per-node sensing probability p , for an example network of $N = 1000$, $T_p = 0.2$ s and $T = 120$ s. As seen in the figure, there is an interplay between the number of measurements and the number of collisions. While increasing p results in a greater number of measurements M , and could thus improve the accuracy of reconstruction, it also increases the probability of collision and after a certain point may even decrease the number of useful packets received at the FC. Hence, there exists a trade-off in choosing the value of p . We will outline the probability distribution of K analytically in Section IV.

In designing a RACS network, the underlying figure of merit is the reconstruction quality. The reconstruction error has to be within an acceptable range in order to obtain a certain reconstruction quality. In addition, among the set of design parameters that meet the required reconstruction quality, our goal is to choose the ones that minimize the average energy consumption of the sensor network. Fig. 7(a) shows the average normalized reconstruction error plotted versus the per-node sensing probability p , for randomly generated sparse data. The normalized error is defined as $\frac{\|\hat{\mathbf{u}}(n) - \mathbf{u}(n)\|_{\ell_2}}{\|\mathbf{u}(n)\|_{\ell_2}}$, where $\mathbf{u}(n)$ is the actual data and $\hat{\mathbf{u}}(n)$ is the recovered data. As noted in the figure, accurate reconstruction is possible for a range of values of p . Fig. 7(b) shows the corresponding normalized average energy consumption of the network versus p . In order to minimize the energy consumption of the network while maintaining the average quality of reconstruction, we choose the smallest value of p for which accurate reconstruction is possible.

IV. NETWORK DESIGN

In the R/D scheme of Section III-A, the number of correctly received packets at the FC, K , is equal to the number of sensor nodes selected for transmission, M . Thus, choosing $M = N_s$ provides a sufficient number of packets at the FC. In the R/R case however, M and K are both random variables. The fact that K is a random variable now implies that there

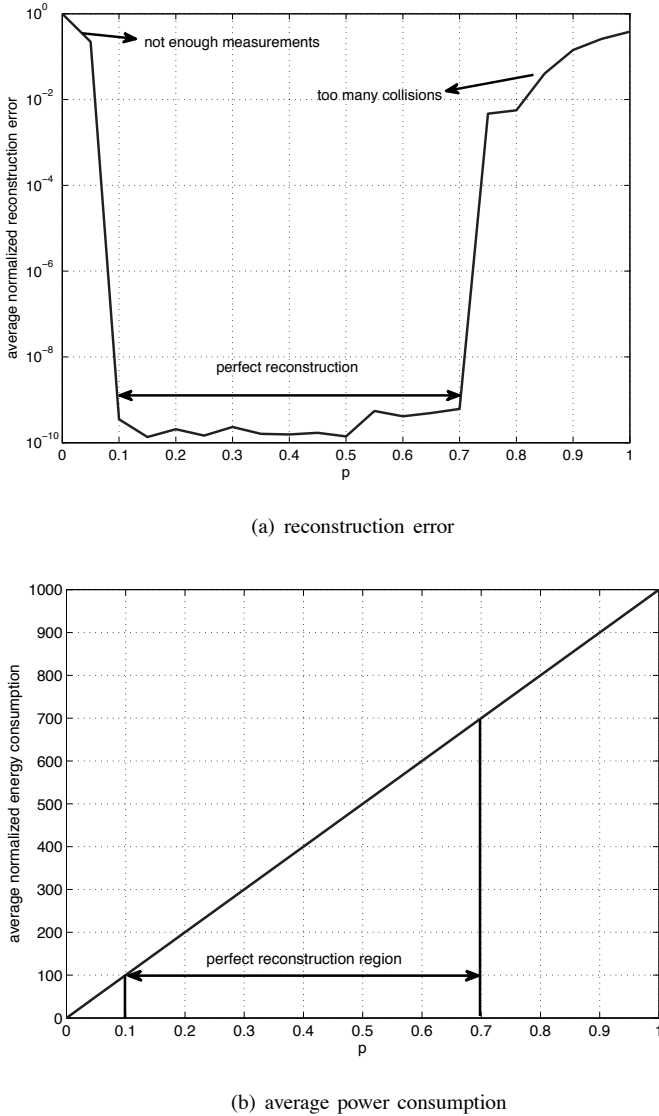


Fig. 7. Average normalized reconstruction error versus p and the corresponding energy consumption. Within the region where perfect reconstruction is achievable we choose the smallest p as this choice results in the least energy consumption.

can be no guarantee that K will be greater than N_s , i.e., obtaining a sufficient number of packets cannot be guaranteed. A probabilistic approach to the system design thus becomes necessary.

