Encoded Sensing for Energy Efficient Wireless Networks

MILEN NIKOLOV, Cornell University
ZYGMUNT J. HAAS, Cornell University

Energy efficient communication is a fundamental design problem in battery powered wireless networks, significantly affecting network performance and lifetime. In this article, we describe *encoded sensing* – a novel approach for collaborative encoding and transmission of data – that drastically reduces communication energy expenditure in the network. For example, encoded sensing achieves comparable or better energy efficiency compared to distributed cooperative transmit beamforming, the latter achieving well-known significant SNR gains. However, unlike distributed transmit beamforming, encoded sensing does not reside at the physical layer and does not require synchronization in frequency and phase across transmitting nodes. There is no feedback (e.g. pilot signals, bit messages, etc.) required from the sink. Encoded sensing maps every distinct message x to a unique subset of source nodes. Given a message x to be transmitted, each source node j determines whether x is mapped to j. Node j transmits a binary signal to the sink only if x is mapped to j. Collecting sufficiently large number of such signals, the sink can determine the actual message x. In the context of wireless sensor networks, the source nodes correspond to a group of sensor nodes. The entire range of possible measurements is quantized into individual intervals, and each interval corresponds to a message x assigned to distinct subset of nodes. The sink can reconstruct any x based on the received signals, and hence estimate the sensor nodes' measurement.

Categories and Subject Descriptors: **C.2.1 [Computer-Communication Networks]**: Network Architecture and Design - Network communication, Wireless communication, Sensor Networks;

General Terms: Design, Performance, Algorithms

Additional Key Words and Phrases: energy efficiency, distributed sensing, cooperative transmissions, spatial correlation, encoded sensing, compressive sensing

ACM Reference Format:

Milen Nikolov and Zygmunt J. Haas, 2014. Encoded Sensing for Energy Efficient Wireless Networks. *ACM Trans. Sensor Networks* x, x, Article xx (2014), x pages. DOI:xxx.xxx

1. INTRODUCTION

Communicating data is among the most energy expensive routines across different types of wireless networks. For instance, receiving and transmitting data in wireless sensor networks (WSN) consisting of Mica2 nodes running TinyDB applications constitutes about 59% of the total energy consumption [Shnayder et al 2004]. In various applications running on similar WSN systems, only transmissions account for up to 50-65% of energy expenditure, depending on radio throughput [Landsiedel 2005 et al.]. The trend is similar for small WiFi connected devices, where radio operation may claim 70% of the total energy budget [Pering et al. 2006]. Reducing the amount of data transmitted and/or energy consumed per transmission could lead to longer network lifetimes and significantly impact network functions.

This work is supported by the National Science Foundation, under grants CNS-1040689, ECCS-1308208.

Author's addresses: M. Nikolov, Electrical and Computer Engineering, Cornell University; and Z. J. Haas, Computer Science Department, University of Texas at Dallas (current address).

DOI:xxx.xxx.xxx

We focus on particular communication scenarios in the context of wireless networks' energy efficiency. In many cases of WSN deployments, for instance, clusters of spatially proximate nodes sense and potentially transmit very highly correlated, almost identical values to the sink [Vuran and Akyildiz 2006; Bouabdallah et al. 2008]. The scenario underlying the operation of cooperative transmission networks (e.g. [Khandani et al. 2007; Elhawary and Haas 2011]) and the related distributed transmit beamforming physical layer protocols [Mudumdai et al. 2010] is analogous: a group of nodes transmits the same message in a carefully coordinated fashion. In both of the above scenarios, a group of source nodes have access to the same or highly correlated pieces of data that need to be sent to a common sink.

We introduce a novel scheme – *encoded sensing* (ES) – that substantially reduces the energy required for transmission of data in various types of wireless networks, where these settings hold. In this section, we first provide a brief description of encoded sensing's basic idea in the context of a WSN, since ES is most intuitive in that setting.

1.1 A Wireless Sensor Network Example

Suppose a set N of sensor nodes is deployed within an event area A, according to some spatial distribution Γ. The event source is continuous and governed by an apriori known spatial statistics. The source statistics model could be dictated by the phenomenon nature, or learnt via training. Although measurements at different positions in A are frequently assumed i.i.d in the research literature, it is more realistic to assume that measurements are instead spatially-correlated and dependent on the distance between sensors [Varshney and van de Ven 2013; Vuran and Akyildiz 2006]. The specifics of the spatial-correlation model are described in Section 4, where we discuss further the application of ES to WSN.

We assume time is discretized in slots. The sensors sample A, and a set of active nodes report measurements to the sink during each timeslot. Nodes are coarse-grained synchronized at a resolution of a timeslot. We assume that nodes transmit over multiple access AWGN channel at the same average power per bit, and that the sink B is within the transmission radius r of the nodes. The latter assumption is made for simplicity of presentation. Extending ES to multihop scenarios is discussed in Section 8.

Finally, in our model we account for imprecisions in sensors' electronics, introducing noise in each node's measurement. The statistical model for this instrumentation noise is provided in Section 5.

WSN problem statement: under the above assumptions, we are seeking a scheme that *minimizes the energy* spent in communicating measurements to the sink, while simultaneously *guaranteeing a level of accuracy* in estimating the value of the measured source. This minimum level of accuracy is prescribed to the system by QoS constraints. Sink's estimate accuracy is modeled via a standard Minimum Square Error (MSE) distortion metric fully specified in Section 4.

1.2 Encoded Sensing Solution Intuition

Suppose the continuous source (e.g. a phenomenon or modality such as temperature) is quantized into a set of short intervals. The intervals are assigned to an index $I = 1, 2, \ldots, M$. For simplicity assume that at a given timeslot, all sensor nodes in a group $G \subset N$ measure phenomenon values falling in the same interval $x \in I$. We relax this assumption later. The goal of group G is to transmit measurement x to the sink, reliably and with minimum energy.

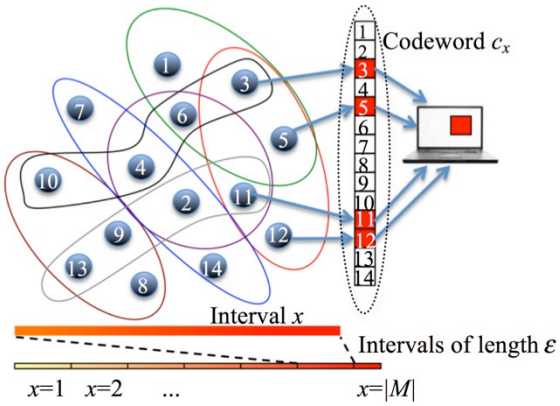

Figure 1. All nodes in the group G sense a value in the same small interval x (in solid red gradient). Only the 4 nodes (circled in red line) that are assigned to x transmit to the sink; the rest 10 nodes are silent (white blocks). The sink identifies the 4 transmitting nodes (red blocks) and recovers the interval x.

Per encoded sensing, the nodes in G are labeled with IDs: 0, 1, …, $|G| - 1$ as shown in Figure 1. Every interval $x \in I$ is assigned to a distinct set of nodes $A_x \subset G$. Each of the nodes in $A_x \subset G$ assigned to x transmits a signal "1" to the sink. The nodes in G/A_x remain silent. Each transmitted signal "1" carries a "signature" of its source node $j \in A_x$. Receiving the signed signals, the sink can identify all j's in A_x. Assuming it knows the assignment of sets A_x to intervals x, $\forall x \in I$, the sink can recover x. Since x is short, it is an accurate estimate of the value measured by the nodes in G.

In essence, the sink receives a "collaborative codeword" c_x of "1"s that *encodes* the interval x. Upon receiving a valid codeword c_x, the sink can decode it and obtain x. The number of available codewords depends on the cardinality of A_x and the size of the group G. For instance, given a group of size $|G| = 14$ and equal codeword lengths $|A_x| = |c_x| = 4$ one could generate ${}^{14}C_4 = 1001$ different codewords; if $|G| = 17$, there are more than double, ${}^{17}C_4 = 2380$, codewords. In the limit case, $|A_x| = |c_x| = 1$: given $|G| = M$ the measurement can be encoded over a single bit sent by a single source node j, given j's "signature"/code. Notice that this limit case is somewhat analogous to utilizing M-ary orthogonal codes (e.g. PPM) to achieve reliable communication at the minimum possible energy per bit E_b as M increases [Tse and Viswanath 2005]. There, the total power is spread over a large time interval. As we will see, in the encoded sensing limit case, where $|G| = M$, the energy per bit is spread over nodes' orthogonal code "signatures".

Although most energy efficient, the limit case is not practical, as M could be on the order of tens of thousands and larger. However, we observe *a trade-off between the number of nodes in the network and network's energy efficiency*. Given the number of nodes $|G|$ as a system constraint/parameter, we find the minimum codeword length $|A_x| = |c_x| = K$, so that ${}^{|G|}C_{|K|} \geq M$, thus *minimizing energy consumption* for communication in group G.

We show that a WSN of N nodes can be partitioned in multiple groups of nodes, so that encoded sensing employed locally in each group maximizes network-wide energy efficiency, while satisfying quality of service constraints.

In the above description, we made a few implicit assumptions regarding ES operation. In what follows we show how these assumptions can be met or relaxed; we study and quantify the benefits of employing ES; and we describe how ES may be generalized outside the scenario of WSNs.