In what follows, we analyze the distribution function of the number of correct packets at the FC. We then study the conditions under which this random variable yields a sufficient number of measurements, $N_s = CS \log N$. These conditions in turn imply a per-node *sensing probability*, p .

A. Packet Reception Model

Given a sensing probability p , the average packet generation rate per node is given by $\lambda_1 = \frac{p}{T-T_p}$. Thus the aggregate arrival rate of packets at the FC is $\lambda = \frac{Np}{T-T_p}$. In order to determine the probability of collision, we note that the packet arrival process resembles a Poisson process. Accordingly, we model the probability of no collision as the probability that

no packet arrives in an interval of length $2T_p$,

$$\text{Prob}\{\text{no collision}\} = e^{-2\frac{NpT_p}{T-T_p}} \quad (7)$$

The probability that a packet is successfully received at the FC within a frame duration T is thus given by

$$q = pe^{-2\frac{NpT_p}{T-T_p}} \quad (8)$$

We now conjecture that K has a binomial distribution with parameter N and probability q , i.e.,

$$P_K(k) = \text{Prob}\{K=k\} = B(N, q) = \binom{N}{k} q^k (1-q)^{N-k} \quad (9)$$

where q is given by Eq. (8). To empirically verify the conjecture, we conduct simulation experiments. Fig. 8(a) shows the histogram of the number of correctly received packets obtained from simulation. In this figure, the $P_K(k)$ obtained from measurements is compared with that of the hypothesized model $B(N, q)$ where q is given by Eq. (8), and an estimated model $B(N, q_{est})$ where

$$q_{est} = \frac{\frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} k(i)}{N}$$

where $k(i)$ is the number of successfully received packets in the i -th simulation run and N_{sim} is the total number of runs. We note that q and q_{est} are very close, and that our conjecture for $P_K(k)$ provides a reasonable match with the simulated data. Fig. 8(b) shows the complementary cumulative distribution function, $Q_K(k) = \text{Prob}\{K \geq k\}$, for the simulated data, as well as for the model (9) and $B(N, q_{est})$. Again, we note a close match. Consequently, we will rely on the model (9) for system design.

B. Performance Requirement

In order to perform field reconstruction, the FC needs to collect at least N_s collision-free packets during one frame. However, since the packet arrival process is random, there is no guarantee that the FC will collect sufficiently many packets. Hence, we define the *probability of sufficient sensing* as the probability that the FC collects N_s or more correct packets, and we specify the performance requirement as the minimum probability of sufficient sensing, P_s . In other words, we ask that the FC collect at least N_s correct packets during one frame with probability P_s or higher. This condition can be expressed as

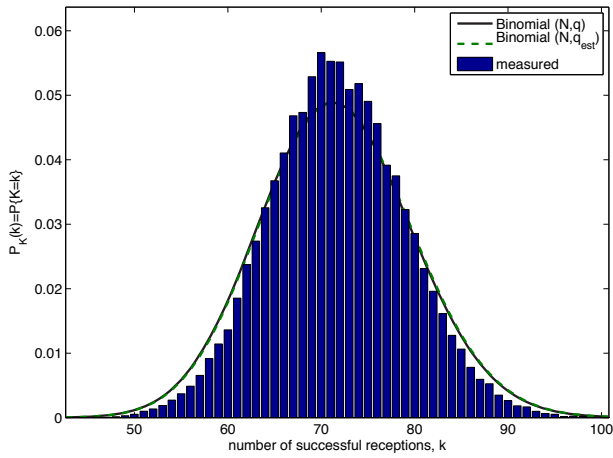
$$\text{Prob}\{K \geq N_s\} = Q_K(N_s) \geq P_s \quad (10)$$

where $Q_K(k)$ is the complementary cumulative distribution function. Using the model (9), we note that

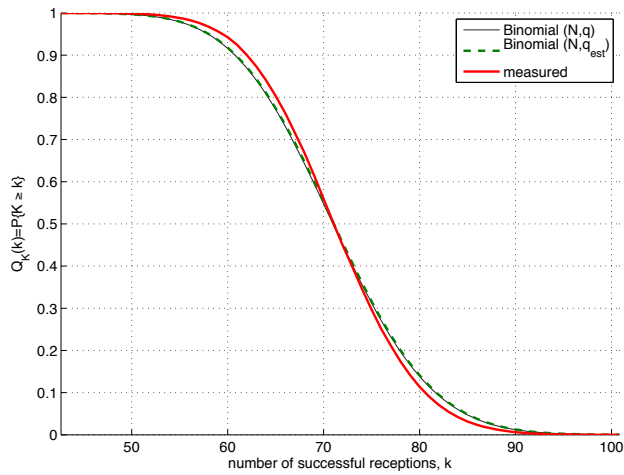
$$Q_K(N_s) \geq P_s \quad \text{for} \quad q \geq q_s \quad (11)$$

C. Design Objective

The design objective is to determine the *per-node sensing probability* p_s that ensures sufficient sensing. The first step in the design approach is to solve for q_s in Eq. (11). This can be done numerically for a given N_s and P_s . The procedure is illustrated in Fig. 9. Note that q_s depends only on P_s and N_s .



(a) Probability distribution



(b) Complementary cumulative distribution function

Fig. 8. The probability distribution $P_K(k)$ and the complementary cumulative distribution function $Q_K(k)$ for $N = 1000$, $T_p = 0.2$ s, $T = 120$ s, $p = 0.1$ and $N_{sim} = 10000$ simulation runs.

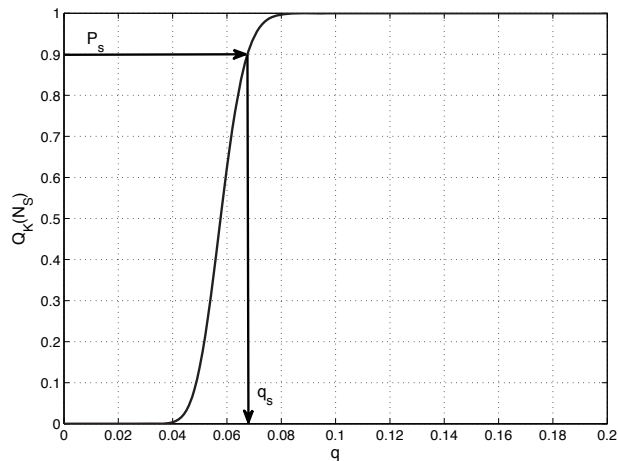


Fig. 9. Given a desired probability of sufficient sensing P_s , and a sufficient number of packets N_s , one can determine the corresponding value q_s . For example, $P_s = 0.9$ and $N_s = 57$ (see Fig. 5) yields $q_s = 0.068$.

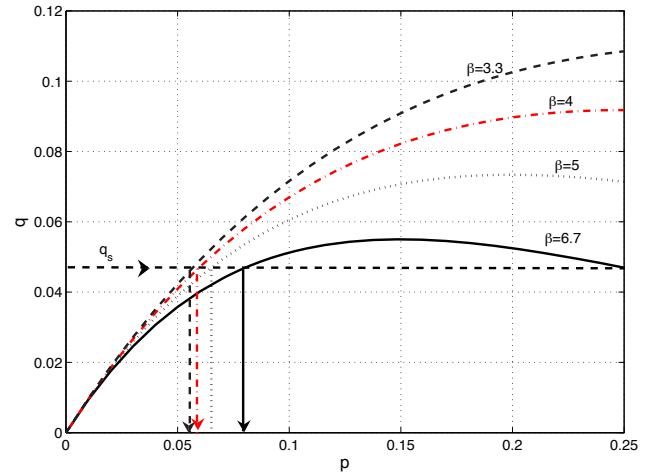


Fig. 10. The probability q of Eq. (8) plotted versus p for different values of $\beta = \frac{2NT_p}{T-T_p}$. For a given q_s a smaller β implies a smaller p_s .