Encoded sensing *collaborative codes*:

- First, we assumed that the sink and nodes "know" of a mapping where any message x is assigned to a distinct set of nodes $A_x \subset G$. In Sections 2 and 3, we discuss decentralized algorithms for finding such mappings and constructing novel *collaborative codes* for energy efficient communication.
- In the case of WSN, we also assumed nodes in the same group G measure values falling in the same interval $x \in I$. In practice this may not always be the case due to measurement imprecisions, for instance. In Section 3 (Theorems 2 – 4), we propose and analyze a novel Minimal Distance Combinatorial Encoding allowing nodes within a group to erroneously measure values falling in an interval x' differing from the true interval x. The nodes in the group still can transmit the correct collaborative codeword for interval x. The same ES collaborative codes can be used to encode general data messages instead of intervals x.

Encoded sensing signatures and *sparse DSSS*:

- We assumed that the sink can identify each transmitting node's "signature" based on a single bit. In Section 4, we discuss a node identification technique at the sink based on standard direct sequence spread spectrum (DSSS) PN sequences. We show how a certain ES *sparseness property* provides insight for a novel, significantly reduced, receiver circuitry (and cost) for DSSS devices based on compressive sensing sparse *support recovery*. Using the latter technique, we achieve node identification w.h.p. (Theorem 5).

Support recovery is sufficient for ES operation, and we do not require detection of the received symbol sign (i.e. detecting if the received bit is 0 or 1). However, the described *sparse* DSSS receiver design is potentially applicable in other, non-ES, DSSS deployment scenarios, too. Hence, for completion, we discuss a MMSE-based scheme for detecting the received symbol sign at the sparse DSSS receiver. We provide bounds for the conventional DSSS receiver design's symbol detection probability of error and compare them with the respective performance of the suggested sparse DSSS MMSE-based scheme.

Energy efficiency and distortion:

- In Section 5, we study ES in a WSN application. We show that ES achieves comparable energy gains to state-of-the-art cooperative distributed transmit beamforming schemes [Mudumdai et al. 2010]. Unlike beamforming, ES does not reside on the physical layer and does not require fine synchronization across transmitting nodes, the latter often being impractical. Also, we show that ES is at least 2 times more energy efficient than non-cooperative [Vuran and Akyildiz 2006] and similar duty cycles schemes, where only one of the nodes in a group G transmits message x. The analytical rationale for the latter is given in Section 6.
- The energy efficiency of the ES scheme is achieved while matching and guaranteeing optimal level of estimate distortion at the sink, within WSN's QoS constraints. To obtain optimal distortion, in Section 6, we employ a vector quantization scheme for determining the sites of *representative groups* of nodes with highly correlated measurements. The scheme can be utilized orthogonal to ES to reduce distortion of measurement either in the context of topology control by placing nodes at the *representative groups* sites; or, for forming groups of highly correlated nodes, given network nodes have already been placed.

Section 7 provides detailed overview of related works; and Section 8 concludes the paper.

2. ENCODED SENSING

In this section, we describe encoded sensing's stages. For clarity, we consider a single group of nodes $G \subset N$. The nodes in G need to send message x to a single sink.

Encoded sensing consists of four major stages: *encoding*, *assignment*, *transmission*, and *decoding*.

- At the *encoding stage*, each node in G locally encodes x into binary codeword c_x of length $|G|$, according to a common decentralized coding algorithm run at all nodes. We discuss such codes, satisfying different properties, below. Let $c_x(j)$, $0 \le j \le |G| - 1$, denote the j-th most significant bit in c_x. Each node forms the set $A_x = \{j : c_x(j) = 1\}$.
- At the *assignment stage* each node j determines whether it is in the set A_x:

$$D_j = \begin{cases} 1, if\ j \in A_x \\ 0, otherwise \end{cases} \tag{2.1}$$

Node j *assigns* itself to x, iff $D_j = 1$.

- Next, at the *transmission stage*, for each node $j \in G$, if $D_j = 1$, j transmits a single bit "1_j" with j's signature, so that upon receiving "1_j" the sink can detect the sender node j. Otherwise, if $D_j = 0$, node j remains silent.
- Finally, at the *decoding stage*, the sink receives the set $\{1_j : j \in A_x\}$. The sink determines the identities of the transmitting nodes based on their signatures and recovers A_x. $c_x(j)$, $0 \le j \le |G| - 1$, is set to 1 if $j \in A_x$ and to 0 otherwise. The sink runs a decoding algorithm on c_x to obtain message x. Notice that codeword c_x received at the sink has been collaboratively constructed by the nodes of group G. We refer to such codewords as *collaborative codewords*.

Minimizing $|A_x|$ reduces the number of nodes that transmit signals, conveying message x to the sink, in turn minimizing communication energy. Notice that $|A_x| = w(c_x)$, where $w(c_x)$ is the weight of codeword c_x[1]. The objective of a collaborative code would be to have a sufficiently large range C of codewords with length $|G|$, so that $C \ge M$. Simultaneously, $w(c_x)$ should be as small as possible for all x. Since we need to encode each x with a distinct codeword c_x, the subsets A_x of G have to be distinct. To illustrate the mechanism of encoded sensing, in the next three subsections we discuss a simple combinatorial code, along with respective encoding/decoding algorithms, that achieves optimal performance for sufficient number of nodes in a group G. However, this combinatorial code may not be practical in certain scenarios. In Section 3, we propose a more sophisticated code preserving the performance of the simple combinatorial code, while avoiding its limitations.

2.1 Encoding and Assignment

To reduce the number of nodes participating at the encoding and transmission stages and achieve optimal energy efficiency, we minimize $|A_x| = w(c_x)$, given M and $|G|$. We assume M and $|G|$ are known locally at all nodes. The constraint here is that all M messages have to be assigned to M distinct codewords. Assuming equiprobable messages at present, we have $|A_x| = w(c_x) = K$ for all x. Then, notice that the number of possible distinct codewords is ${}^{|G|}C_K$, and we require ${}^{|G|}C_K \ge M$. For a given fixed group size $|G|$, we need to find

$$K^* = \min_{1 \le K \le |G|/2} K \quad s.t. \quad {}^{|G|}C_K \ge M \tag{2.2}$$

[1] The weight of a binary codeword as usual equals the number of 1's in it.

ALGORITHM 1 Assign *msg x* to a K^*- subset A_x of G

Input: *value of binary x in decimal,* K^*
Output: $A_x = \{n_{K^*}, n_{K^*-1}, \ldots, n_1\}$
Algorithm:
1: $A_x \leftarrow \{\varnothing\}$
2: $m \leftarrow K^*$
{At each iteration add a node ID to the assignment A_x of nodes to measurement x}
3: **while** $m \geq 1$ **do**
4: $n_m \leftarrow$ maximum integer such that ${}^{n_m}C_m \leq x$
5: $x \leftarrow x - {}^{n_m}C_m$
6: $m \leftarrow m - 1$
7: $A_x \leftarrow A_x \cup n_m$

and then construct a code, where $w(c_x) = K^*$ for all x. The range C of this code is ${}^{|G|}C_K$. For any given $|G|$ and M, K^* is easily computed numerically at each node. Picking K^* for the size of the subset of nodes A_x assigned to each message x guarantees that the number of nodes participating in the encoding of x is the minimum possible.

We index each message x so that: $0 \leq x \leq M - 1$. Given message x, in the *encoding step*, each node runs **Algorithm 1** locally. **Algorithm 1** performs basic combination unranking, hence we dub this simple code *combinatorial encoded sensing* (*ES-C*). For each distinct input pair (x, K^*) it outputs a distinct set A_x of K^* integers taking values from 0 to at most $|G| - 1$. Stated differently, each distinct message x is mapped to a set A_x containing the IDs of the nodes in G responsible for x. A_x is the minimal possible such subset under the constraints of (2.2).

At the *assignment stage*, each node $j \in G$ knows the set A_x assigned to the message x, and j knows its own ID. If j's ID is in A_x, j sets $D_j = 1$ in (2.1). Notice that without message passing between each other, the nodes in G have at this point locally and collectively encoded the message x to a codeword c_x, where $c_x(j)$, $0 \leq j \leq |G| - 1$, is set to 1 if $j \in A_x$ and to 0 otherwise.

2.2 Transmission and Optimality

At the transmission stage, the nodes in A_x can transmit their respective single-bit signals "1$_j$". The sink can only recover the codeword c_x if nodes' *transmission*s indicate nodes' identities as well. This information however is often implicit in many standard physical layer protocols employing variants of spread spectrum wave signatures embedded in each node's signal. For instance, the physical layer of IEEE 802.15.4 utilizes Direct Sequence Spread Spectrum (DSSS). Similarly practical and efficient DSSS architectures for low-powered WSNs have also been discussed in [Chien et al. 2001]. Long Term Evolution 3GPP networks are another example where each node is assigned a distinct, in this case, Zadoff-Chu sequence [Song and Shen 2010]. In general, some form of a signature identifying nodes' signals at the receiver is a common feature in many "off-the-shelf" multi-user systems [Viterbi 1995; Madhow and Pursley 1993]. In the above instances, each node's signal (bit) is multiplied by the node's unique sequence and transmitted. The received signal is correlated at the sink, via a bank of matched filters, with the set of available sequences' waveform signatures. The sink can obtain the identity of the transmitting nodes in this sequence acquisition stage. This is the only assumption ES makes regarding the physical layer. Section 4 expounds the properties of our DSSS design.