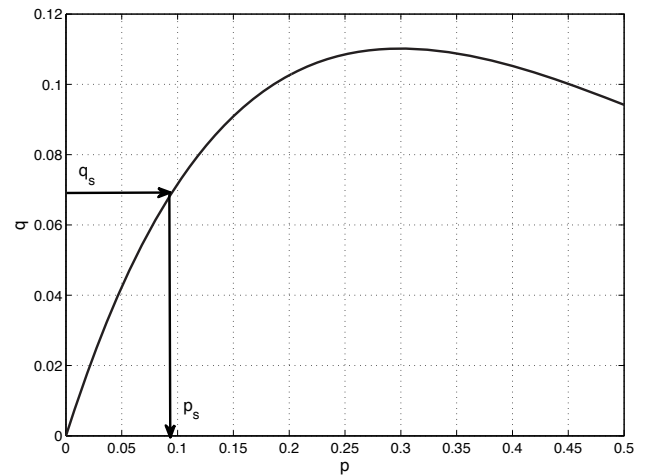


Fig. 11. Given q_s , the per-node sensing probability p_s is determined using the model (13). For example, $q_s = 0.068$ implies $p_s = 0.093$.

Now, using Eq. (8) we have that $q_s = p_s e^{-\beta p_s}$ where $\beta = \frac{2NT_p}{T-T_p}$. Given a specific value of q_s , this relationship is used to determine the underlying p_s and β . Our design approach is to minimize the energy consumption; hence, we want to identify that solution which yields the smallest p_s , since, as we will see in Section V, it yields the least energy consumption. Fig. 10 shows the plots of q versus p for various values of β . As seen in this figure, for a given q_s , the curve with a smaller β yields a smaller p_s . The smallest β corresponds to $T = T_{coh}$ and is determined as

$$\beta_{min} = \frac{2NT_p}{T_{coh} - T_p}. \quad (12)$$

Using this value, we find p_s as the solution of

$$q_s = p_s e^{-2 \frac{p_s N T_p}{T_{coh} - T_p}} \quad (13)$$

This procedure is illustrated in Fig. 11.

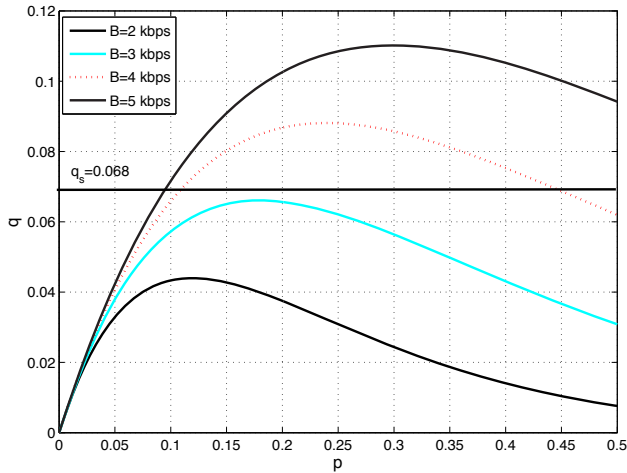


Fig. 12. The probability q versus p for $B = 2$ kbps – 5 kbps; System parameters are $N = 1000$, $T = 120$ s, $L = 1000$ bits. We note that in order for a solution to exist a minimum bandwidth is required.

The following example describes the complete design procedure. Let us assume a network of size $N = 1000$, measuring a phenomenon with sparsity $S = 10$ in the frequency domain. Fig. 5 implies that the required number of collision-free packets for perfect recovery is $N_s = 57$. For the given N_s and a desired sufficient sensing probability of say $P_s = 0.9$, Fig. 9 implies that $q_s = 0.068$. Let us assume that the packet duration is $T_p = 0.2$ s and that the coherence time of the process is $T_{coh} = 120$ s. Using Fig. 11, which shows q as a function of p based on Eq. (13), one can determine the per-node sensing probability $p_s = 0.093$.

Note that there is a possibility that q_s is too high for a solution p_s to exist. Fig. 12 shows q versus p for different bandwidths $B = 2$ kbps–5 kbps. For the given $q_s = 0.068$, we note from this figure that if $B = 2$ kbps or 3 kbps, there is no solution for p_s ; however, for $B = 4$ kbps or 5 kbps a solution exists. Thus, in order for a solution to exist, a minimum bandwidth is required. There is also a possibility of having two solutions for p_s ; if this occurs, we choose the smaller of the two as it corresponds to fewer sensors transmitting, which in turn translates into lower energy consumption as we will see in Section V.

The resulting complementary cumulative distribution function $Q_K(k)$ is shown in Fig. 13, which confirms that the choice of $p = 0.093$ satisfies Eq. (11), i.e., that the desired sufficient sensing is achieved.

In summary, we have a design approach that avails itself of a simplified model. For a given N , a coherence time T_{coh} , and a packet duration T_p , the model (13) is used to determine the per-node sensing probability p such that the desired probability of sufficient sensing P_s is met.

V. PERFORMANCE ANALYSIS

In this section, we compare the performance of the RACS scheme with that of a conventional network (Section II-A). In an underwater deployment, network lifetime is of utmost importance since re-charging batteries is difficult. Energy

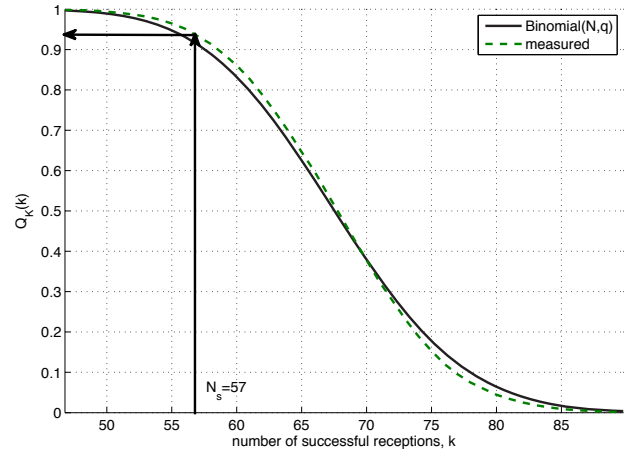


Fig. 13. Complementary cumulative distribution function $Q_K(k)$ plotted for $p_s = 0.093$ confirms that the desired sensing probability is achieved, i.e., $Q_K(N_s) \geq P_s$ for $N_s = 57$ and $P_s = 0.9$.

per successfully delivered bit of information thus naturally emerges as a figure of merit for system performance. In light of a sensor network based on compressed sensing, we define a figure of merit as the total average energy required for one field reconstruction. One of the performance measures that we consider is thus the average energy consumption of the network needed to sense a given area. Since bandwidth is severely limited in an underwater acoustic network, another measure of performance is the minimum bandwidth required. In what follows we analytically derive these performance metrics for the two schemes based on compressed sensing (R/D and R/R) and compare the results to those of a conventional system.

If by P_T we denote the per-node transmit power, the consumed energy per node is given by

$$E_0 = P_T T_p$$

where T_p is the packet duration, i.e., the time during which a node is active. The total consumed energy in the conventional TDMA network is given by

$$E_{conv} = N E_0 = \frac{A}{d^2} E_0 \quad (14)$$

where, $A = N d^2$ is the coverage area of the network. The frame duration in a conventional network is $T = N T_p \leq T_{coh}$. Hence, noting that $T_p = L/B_{conv}$, the bandwidth requirement is given by

$$B_{conv} \geq \frac{N L}{T_{coh}} \quad (15)$$

For the R/D scheme, the total energy required for one field reconstruction is given by

$$E_{R/D} = M E_0 = C S \log(N) E_0 = C S \log\left(\frac{A}{d^2}\right) E_0 \quad (16)$$

The frame duration in R/D is given by $T = 2\tau_{max} + M T_p$. Moreover, the network needs updated data every T_{coh} ; therefore, $M T_p + 2\tau_{max} \leq T_{coh}$. This condition results in a

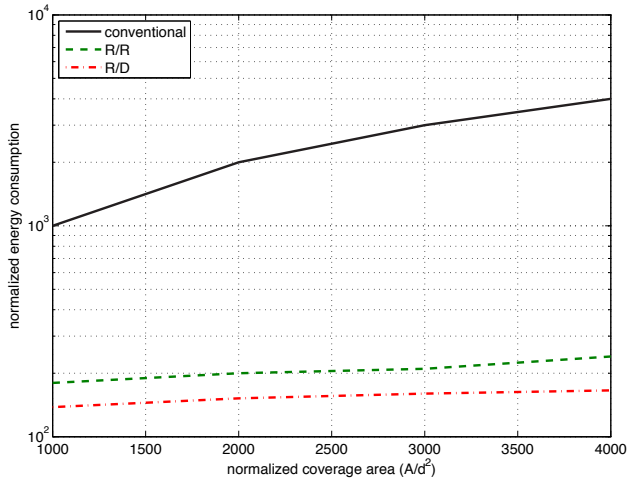


Fig. 14. Total network energy consumption, normalized with respect to the energy E_0 needed by one node to transmit one data packet, plotted versus normalized coverage area A/d^2 . System parameters are $T = 800$ s and $T_p = 0.2$ s. The sparsity level $S = 10$ is assumed to remain fixed.

minimum bandwidth requirement of

$$B_{R/D} \geq \frac{LCS \log(N)}{T_{coh} - 2\tau_{max}} \quad (17)$$

Finally, in the R/R scheme, the average consumed energy for one field reconstruction is

$$E_{R/R} = p_s N E_0 = p_s \frac{A}{d^2} E_0 \quad (18)$$

where $p_s N$ is the average number of nodes that transmit in one frame. Note that p_s in the above expression is implicitly dependent on N , through the design procedure outlined in Section IV.