Notice that given $|G| \geq M$ each message x can be encoded over a single bit, $w(c_x) = K^* = 1$, sent by a single source node j, where the bit is spread over j's DSSS spreading sequence. This limit case is similar to utilizing M-ary orthogonal codes (e.g. PPM) to achieve reliable communication over AWGN channels at the minimum possible energy per bit E_b, as M increases [Tse and Viswanath 2005]. There, the total power is spread over a large time interval to achieve the Shannon limit of $E_b/N_0 = -1.6$dB. In our case, DSSS spreading (e.g. PN, orthogonal, etc.) sequences have very low cross-correlation; ideally they are orthogonal to each other. In the *ES-C* limit case, where $|G| \geq M$ and $w(c_x) = 1$, the energy per the single bit transmitted over AWGN is spread over the spreading sequence of the single channel user j. Node j's sequence is orthogonal to the spreading sequences of the nodes in $G/\{j\}$ and ES-C operates at the optimum energy per bit level in this case.

2.3 Decoding

After receiving and correlating nodes' transmissions according to the physical layer properties, the sink identifies the nodes j in A_x each of which has transmitted only a single bit "1$_j$". Assuming the system operates under DSSS acquisition capacity, so that w.h.p. there are no errors in transmission, the sink can *decode* the codeword c_x. The following *Theorem* 1 ensures that the IDs of the nodes in A_x available at the sink are sufficient to recover the index of the message x.

THEOREM 1**:** *For every number* $x \in \mathbb{N}$, $\exists$ *a unique set* $\{n_m, n_{m-1}, \ldots, n_1\}$, $n_i \in \mathbb{N}$ *and* $n_m > n_{m-1} > \ldots > n_1 \geq 0$, *such that for any* $m \in \mathbb{N}$, *where* $m \leq n_m$

$$x = {}^{n_m}C_m + {}^{n_{m-1}}C_{m-1} + \ldots + {}^{n_1}C_1 = \sum_{i=1}^{m} {}^{n_i}C_i \tag{2.3}$$

PROOF: See [Lemher 1964]. □

Notice that Theorem 1 guarantees existence as well uniqueness. Namely, after ordering the set of IDs in A_x into the sequence $(n_m, n_{m-1}, \ldots, n_1)$ and having $m = K^* \leq n_m$ we obtain that

$$x = \sum_{i=1}^{m} {}^{n_i}C_i \tag{2.4}$$

This is the only value x that can be assigned to the set A_x. The ES decoding process is simply a combination ranking of A_x.

2.4 Example of a Basic ES-C Run

Suppose a group of $|G| = 6$ nodes is required to send, among a set of $M = 16$ messages, the message $x = 0101_2 = 5_{10}$. It is easy to compute that $K^* = 3$ from (2.2). (By inspection, if $K^* = 2$, ${}^6C_2 = 15 < M = 16$ and (2.2) is not satisfied.) Then, at the *encoding step*, all 6 nodes run **Algorithm 1** with input pair $x = 5$ and $K^* = 3$. The output of **Algorithm 1** is the set $A_5 = \{4, 2, 0\}$. Nodes in G with IDs 4, 2 and 0 are assigned to $x = 5$. At the *assignment stage*, the binary decisions from (2.1) of the nodes in G are $D_4 = 1$, $D_2 = 1$, $D_0 = 1$, $D_5 = 0$, $D_3 = 0$, $D_1 = 0$. Collectively the codeword to be transmitted is $c_5 = 010101$. At the *transmission stage* each node in A_5 sends a single bit utilizing a physical layer allowing the nodes' identification at the receiver (e.g. DSSS as discussed in the previous section). After receiving the signals and identifying the transmitting nodes' IDs in A_x, the sink can *decode* the codeword c_5

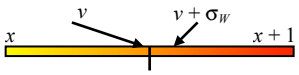

Figure 2. The value v of the phenomenon falls close to one of the endpoints of interval x. Due to sensing instrumentation imprecision σ_W, the value is incorrectly measured to fall in interval x+1.

using (2.4). Knowing $m = K^* = 3$ and ordering the IDs in A_x so that $(n_1, n_2, n_3)=(4, 2, 0)$ the sink obtains

$$x = \sum_{i=1}^{3} {}^{n_i}C_i = {}^4C_3 + {}^2C_2 + {}^0C_1 = 5_{10} = 0101_2$$

This is the correct message and it has been transmitted utilizing only 3 signals.

ES-C is significantly more energy efficient than current communication schemes deployed in WSNs, as we show in Sections 5 and 6. However, ES-C implicitly requires that *all* nodes in group G have access to message x that needs to be sent. This need not be the case in the context of WSNs where each message x is uniquely mapped to one of M intervals spanning a continuous source, as described in Section 1.2[2]. Nodes in the same group may measure different values falling in intervals x and x'. Independent on the source quantization (i.e. the length of the intervals), the probability of the latter event is not zero. Figure 2 illustrates the scenario leading to nodes in the same group reporting different measurements due to instrumentation noise σ_W. Theoretically we cannot preclude the event shown in Figure 2. The actual value v of the phenomenon may be arbitrarily close to either of interval x endpoints, and even small instrumentation noise may lead some number k of nodes in G to measure values in the adjacent but wrong interval x'. The resulting ES-C collaborative codewords sent to the sink may be invalid or, worse, wrong causing large distortion in the sink's estimate even if a single node is in error.

The instrumentation imprecision of sensors is modeled statistically in the technical literature as an additive Gaussian random variable $N(0, \sigma_W)$, independent at each node. We can set the length ε of each interval x so that the probability of a measurement being shifted to *more than one* intervals due to sensing imprecision is bounded and very low (i.e. by setting ε to a couple of standard deviations σ_W).

If a number k of nodes in a group G erroneously measure a value v' falling in interval x' and the rest $|G| - k$ nodes sense a value v falling in the correct interval x, could we find an encoding scheme with the property that almost surely the correct codeword is received at the sink, given $|x - x'| = 1$?

We answer the question in the positive, proposing the *minimum distance combinatorial encoding* presented in the next section. We avoid the above shortcoming of ES-C, while preserving ES-C's energy efficiency.

3. MINIMUM DISTANCE COMBINATORIAL ENCODING

Suppose we have any two intervals x and x' within the phenomenon's range of values such that $|x - x'| = 1$. As above, x and x' are assigned to two distinct binary codewords. However, the assignment/encoding algorithm is different. We insist that if $|x - x'| = 1$, then the hamming distance between the codewords assigned to x and x' is bounded by a constant. Since we consider a constant number $|A_x|$ active nodes out of $|G|$ nodes in a group, we can obtain a new codeword from a given valid codeword flipping *at least* two bits. One of the bits flips to 0 and the other bit flips to 1 to preserve the number of active nodes. (E.g. if only one bit flips to 1, $|A_x|$ would

[2] We use 'message x' and 'interval x' interchangeably in the context of WSN.

increase by 1; and if only 1 bit flips to 0, $|A_x|$ would decrease by one.) More specifically, let HD(c_x, $c_{x'}$) be the hamming distance between the codewords c_x and $c_{x'}$ assigned to measurements in intervals x and x' respectively. We require that

$$|x - x'| = 1 \Rightarrow \text{HD}(c_x, c_{x'}) = 2. \tag{3.1}$$

Notice that similarly to ES-C, $|c_x| = |c_{x'}| = |G|$ and $w(c_x) = w(c_{x'}) = |A_x|$.

We, next, first show the benefit of such *minimum distance combinatorial encoding* (MDCE) with the latter property; and, second, provide the algorithms to construct MDCE, and accordingly encode and decode MDCE words.

3.1 MDCE Probability of Error

Suppose the event in Figure 2 occurs with probability p_e independently at each node. Namely, the phenomenon's actual value v at a sensor's node location is within, but close to, the boundaries of interval x. Due to instrumentation noise, the sensor node at that location erroneously measures value in the interval x' such that $|x - x'| = 1$. Then, the probability P_k that k nodes are in error is given by

$$P_k = \binom{|G|}{k} p_e^k \left(1 - p_e\right)^{|G|-k} \tag{3.2}$$

Suppose node j is one of the k nodes that erroneously determines the collaborative codeword to be transmitted is $c_{x'}$. Node j checks the bit at position j in $c_{x'}$ to determine its decision. Notice that node j's decision would be erroneous *only if the bit at position* j in $c_{x'}$ is erroneous. Since $|x - x'| = 1$ we are guaranteed that $HD(c_x, c_{x'}) = 2$ from (3.1). Also, we have that $|c_{x'}| = |G|$. Hence, the probability p_j that the bit at position j in $c_{x'}$ is erroneous is given by

$$p_j = \frac{2}{|G|} \tag{3.3}$$

Given we have k nodes in error, the probability P_R that either of the k nodes reports erroneously (causing at least a single error in the codeword received at the sink) can be bounded as

$$P_R = \bigcup_{j=1}^{k} p_j \le \sum_{j=1}^{k} p_j = \sum_{j=1}^{k} \frac{2}{|G|} = \frac{2k}{|G|} \tag{3.4}$$

where the inequality follows from the union bound. Since the probability that k nodes are in error is P_k, we have that overall the probability P_E of error in the message is given by

$$P_E = P_K P_R \le \frac{2k}{|G|} \binom{|G|}{k} p_e^k \left(1 - p_e\right)^{|G|-k} \tag{3.5}$$

Figure 3 shows the values of P_E for varying numbers of k and $|G|$.

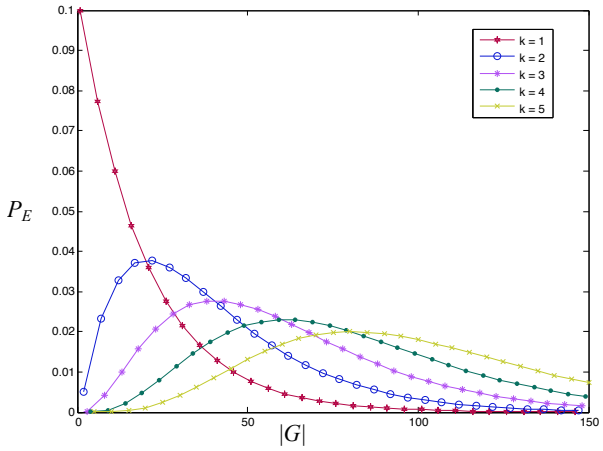

Figure 3. MDCE probability of encoding invalid or erroneous codeword given different number k of sensors with erroneous measurements, for $p_e = 0.05$.

As the instrumentation noise decreases, the probability of inaccurate measurements decreases as well. Assuming p_e is small, k is a constant, and $|G|$ is large, we have

$$P_E \le \frac{2k}{|G|} \frac{e^{-|G|p_e}(|G|\,p_e)^k}{k!} \tag{3.6}$$

$P_E \to 0$ as $|G|$ increases, and the probability that a MDCE codeword transmitted by group G is invalid or wrong can be made arbitrarily low.

3.2 MDCE Construction

Our MDCE construction starts with the well-known Gray binary code with the property that the hamming distance between two consecutive codewords is one. Suppose we have a set of M messages and consider a Gray codebook with words of length $n = |G|$. Let $x = 0$. For each consecutive word c in the Gray codebook, we check if $m = \mathrm{w}(c) = |A_x| = K^*$, so that (2.2) is satisfied; if true, we add c to our MDCE codebook, setting $c_x = c$ and then incrementing x by one.

THEOREM 2**:** *The resulting MDCE code satisfies (3.1) and hence is indeed a minimum distance combinatorial encoding.*

PROOF: We first note that the standard reflexive binary Gray code of length n is given recursively as

$$Gray_n = 0\ Gray_{n-1},\ 1\ reverse(Gray_{n-1})$$

where the *reverse*() operation simply reverses the binary sequence. Let $MDCE_{m,n}$ be the subsequence of the $Gray_n$ code where $w(c) = m$ for $c \in Gray_n$. Then,

$$MDCE_{m,n} = 0\ MDCE_{m,n-1},\ 1\ reverse(MDCE_{m-1,n-1})$$

Notice that $MDCE_{0,n} = \{000\ldots0\}$: a run of m 0s, denoted as 0^m. Also, $MDCE_{n,n} = 1^m$.

By induction, we have that the first word in $MDCE_{m,n}$ is $0^{n-m}1^m$ and the last word in $MDCE_{m,n}$ is given by $10^{n-m}1^{m-1}$. For example, $MDCE_{1,2} = \{0\ MDCE_{1,1},\ 1\ reverse(MDCE_{0,1})\} = \{0^{2-1}1^1,\ 10^{2-1}1^0\} = \{01,10\}$, which is true.

Invariantly, to obtain

$$1\ reverse(MDCE_{m-1,n-1})$$

we need to only flip 2 bits in

$$0\ MDCE_{m,n-1}.$$

To see that, we observe that $0\ MDCE_{m,n-1} = 010^{n-m-1}1^{m-1}$ and $1\ MDCE_{m-1,n-1} = 110^{n-m-1}01^{m-2}$ by induction $\forall\ m \geq 2$. Trivially, for $m = 1$ we again only flip 2 bits to transition from $0\ MDCE_{m,n-1}$ to $1\ reverse(MDCE_{m-1,n-1})$.

Therefore every two consecutive words in our code differ by the signs of two bits, and hence $|x - x'| = 1 \Rightarrow HD(c_x, c_{x'}) = 2$ is satisfied. Since we only selected codewords c where $m = w(c) = |A_x|$ from the Gray code, we have constructed minimum distance combinatorial encoding of the measurements x. □

Notice that the MDCE codewords generated in this manner have weight equal to the energy efficient codewords of ES-C generated by **Algorithm 1** for a given M and $|G|$. In fact, the set of MDCE codewords and that of ES-C are identical, for any given M and $|G|$. However, the *mapping* between codewords and messages is rather different. Table I lists a subset of the MDCE and ES-C codewords for the case $|G| = 6$, $M = 20$, where $|A_x| = K^* = 3$ respectively. Notice that the hamming distance between two consecutive codewords is exactly equal to two.

The above MDCE procedure is not rather efficient, as it is based on the binary reflexive Gray codes, which are generated recursively. **Algorithm 2**, listed on thte next page, alleviates that problem and efficiently constructs MDCE codebooks for any given $|G|$ and K^*.

THEOREM 3**: Algorithm 2** *with input* $K^* = |A_x| = m$ *and* $|G| = n$ *generates a valid* $MDCE_{m,n}$.

PROOF: Consider the binary sequence of codewords in $MDCE_{m,n}$. Suppose we index each bit in each codeword, so that the first bit indicates whether node n – 1 transmits, the second bit indicates whether node n – 2 transmits, etc. so that the last bit indicates whether node 0 transmits. For instance, the binary codeword $1^5 1^4 1^3 0^2 0^1 0^0$ is equivalent to nodes 5, 4, and 3 transmitting; the binary $1^5 1^4 1^3 0^2 0^1 0^0$ can then be interpreted as the codeword 543. Then, we can convert

$$MDCE_{m,n} = 0\ MDCE_{m,n-1},\ 1\ reverse(MDCE_{m-1,n-1})$$

to

$$MDCE_{m,n} \equiv MDCE^I_{m,n} = MDCE_{m,n-1},\ \{n-1\} \cup reverse(MDCE_{m-1,n-1})$$

Let $(n_m, n_{m-1}, \ldots, n_2, n_1)$ be a codeword in $MDCE^I_{m,n}$. Note that the sequence of these codewords in $MDCE^I_{m,n}$ is sorted in the lexicographic order of

$$(n_m, -n_{m-1}, \ldots, (-1)^{m-1} n_1).$$

This follows directly by induction on the structure of the $MDCE^I_{m,n}$ code. It is straightforward to check that **Algorithm 2** generates the codewords of $MDCE^I_{m,n}$ exactly in the lexicographical order of $(n_m, -n_{m-1}, \ldots, (-1)^{m-1} n_1)$ as well. □

Table I below illustrates the encoding of $MDCE^I_{3,6}$ generated by **Algorithm 2** as an example. The basic combinatorial encoding from Section 2 is also given in Table I, for comparison.

Algorithm 2 could be used to encode/decode measurements. However, that would require storing an assignment table similar to Table I at the sensor and sink nodes, and then looking up values of x and c_x respectively for encoding and decoding. The next theorem exploits the structure of the $MDCE^I_{m,n}$ code so that encoding and decoding are done much more efficiently, without requiring extra space for code tables.

Let the range of $MDCE^I_{m,n}$ be M, so that (2.2) is satisfied.

ALGORITHM 2 List *all minimum-distance* assignments to K^* out of $|G|$ nodes

Input: K^* and $|G|$
Output: *L:* list of assignments $A_x = (n_{K^*}, n_{K^*-1}, \ldots, n_1)$ for $\forall x \in \mathrm{I}$
Algorithm:

```
1: m ← K*
2: n_{m+1} ← |G|
3: L ← {∅}
4: j ← m
5: while j ≥ 1 do
6:    n_j ← j − 1
7:       j ← j − 1
8: j ← 1
{At each iteration add the m-tuple assignment to the list L, where the m-tuple comprises
the values of n_m, n_{m-1}, …, n_2, n_1 at the beginning of that iteration; note m = K*}
9: while j ≤ m do
10:    L ← L ∪ (n_m, n_{m-1}, …, n_2, n_1)
11:    increment_n_j ← false
12:    decrement_n_j ← false
13:    if m is odd then
14:       if n_1 + 1 < n_2 then
15:          n_1 ← n_1 + 1
16:          continue {add assignment to L}
17:       else
18:          j ← 2
19:          decrement_n_j ← true
20:    else {m is even}
21:       if n_1 > 0 then
22:          n_1 ← n_1 − 1
23:          continue {add assignment to L}
24:       else {n_1 is zero}
25:          j ← 2
26:          increment_n_j ← true
27:    while true do
28:       if decrement_n_j = true then
29:          if n_j ≥ j then
30:             n_j ← n_j − 1
31:             n_{j−1} ← j − 2
32:             break {go to 9; add assignment to L}
33:          else {n_j < j}
34:             j ← j + 1
35:             increment_n_j = true
36:       if increment_n_j = true then
37:          if n_j + 1 < n_{j+1} then
38:             n_{j−1} ← n_j
39:             n_j ← n_j + 1
40:             break {go to 9; add assignment to L }
41:          else
42:             j ← j + 1
43:          if j ≤ m then
44:             decrement_n_j = true
45:          else { j > m }
46:             break {go to 9; terminate}
```

Table I. MDCE and Combinatorial Encoding Example

x	MDCE	$MDCE^{I}_{3,6}$	Combinatorial encoding	Combinatorial encoding index
0	000111	2,1,0	000111	2,1,0
1	001101	3,2,0	001011	3,1,0
2	001110	3,2,1	001101	3,2,0
3	001011	3,1,0	001110	3,2,1
4	011001	4,3,0	010011	4,1,0
5	011010	4,3,1	010101	4,2,0
6	011100	4,3,2	010110	4,2,1
7	010101	4,2,0	011001	4,3,0
8	010110	4,2,1	011010	4,3,1
9	010011	4,1,0	011100	4,3,2
10	110001	5,4,0	100011	5,1,0
11	110010	5,4,1	100101	5,2,0
12	110100	5,4,2	100110	5,2,1
13	111000	5,4,3	101001	5,3,0
14	101001	5,3,0	101010	5,3,1
15	101010	5,3,1	101100	5,3,2
16	101100	5,3,2	110001	5,4,0
17	100101	5,2,0	110010	5,4,1
18	100110	5,2,1	110100	5,4,2
19	100011	5,1,0	111000	5,4,3

THEOREM 4**:** Any integer value x in the range of $MDCE^{I}_{m,n}$ can be uniquely represented as

$$x = \sum_{j=1}^{m} (-1)^{m-j} \left({}^{n_j}C_j - 1 \right)$$

where $(n_m, n_{m-1}, \ldots, n_2, n_1)$ is a codeword in $MDCE^{I}_{m,n}$.