The observations that we made from Fig. 12 imply that in order for a set of design parameters to satisfy the sufficient sensing condition, a minimum bandwidth is required. The minimum required bandwidth is obtained by identifying the maximum of q , i.e., by taking the derivative of q with respect to p and setting it equal to zero. Let us assume that $\beta_{min} > 1$, which is the case of our interest.³ The maximum value of q is then obtained as $q_{max} = 1/e\beta_{min}$. In order for sufficient sensing to occur, we need to have $q_s \leq q_{max}$, which results in the minimum bandwidth requirement as

$$B_{R/R} \geq (2eNq_s + 1) \frac{L}{T_{coh}}$$

Fig. 14 shows the energy consumption normalized with respect to E_0 , versus the normalized coverage area $\frac{A}{d^2}$ for the three schemes above. Note that $T_{coh} = 800$ s results in a maximum of $N = 4000$ nodes in a conventional network. As seen in the figure, for the same coverage area, RACS offers energy savings of an order of magnitude compared to the conventional scheme. By reducing the energy consumption, RACS extends the life-time of the sensor network. Fig. 15 shows the minimum bandwidth required, versus the size of the network N . For the same network size, RACS

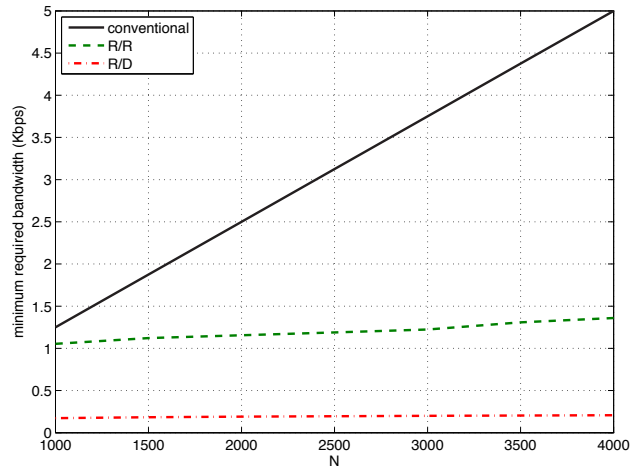


Fig. 15. Minimum required bandwidth versus the size of the network N for the conventional, R/D and R/R schemes. System parameters are $S = 10$, $T = 800$ s, $L = 1000$ bits and $\tau_{max} = 0.33$ s.

requires lower bandwidth compared to the conventional scheme. For example, in a network of $N = 2500$ nodes, R/R scheme requires only a bandwidth of 1.2 kbps, whereas the conventional network requires 3.1 kbps. The savings in bandwidth are a significant feature from the viewpoint of acoustic communications.

VI. REAL DATA EXAMPLE

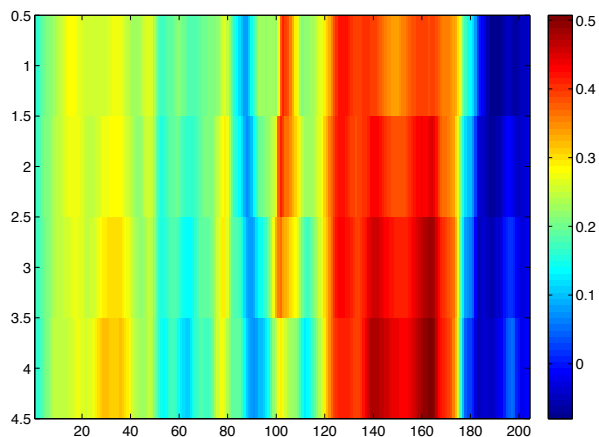
To visually illustrate the field recovery process, we employ RACS to sense a real field. We consider zonal current data collected at Southern California bight at 3 GMT on August 19, 2010 at latitudes $[32.5^\circ, 32.58^\circ]$ and longitudes $[238.8^\circ, 243^\circ]$. This data set is accessible at <http://ouocean.jpl.nasa.gov> and is shown in Fig. 16(a). We note that almost 99% of the energy of the signal is contained in $S = 17$ Fourier coefficients. For $N_s = 285$, assuming a sufficient sensing probability $P_s = 0.9$, a desired updating interval $T = 1000$ s, and a packet duration $T_p = 0.2$ s, following the design approach of Section IV, the per-node sensing probability is determined to be $p_s = 0.439$. Fig. 16(b) shows the map of the field recovered using RACS with this probability.⁴ In this example recovery has been achieved consuming less than half the energy of a conventional network, $E_{R/R}/E_{conv} \approx 0.4$. In order to study the error behavior of the scheme as a function of sensing probability, Fig. 17 shows the normalized reconstruction error versus the per-node sensing probability p . Saturation region is not present in this figure, as it is in Fig. 7(a), since the updating interval T is long enough to keep the number of packet collisions from dominating the error.