PROOF: We note that the tuple $(n_m, n_{m-1}, \ldots, n_2, n_1)$ is the x-th codeword generated by **Algorithm 2**, where

$$x = \sum_{j=1}^{m} (-1)^{m-j} \left({}^{n_j}C_j - 1 \right) \tag{3.7}$$

To see that, consider the codewords $(n_j, n_{j-1}, \ldots, n_2, n_1)$ and $(n_{j-1}, n_{j-2}, \ldots, n_2, n_1)$ of the two codes $MDCE^{I}_{j,n}$ and $MDCE^{I}_{j\text{-}1,n'}$. Let the codeword $(n_j, n_{j-1}, \ldots, n_2, n_1)$ be the x-th codeword generated by **Algorithm 2** and codeword $(n_{j-1}, n_{j-2}, \ldots, n_2, n_1)$ be the y-th codeword generated by **Algorithm 2**.

By induction,

$$x = {}^{n_j+1}C_j - 1 - y \text{ for any } j > 0.$$

Then, starting with $j = 1$ and summing to $j = m$ using the above equation we obtain that

$$x = \sum_{j=1}^{m} (-1)^{m-j} \left({}^{n_j}C_j - 1 \right) \qquad \square$$

Algorithm 3 and **Algorithm 4**, listed below, utilize directly *Theorem* 4 respectively to encode measurement x, in the range of code $MDCE^{I}_{m,n}$, and then decode x at the sink. Similarly to **Algorithm 1**, **Algorithm 3** is run at each node of group G, upon measuring x. If node j's ID is in the output A_x then j transmits a 1. The sink receives the resulting 1's from all nodes in A_x (thus determining A_x). The sink then runs **Algorithm 4** to recover x; **Algorithm 4** computes the formula in (3.7).

ALGORITHM 3 Assign *msg x* to a K^*- subset A_x of G, where A_x forms a MDCE codeword

Input: x, K^*, $|G|$
Output: $A_x = \{n_{K^*}, n_{K^*-1}, \ldots, n_1\}$
Algorithm:
1: $A_x \leftarrow \{\varnothing\}$
2: $m \leftarrow K^*$
3: $n \leftarrow |G|$

{At each iteration, add a node ID to the assignment A_x of nodes to measurement x}
3: **while** $m \geq 1$ **do**
4: **while** ${}^{n}C_{m} > x$ **do**
5: $n \leftarrow n - 1$
6: $n_m \leftarrow n + 1$
7: $A_x \leftarrow A_x \cup n_m$
8: $x \leftarrow {}^{n+1}C_{m} - x - 1$
9: $m \leftarrow m - 1$

ALGORITHM 4 Decode a MDCE codeword to measurement x

Input: $A_x = \{n_{K^*}, n_{K^*-1}, \ldots, n_1\}$, K^*
Output: x
Algorithm:
1: $m \leftarrow K^*$
2: **if** m is even **then**
3: $x \leftarrow 0$
4: **else** {m is odd}
5: $x \leftarrow -1$
6: $c \leftarrow 1$
7: **while** $m \geq 1$ **do**
8: $x \leftarrow x + c\,{}^{n_m}C_{m}$
9: $c \leftarrow -c$

Notice that both **Algorithms 3** and **4** only work on a *single* codeword. The code table computed by **Algorithm 2** is no longer required. In the case of $MDCE^{I}_{m,n}$, the runtime of **Algorithm 4** is $O(m)$; here $m = |A_x| = K^*$. The worst case runtime input for **Algorithms 3** is $|A_x| = K^* = 1$ and $x = M - 1$, yielding $O((M - 1)/|G|)$ performance. However, in practice the average runtime of **Algorithm 3** is much lower depending on the distribution of x and the values of $|A_x|$ and $|G|$.

We dub ES-MDCE the resulting scheme employing **Algorithms 3** and **4** to construct MDCE.

4. SPARSE DSSS RECEIVER

In this section we explore an interesting sparseness property of the ES signal received at the sink. This allows us to suggest a novel DSSS design that significantly reduces the complexity and cost of the status quo DSSS receivers.

In regular DSSS systems, each PN sequence corresponds to a unique signature waveform $s_i(t)$ assigned to node i, $0 \leq t \leq T$ and $1 \leq i \leq N$, where T is the symbol duration and N is the total number of users in the system. At any given timeslot, the

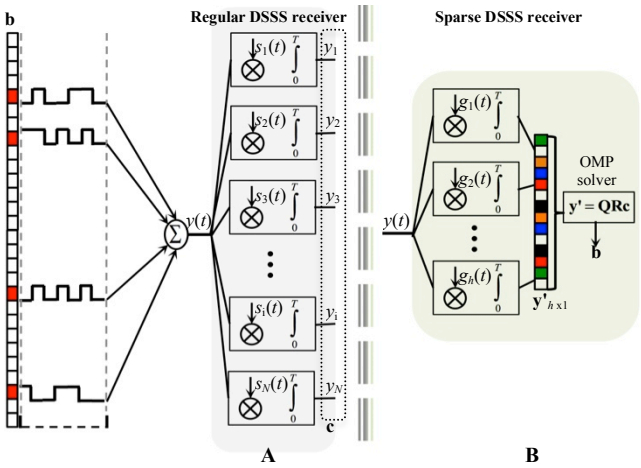

Figure 4. Receiving signal $y(t)$ with K-sparse support vector b. A – regular DSSS receiver consisting of N matched filters; B – sparse DSSS receiver consisting of $K\log(N/K)$ matched filters.

sink receives a signal $y(t)$ as a superposition of the transmitted signature waveforms. As shown in Figure 4A, the sink employs a set of N matched filters (MFs) for decorrelation of the received signals. Each MF_i correlates $y(t)$ with the signature waveform $s_i(t)$ of node i. Each MF_i comprises high complexity and high precision analogue circuitry driving the cost of a DSSS receiver proportional to the number of MFs/nodes in the system. Thus, along reducing circuit complexity, reducing the number N of MFs also reduces the actual cost of a DSSS receivers' deployment.

4.1 Sparse Support Recovery

In a general setting, the task of reducing the number of MFs is hard, if at all feasible. We need information for all the N possible signature waveforms at the receiver in order to distinguish between the N different transmitted signals that can be received at any time. The crux here is that the ES communication scheme allows the number of signals transmitted and received during any timeslot to be much lower than N. Consider an ES system consisting of a single group of size $N = |G|$ with $|A_x| = K$ active nodes. We can encode ${}^{|G|}C_K$ different messages; for instance if $K = 3$ and $|G| = 40$ we can generate ${}^{|G|}C_K = 9880$ different codewords. At the receiver, each codeword contains exactly K non-zero bits and can be represented as a binary K-sparse vector $\mathbf{b}_{N\text{x}1}$. The received signal is

$$y(t) = \sum_{i=1}^{N} b_i s_i(t)$$

containing the signature waveforms of the transmitting nodes. The task of the MFs at the receiver is only to recover the *support* of **b** and thus find the identities of the transmitting nodes. As discussed above, this is sufficient for the operation of ES, which does not require symbol detection. Let l denote the length of the PN sequences. I.e. l equals the number of chips per PN sequence.

THEOREM 5**:** The support of **b** is recovered correctly *w.h.p.* at the DSSS receiver utilizing on the order of $h = K\log(N)$ matched filters, instead of N matched filters, iff $l > 2K\log(N)$.

PROOF: To show this, suppose each MF now correlates the received signal $y(t)$ with a signature waveform $g_j(t)$ (Figure 4B), instead of utilizing the regular signature waveform $s_i(t)$. The signature $g_j(t)$ is obtained as follows:

$$g_j(t) = \sum_{i=1}^{N} q_{ji} s_i(t), \text{ where } 1 \le j \le h$$

where q_{ji} are the random ±1 entries of a Rademacher matrix $(\mathbf{Q})_{h \times N}$. Now, we have h signature waveforms at h MFs. Each signature waveform $g_j(t)$ is a random linear combination of the original N signature waveforms. After receiving $y(t)$, the output of the j^{th} MF is given by the crosscorrelation between $y(t)$ and $g_j(t)$, namely the inner product $y'_j = \langle g_j(t), y(t) \rangle$.

The overall output of all h MFs, represented in vector form, is $\mathbf{y'} = \mathbf{QRb}$, where $(\mathbf{R})_{NxN}$ is the cross-correlation matrix of the original signature waveforms $s_i(t)$.

Due to the *universality* property and *restricted isometry property* (RIP) of $\mathbf{Q}$, the matrix $\mathbf{\Theta} = \mathbf{QR}$ also has the RIP w.h.p., if $h \ge \alpha K \log(N/K)$, where α is a small constant [Baraniuk 2007; Donoho 2006]. We then recover the support of $\mathbf{b}$ using a standard Orthogonal Matching Pursuit (OMP) [Tropp and Gilbert 2007]. Applying the argument in [Tropp and Gilbert 2007], we require that $l > 2K\log(N)$, and this is sufficient for OMP to recover the support of $\mathbf{b}$ with high probability. □

Suppose the PN sequence has length $l = 32$ chips (similarly to IEEE 802.14.5). This implies that, for instance, with $K = 3$ and $N = |G| = 40$, we require 16 MFs to identify the 3 transmitting nodes w.h.p., in comparison to 40 MFs in a regular DSSS receiver. We achieve a 60% reduction in circuitry! Notice that we have $l = 32 > 2*3\log(40) = 31.93$ and satisfy the sufficient condition of Theorem 5. We note that a compressive sensing argument similar to the above can readily be devised in order to reduce significantly the *length* of the PN sequences generated in a DSSS system *instead of* reducing the complexity of the receiver. This could potentially increase the energy efficiency of ES further. We do not pursue this avenue here, but it could be of interest for future research.

4.2 Symbol Detection

In the preceding subsection, we showed that due to the sparseness of ES collaborative codewords, we can recover nodes' identities *w.h.p* at the sink, efficiently utilizing fewer MFs, which is sufficient for ES operation. However, other potential applications of the sparse DSSS receiver may utilize symbol detection (i.e. determining whether the bit transmitted by a given node j is 0 or 1).

4.2.1. Sparse DSSS Symbol Estimator. The receiver detector can be constructed using different approaches. The scenario here is very similar to multi-user detection (MUD). It has been shown that the optimal maximum likelihood detector for MUD is intractable in practice [Verdu 1986]. Therefore various lower complexity approximate detectors have been proposed (e.g. section 13.4.4 of [Goldsmith 2005]). The MMSE estimator is in the category of efficient low complexity, linear MUD detectors. Therefore, in what follows we derive the optimal MMSE estimator for the sparse DSSS receiver.

The value of the bit that node j sends have so far been assumed to be 1_j, given $D_j = 1$ in (2.1). Instead, suppose now node j sends either a 0 or 1 (equivalently -1 or 1) when $D_j = 1$. The sink needs to detect the correct sign of the transmitted bit. That is, K nodes transmit in group G; next we recover the identities of the transmitting nodes correctly, per Theorem 5; and finally, we require that the receiver determine the sign of the bit transmitted by each of the K nodes. Notice that once the support vector $\mathbf{b}_S$ of $\mathbf{b}$ is determined, errors can still occur while detecting each of the K bits in $\mathbf{b}$ due to multi access interference (MAI).

Assume that we have already recovered correctly the sparse support of the received signal, as in the previous subsection. Then, also notice that we need to estimate only the signs (-1 or 1) of the bits in the *support* of **b**; the rest of b_j's are 0.

As before, let **R** be the normalized cross-correlation matrix of the original $N = |G|$ signature waveforms:

$$\mathbf{R} = \left[\left\langle s_i(t), s_j(t)\right\rangle\right]_{i,j=1}^{N}$$

Suppose the channel gains p_i, $i = 1, 2, \ldots, N$, are known at the receiver and

$$p_i = a_i\sqrt{P_i}$$

where a_i is the channel amplitude also known at the receiver. Let **P** be a diagonal matrix with entries $p_1, p_2, \ldots, p_N$. The output of the sparse DSSS receiver in vector form is then given by

$$\mathbf{y'} = \mathbf{QPb} + \mathbf{z}. \tag{4.1}$$

Here, **z** captures the effect of MAI and is a Gaussian r.v. with zero mean and covariance $\sigma_Z^2\,\mathbf{QRQ}^T$.

As per the typical MUD MMSE ([Poor and Verdu 1997]), the MMSE detector at the sparse DSSS receiver determines the sign of each bit by applying a linear transform **D** to the output of the h MFs. I.e.

$$\hat{b}_j = \mathrm{sgn}((\mathbf{Dy'})_j)$$

so that the mean square error between **b** and $\hat{\mathbf{b}} = (\hat{b}_j)_N$ is minimized.

The challenge in our case is to find $\mathbf{D}^*$ *for the sparse DSSS receiver so that*

$$\mathbf{D}^* = \underset{\mathbf{D}}{\mathbf{argmin}}\,\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)^2\right].$$

We start by noting that

$$\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)^2\right] = tr\left[\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)\left(\mathbf{b} - \mathbf{Dy'}\right)^T\right]\right]$$

Carrying the multiplication, substituting **y'** form eq. (4.1) and the covariance expression of **z**, we obtain

$$\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)^2\right] = tr\left(\mathbf{I} + \mathbf{D}\left(\mathbf{QP}^2\mathbf{Q}^T + \sigma_Z^2\mathbf{QR}^{-1}\mathbf{Q}^T\right)\mathbf{D}^T - \mathbf{PQ}^T\mathbf{D}^T - \mathbf{DQP}\right)$$

Suppose

$$\mathbf{D}^* = \mathbf{PQ}^T\left(\mathbf{QP}^2\mathbf{Q}^T + \sigma_Z^2\mathbf{QR}^{-1}\mathbf{Q}^T\right)^{-1} \tag{4.2}$$

It can be shown by straightforward arithmetic that

$$\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)^2\right] = tr\left(\mathbf{I} - \mathbf{D}^*\mathbf{B}\left(\mathbf{D}^*\right)^T + \left(\mathbf{D} - \mathbf{D}^*\right)\mathbf{B}\left(\mathbf{D} - \mathbf{D}^*\right)^T\right)$$

where $\mathbf{B} = \left(\mathbf{QP}^2\mathbf{Q}^T + \sigma_Z^2\mathbf{QR}^{-1}\mathbf{Q}^T\right)$. Substituting $\mathbf{D}^*$ from (4.2) in the second term above

$$\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy'}\right)^2\right] = tr\left(\mathbf{I} - \mathbf{PQ}^T\mathbf{B}^{-1}\mathbf{QP} + \left(\mathbf{D} - \mathbf{D}^*\right)\mathbf{B}\left(\mathbf{D} - \mathbf{D}^*\right)^T\right)$$

Notice that $\mathbf{QP}^2\mathbf{Q}^T$ is a positive definite matrix. Also, the cross correlation matrix **R** is positive definite, hence $\sigma_Z^2\mathbf{QR}^{-1}\mathbf{Q}^T$ is positive definite. Then, **B** is positive definite.

We then have $tr\left(\mathbf{PQ}^T\mathbf{B}^{-1}\mathbf{QP}\right) > 0$. Consequently,

$$\mathbf{D}^* = \mathbf{PQ}^T\left(\mathbf{QP}^2\mathbf{Q}^T + \sigma_Z^2\mathbf{QR}^{-1}\mathbf{Q}^T\right)^{-1} = \underset{\mathbf{D}}{\mathbf{argmin}}\,\mathrm{E}\left[\left(\mathbf{b} - \mathbf{Dy}'\right)^2\right].$$

Thus, we have obtained the MMSE linear transformation, and can estimate the signs of the transmitted symbols in the support of **b**.

4.2.2. Conventional MUD Symbol Estimator.

Let $Pr_j(error)$ be the probability of error in detecting node's j bit sign at the receiver. For completion, we next derive bounds on $Pr_j(error)$ of the conventional multi-user detection receiver. We show the conditions that need to be met so that $Pr_j(error) \to 0$ and compare $Pr_j(error)$ of the conventional symbol estimator with the respective performance of the sparse DSSS receiver developed in 4.2.1 above.

As above, we utilize the MMSE estimator. Carrying over notation from the previous section, this time we have

$$\mathbf{y}' = \mathbf{RPb} + \mathbf{z}. \tag{4.3}$$

Here, **z** is a vector representing only white Gaussian noise variables with power σ_Z^2, and now **y'** has N entries. Similarly to the above the MMSE detector determines

$$\hat{b}_j = \mathrm{sgn}((\mathbf{Dy}')_j)$$

so that the mean square error between **b** and $\hat{\mathbf{b}} = (\hat{b}_j)_N$ is minimized. It is well-known (e.g. [Poor and Verdu 1997]) that the linear transformation **D** here is given by

$$\mathbf{D} = \left(\mathbf{R} + \sigma_Z^2\mathbf{P}^{-2}\right)^{-1}$$

From Proposition 3.1 in [Poor and Verdu 1997], asymptotically, as P_j/σ_j grows, the probability of error in detecting node's j sent bit is given by

$$\Pr_j(error) = Q\left(P_j / \left(\sigma_Z\sqrt{\left(\mathbf{R}^{-1}\right)_{j,j}}\right)\right) \tag{4.4}$$

The impact of the PN sequence ensemble choice on the system performance is not immediately obvious from the term $(\mathbf{R}^{-1})_{j,j}$ as given in (4.4). We, next, provide closed form bounds on this expression, explicitly as a function of the maximum cross-correlation between PN sequence signatures used in the system.

LEMMA 1**:** Let N be the number of users in the system and χ_{max} be the highest cross-correlation[3] between any two PN sequences in the system. Then

$$[1 + (N-1)\chi_{\max}]^{-1} \le \left(\mathbf{R}^{-1}\right)_{j,j} \le [1 - (N-1)\chi_{\max}]^{-1}.$$

PROOF: As observed earlier, the cross-correlation matrix **R** is positive definite and its eigenvalues can be ordered as $0 < \lambda_1 \le \lambda_2 \le \ldots \le \lambda_N$. Let $\Delta \ge \lambda_N$ and $0 < \nabla \le \lambda_1$. Then, from [Robinson and Wathen 1992], we know that

$$\frac{1}{\Delta} \le \left(\mathbf{R}^{-1}\right)_{j,j} \le \frac{1}{\nabla} \tag{4.5}$$

Let **R** = **E** + **F**, where **E** is diagonal matrix with entries e_1, …, e_m and **F**'s diagonal entries are 0's. From *Gershgorin's Circle Theorem* (GCT), ([Golub, van Loan 1996]), if

[3] χ_{max} depends on the PN sequence type (e.g. Gold, Kasami, etc.) and can be computed or obtained from tables as in [Goldsmith 2005, Section 13.4.1] and [Stuber 2001, tables 9.1, 9.2].