VII. CONCLUSION

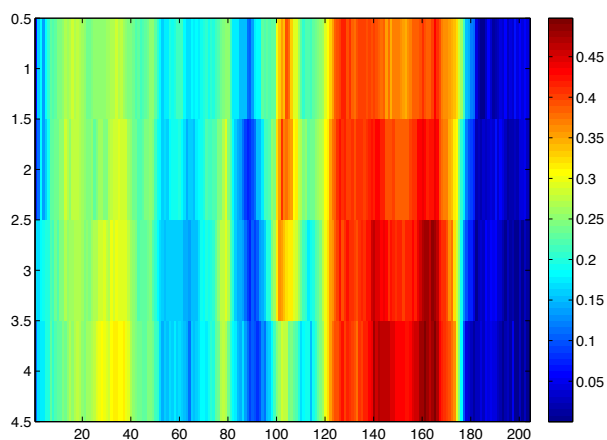
We proposed a networking scheme that combines the concepts of random channel access and compressed sensing to achieve energy and bandwidth efficiency. This scheme is

³In the case that $\beta_{min} \leq 1$, $q_{max} = e^{-\beta_{min}}$ and the analysis follows similarly.

⁴For reconstruction, we used CVX, a package for specifying and solving convex programs [19].



(a) The map of the original zonal current field.



(b) The map of the reconstructed zonal current field.

Fig. 16. The sensing field is recovered employing RACS with $p_s = 0.439$, $T = 1000$ s, and $T_p = 0.2$ s.

suitable for large networks, deployed for long-term monitoring of slowly varying phenomena. The underlying condition is that the measured physical phenomenon has compressible (sparse) representation in the frequency domain, which is the case in many natural fields. The proposed method is completely decentralized, i.e., sensor nodes act independently without the need for coordination with each other or with the FC. The only downlink feedback needed is an occasional synchronization beacon.

To account for the random packet loss caused by collisions, it becomes necessary to employ a probabilistic approach in the system design, thus we introduced the concept of sufficient sensing probability. With this probability, which is the system design target, the FC is guaranteed to acquire a sufficient number of observations per frame to reconstruct the measured field. A desired probability of sufficient sensing then points to the necessary per-node sensing probability. The performance of RACS was assessed analytically in terms of the energy consumption and bandwidth requirement, demonstrating substantial savings over a conventional scheme based on deterministic sensing and access.

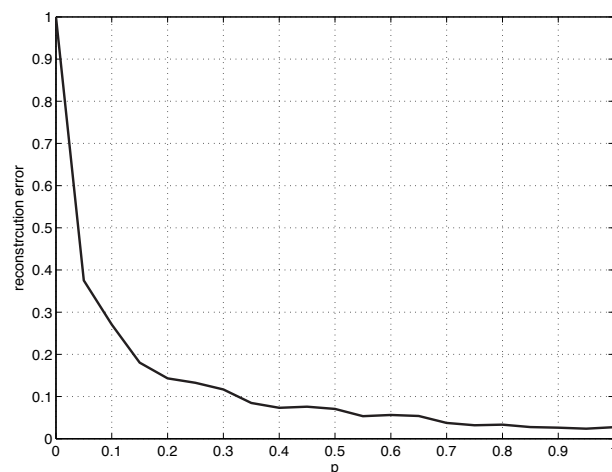


Fig. 17. Normalized reconstruction error versus p for the zonal current data of Fig. 16.