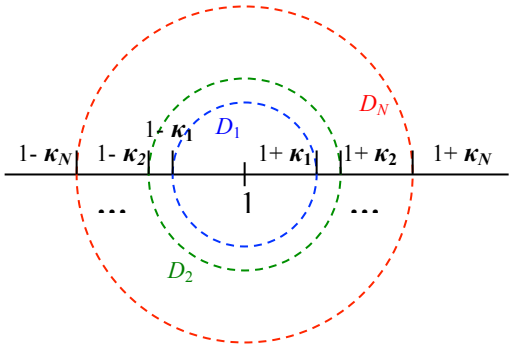

Figure 5. Possible positions of R's eigenvalues on the real line. Here, interval $[1-\kappa_N,\ 1+\kappa_N]$ is described by circle D_N.

$\lambda(\mathbf{R})$ is the set of all eigenvalues of $\mathbf{R}$, then

$$\lambda(\mathbf{R}) \subseteq \bigcup_{i=1}^{N} D_i$$

where

$$D_i = \left\{ z \in \mathbb{C} : |d_i - z| \le \sum_{j=1}^{N} \left| (\mathbf{F})_{i,j} \right| = \kappa_i \right\}.$$

That is, all eigenvalues of $\mathbf{R}$, lie within the union of closed disks D_i, in the complex plane. Each disk D_i is centered at the corresponding diagonal entry of $\mathbf{R}$ and has radius equal to the sum of the non-diagonal entries at row/column i of $\mathbf{R}$. In our case, all eigenvalues of $\mathbf{R}$ are real and D_i's are intervals on the real line centered around $(\mathbf{R})_{i,i} = 1$. Figure 5 illustrates the possible positions of $\mathbf{R}$'s eigenvalues.

$\lambda(\mathbf{R})$ lie within the interval $D_N = [1-\kappa_N,\ 1+\kappa_N]$, where $\kappa_1 < \kappa_2 < \ldots < \kappa_N$. Then,

$$\kappa_N = \max_{1 \le i \le N} \arg \left\{ \sum_{j=1}^{N} \left| (\mathbf{F})_{i,j} \right| \right\} \le (N-1)\chi_{\max}.$$

Suppose $\kappa_m = (m-1)\chi_{\max}$. From *GCT*, we have

$$\min\left\{\lambda(\mathbf{R})\right\} \ge 1-\kappa_m = \nabla > 0 \ \ and \ \ \max\left\{\lambda(\mathbf{R})\right\} \le 1+\kappa_m = \Delta$$

Therefore,

$$\nabla = 1-(m-1)\chi_{\max} \ \ and \ \ \Delta = 1+(m-1)\chi_{\max}$$

Then, combining with (4.5) we obtain

$$[1+(N-1)\chi_{\max}]^{-1} \le \left(\mathbf{R}^{-1}\right)_{j,j} \le [1-(N-1)\chi_{\max}]^{-1} \qquad \square$$

Observe that as the highest cross-corelation between two PN sequences goes to 0 (as it should ideally) $\left(\mathbf{R}^{-1}\right)_{j,j} \to 1$, from the bounds in Lemma 1.

THEOREM 6**:** The asymptotic symbol detection probability, $Pr_j(error)$, as P_j/σ_j increases, of the conventional MMSE MUD detector is upper bounded by

$$U = \frac{\sigma_Z}{P_j \sqrt{2\pi}(1-(m-1)\chi_{\max})} \exp\left\{ \left(\frac{P_j}{\sigma_Z} \right)^2 \frac{(m-1)\chi_{\max} - 1}{2} \right\}$$

and lower bounded by

$$L = \frac{\sigma_Z}{P_j \sqrt{2\pi}(1+(m-1)\chi_{\max})} \exp\left\{-\left(\frac{P_j}{\sigma_Z}\right)^2 \frac{1+(m-1)\chi_{\max}}{2}\right\}$$

PROOF: Using

$$Q(x) \le \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$$

and substituting the result from Lemma 1 in 4.4 yields the bounds.

Figure 6 plots numerically the symbol estimation probability of error for the conventional DSSS MUD detector based on MMSE (along with its upper bound) and the sparse DSSS receiver based on the MMSE with the optimal linear transform $\mathbf{D}^*$ derived in (4.2). We set $\sigma_Z = 0.2$, $P_j = 1$, $\chi_{max} = 0.02$, $N = 40$. The number of MFs varies in the case of the sparse DSSS receiver. Notice, that as the number of MFs increases the symbol estimation probability of error decreases for the sparse DSSS receiver and becomes comparable to that of the conventional DSSS receiver. We note *that this is conditional on the support of* **b** *being correctly computed*. For that, we have to ensure greater number of MFs than shown on the figure here. We would like to emphasize again that the ES scheme, does not require the symbol sign estimation and support recover as given in section 4.1 suffices.

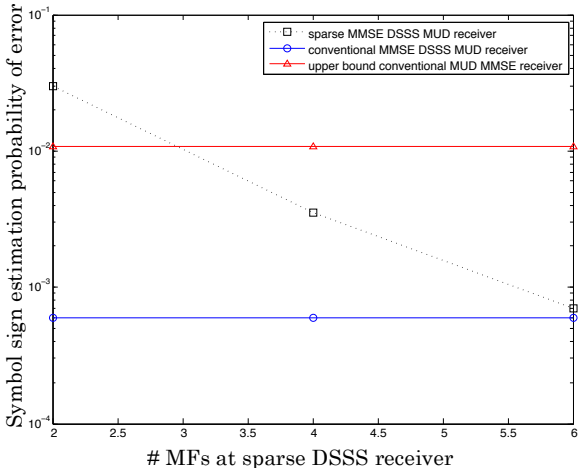

Figure 6. Symbol sign estimation probability of error for varying number of MFs at the sparse DSSS receiver's. The number $N = 40$ of MFs for the conventional receiver numerical performance and theoretical upper bound is a constant. Notice that ES does not require symbol sign estimation and requires only support recovery as described in subsection 4.1.

5. A WIRELESS SENSOR NETWORK APPLICATION

Next, we demonstrate the benefits of encoded sensing in a practical WSN application. All N sensor nodes utilize ES to communicate their measurements with a sink.

5.1 Phenomenon and Sensing Models

We assume a continuous source **S** giving rise to a space-time random field $s(l, x, y)$. In the l-th timeslot the random field is defined as $\{S_j[l] = s(l, x_j, y_j): (x_j, y_j) \in A\}$ at spatial position (x_j, y_j), according to a spatial-correlation model. The instances $S_j[l]$ are

modeled as joint Gaussian random variables (JGRV). Considering a single discrete-time interval sample, the time index l can be dropped. The JGRV is characterized by:

$$E[S_j]=0\,,\ Var[S_j]=\sigma_S^2\,,\ \rho_{i,j}=E[S_iS_j]/\sigma_S^2\,,$$

where $\rho_{i,j}$ are the correlation coefficients of the JGRV[4].

Often physical phenomena's values at different points in space are dependent on one another and are correlated via some function of the distance between them. Formally, this function is represented by a parameterized spatial covariance model, which reflects the nature of the specific phenomenon and determines the correlation coefficient. The covariance model we selected is the *Power Exponential*[5] since different physical phenomena monitored by sensor networks could be approximated this way [Vuran and Akyildiz 2006; Berger et al. 2001; Yuan et al. 2008]:

$$\rho_{i,j}=K_\theta\left(\left\|\mathbf{i}-\mathbf{j}\right\|\right)=e^{\left(-\left\|\mathbf{i}-\mathbf{j}\right\|/\theta_1\right)^{\theta_2}} \tag{5.1}$$

where $\|\bullet\|$ denotes Euclidean distance between sensors i and j at (x_i,y_i) and (x_j,y_j).

A sensor node j with coordinates $(x_j,\ y_j)$ does not have direct access to the phenomenon value S_j. Rather, it obtains a distorted measurement X_j of S_j due to inherent *sensing imprecision/noise* W_j. W_j's are assumed to be *i.i.d.* Gaussian random variables, such that $E[W_j]=0$ and $Var[S_j]=\sigma_W^2$.

Ultimately, j measures $X_j = S_j + W_j$.

5.2 Encoding Sensing Setup

To utilize ES we need to ensure that the network is partitioned in groups of $|G|$ nodes such that the nodes in each group have access to the same message x. To this end, there are two conditions we need to meet simultaneously.

- First, each group occupies a site where the *actual values* of the source at the different points by the sensors fall in the same interval $x\in I$ w.h.p..
- Second, noise due to electronic imprecision should not offset node j's measurement in an interval x' different from x, for all $j\in G$, where $|x-x'|>1$.

Source quantization: To meet the *second condition*, we quantize the interval I so that the probability is low that the Gaussian measurement imprecision W_j offsets the measurement of any $j\in G$ outside interval x', where $|x-x'|\le 1$. We set the size of each interval to $\varepsilon=2\beta\sigma_W$, $\beta\ge 2$.