REFERENCES

- [1] J. Heidemann, W. Ye, J. Wills, A. Syed, and Y. Li, "Research challenges and applications for underwater sensor networking," in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2006, pp. 228–235.
- [2] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Networks (Elsevier)*, 2005.
- [3] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, pp. 21–30, March 2008.
- [4] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, pp. 118–121, July 2007.
- [5] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *5th Int. Conf. Information Processing in Sensor Networks (IPSN'06)*, April 2006, pp. 134–142.
- [6] W. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *4th Int. Conf. Information Processing in Sensor Networks (IPSN'05)*, April 2005, pp. 332–339.
- [7] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3629–3653, October 2007.
- [8] Y. Mostofi and P. Sen, "Compressive cooperative sensing and mapping in mobile networks," in *Proc. American Control Conference (ACC)*, St. Louis, Missouri, June 2009.
- [9] J. Wu, "Ultra-low power compressive wireless sensing for distributed wireless networks," in *Military Communications Conference (MILCOM)*, Oct 2009, pp. 1–7.
- [10] C. T. Chou, R. Rana, and W. Hu, "Energy efficient information collection in wireless sensor networks using adaptive compressive sensing," in *IEEE 34th Conference on Local Computer Networks (LCN)*, Oct 2009, pp. 443 – 450.
- [11] J. Meng, H. Li, and Z. Han, "Sparse event detection in wireless sensor networks using compressive sensing," in *43rd Annual Conference on Information Sciences and Systems (CISS)*, March 2009, pp. 181 – 185.
- [12] Q. Ling and Z. Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3816–3827, July 2010.
- [13] A. Griffin and P. Tsakalides, "Compressed sensing of audio signals using multiple sensors," in *Proc. 16th European Signal Processing Conference (EUSIPCO '08)*, August 2008.
- [14] S. Hu and J. Tan, "Compressive mobile sensing for robotic mapping," in *IEEE International Conference on Automation Science and Engineering*, Aug 2008, pp. 139 – 144.
- [15] S. Poduri, G. Marcotte, and G. S. Sukhatme, "Compressive sensing based lightweight sampling for large robot groups." [Online]. Available: http://cres.usc.edu/pubdb_html/files_upload/647.pdf
- [16] A. K. Fletcher, S. Rangan, and V. K. Goyal, "On-off random access channels: A compressed sensing framework," *Submitted to IEEE Trans. Inf. Theory*, arXiv:0903.1022v2 [cs.IT].

- [17] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, March 2006.
- [18] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE, special issue, Applications of sparse representation and compressive sensing*, vol. 98, no. 6, pp. 948–958, June 2010.
- [19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming." [Online]. Available: <http://cvxr.com/cvx/>



Maryam Fazel received her BS degree from Sharif University, Iran, and her MS and PhD degrees from Stanford University in 2002. She was a postdoctoral scholar and later a Research Scientist in the Control and Dynamical Systems Department at Caltech until 2007. She is currently an assistant professor in the Electrical Engineering Department at the University of Washington, Seattle, with adjunct appointments in Mathematics and in Computer Science and Engineering. She is the recipient of a 2009 NSF CAREER Award and an Outstanding Teaching Award

at the University of Washington.



Fatemeh Fazel received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran in 2000 and the M.S. degree from University of Southern California in 2002. She received her Ph.D degree from the Department of Electrical Engineering and Computer Science at University of California, Irvine. She is currently a Postdoctoral Associate with the Electrical and Computer Engineering Department at Northeastern University, Boston. Her research interests are in wireless communications, with a focus on multiple-

input multiple-output (MIMO) systems, space-time coding, and energy-efficient sensor networks.



Milica Stojanovic (Sm'08, F'10) graduated from the University of Belgrade, Serbia, in 1988, and received the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston, MA, in 1991 and 1993. After a number of years with the Massachusetts Institute of Technology, where she was a Principal Scientist, she joined the faculty of Electrical and Computer Engineering Department at Northeastern University in 2008. She is also a Guest Investigator at the Woods Hole Oceanographic Institution, and a Visiting Scientist at MIT. Her

research interests include digital communications theory, statistical signal processing and wireless networks, and their applications to underwater acoustic communication systems. Milica is an Associate Editor for the IEEE Journal of Oceanic Engineering and the IEEE Transactions on Signal processing.