Vector quantization: To meet the *first condition*, the network is partitioned in disjoint groups of highly correlated nodes. We utilize the distributed vector quantization algorithm underlying the CC-MAC protocol presented in [Vuran and Akyildiz 2006]. Given the nodes' spatial statistical distribution Γ and the spatial correlation model as an input, the algorithm selects k *representative* nodes out of the N nodes in the WSN. In [Vuran and Akyildiz 2006], the selected k nodes are chosen so that the spatial correlation between their measurements is reduced. Each of the

[4] We assume memoryless source and do not account for temporal correlations similarly to [32]. ES is readily adjustable to temporal correlations: if some readings are more likely than others **Algorithms** 3 and **4** can be easily extended to allow for a round-robin assignment of nodes to x in order to improve the energy balance in a group.

[5] The following discussion is independent of the specific correlation function of choice.

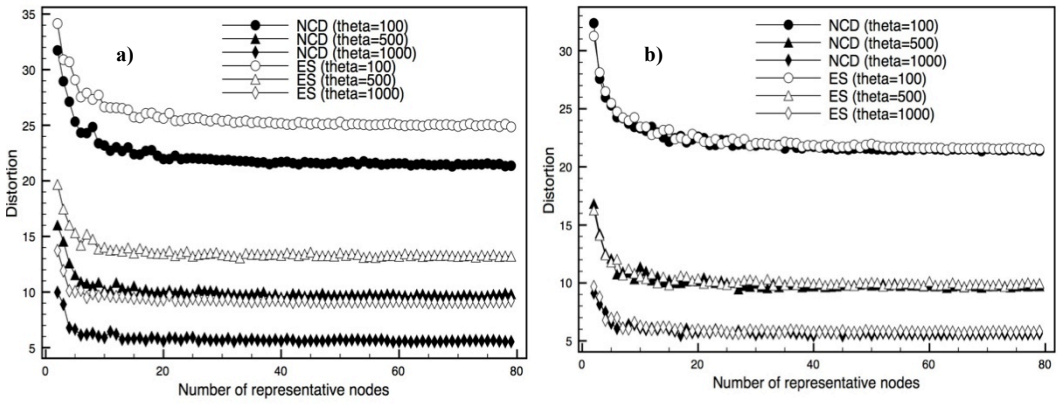

Figure 7. Distortion of measurement for *ES* and *NCD*; σ_W = 0.25, 0.05 respectively in a) and b); σ_S = 5 throughout. ES and NCD converge for lower values of σ_W.

selected k nodes transmits (non-cooperatively) a measurement to the sink every timeslot; the rest of the N nodes' measurements can be eliminated. This describes the basic operation of the *non-cooperative transmission with decorrelation* (NCD) scheme. The k nodes are selected so that distortion in the sink's estimate of the source **S** is *minimized to a given QoS threshold*.

Here we exploit a related but different property of the algorithm in [Vuran and Akyildiz 2006] to satisfy the ES requirement that a group G of nodes all sense values in the same interval x. All nodes within distance $r_{corr} < r$ of each representative node, have highly spatially-correlated measurements and report almost identical values. r_{corr} is an output of the representative node vector quantization algorithm used in [Vuran and Akyildiz 2006]. Consider a disk, O_i, with radius $r_{corr,}$ centered at representative node i.

Definition 1 (Representative Group). A group of nodes, G_i, situated in disk O_i is called a *Representative Group*[6].

Nodes within each representative group G_i utilize ES-MCDE to transmit their measurements to the sink. Encoding sensing operates in each G_i, as described in Sections 1.1 and 2: each short interval x is assigned to a distinct subset of nodes A_x in G_i according to **Algorithm 3**. Upon measuring a value in interval x, each node $j \in G_i$ checks if it is in A_x. If so, j transmits a signal. The sink receives the codeword c_x and recovers the interval x using **Algorithm 4**. In the Appendix, we show analytically that as the sensing imprecision due to W_j decreases, the distortion $D_{ES}(k)$ in the sink's estimate due to ES converges to the *optimal distortion* $D_{NCD}(k)$ *that can be obtained within QoS constraints, via non-cooperative transmission of k representative nodes* [Vuran and Akyildiz 2006]. Figure 7 illustrates the numerical evaluation of $D_{ES}(k)$ and $D_{NCD}(k)$ derived in the Appendix. The behavior of the two distortion functions is evaluated for varying number of representative nodes/groups, given the covariance model (5.1) and its parameters: $\theta_1 \in \{100, 500, 1000\}$, and $\theta_2 = 1$.

5.3 Performance Evaluation

5.3.1. Simulation System. The simulation environment and schemes' implementations are programmed in JAVA and utilize the BLOG Inference Engine [Bayesian Logic (BLOG) Inference Engine 2014], available online. In each simulation run, N nodes

[6] It is assumed that i can estimate the distances to its neighboring nodes, as in [Vuran and Akyildiz 2006].

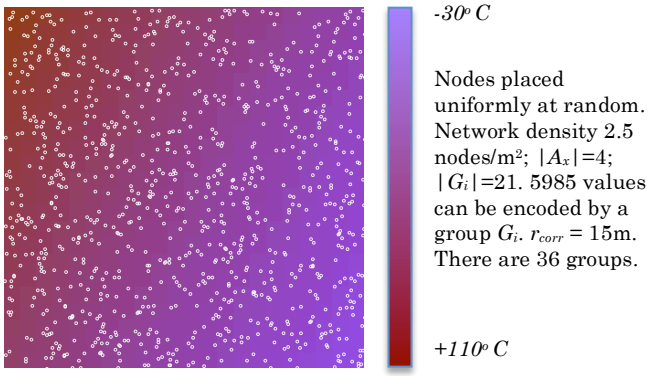

Figure 8. A WSN instance: θ_1=10000 and θ_2 =2 in K_θ

are placed within a 100[m] × 100[m] square area. S_j's are modeled as spatially correlated JGRVs with covariance model K_θ (5.1).

- The phenomenon's values are assumed to be in the range [-4σ_S, 4σ_S]; the standard deviation is σ_S =625.

- Notice that if nodes' sensors measure temperature with precision 0.1 degree, the range in our setup would allow for temperature measurements between -250 and 250 degrees for *all schemes*. A sample output of the system is shown in Figure 8. The BLOG Inference Engine is used to generate the set of S_j's; θ_1 = 10000 and θ_2 = 2 are the parameters of K_θ.

- For all simulated schemes, the transmit power per bit, E_b, is standard: E_b= 14[dBm]=25[mW]. All schemes utilize a standard DSSS with processing gain F = 512, at the physical layer. Thus the acquisition capacity of the system is approximately 36 users/nodes. An exception here is the distributed transmit beam forming, which is itself a physical layer protocol. We describe the evaluation of its performance in our system with more detail later in the paper.

5.3.2. Simulated Schemes. We analyze the energy efficiency of the system, the distortion of the sink measurement, and compare ES performance to systems communicating via *distributed source coding* (DSC) [Yuan et al. 2008; Critescu et al. 2005]; *non-cooperative transmission with decorrelation* (NCD) [Vuran and Akyildiz 2006]; and *cooperative distributed transmit beamforming* (BF) [Mudumbai et al. 2010].

- Per *ES*, the range of phenomenon values is quantized into a sequence of intervals where M = 5000. Each interval is of length ε = 1, with σ_W = 0.25.

- Per *NCD*, *k representative nodes* are selected by the CC-MAC algorithm as above, and each of the k nodes transmits, non-cooperatively, the entire measurement. The transmissions of highly correlated nodes are eliminated with the goal of better energy efficiency. The scheme's energy efficiency is equivalent to a duty cycle scheme. In the latter, only one node is selected to transmit the entire measurement from each of the *k representative groups,* at each timeslot.

- The algorithm underlying the scheme *DSC* is a single hop variation of **Algorithm 2** in [Critescu et al. 2005] and **Algorithm 1** in [Yuan et al. 2008]. The algorithm is frequently used in more recent works on DSC in wireless sensor and ad-hoc networks. Given a neighborhood radius r_i around a representative node i, the algorithm constructs an ordered sequence C_i of the nodes that are within distance r_i of i. Next, in the order of the sequence, the algorithm allocates rates as follows: the

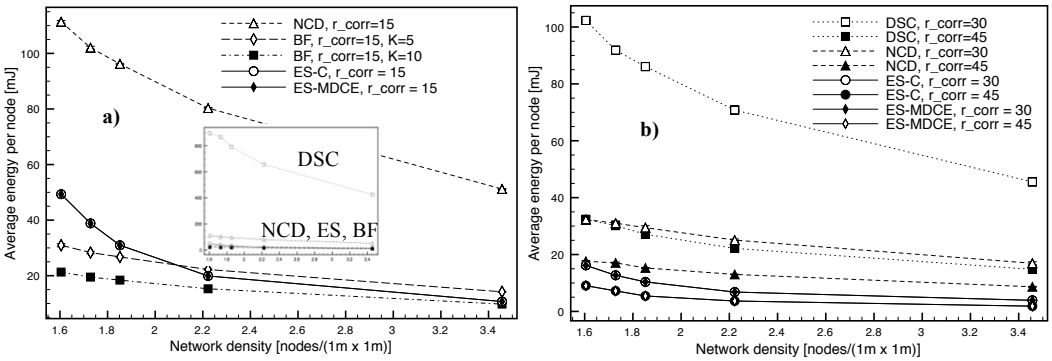

Figure 9. Energy consumption for 150 rounds and different r_{corr} a) r_{corr}=15m; inset: DSC performancs; b) r_{corr}=30; 45m. Energy consumption decreases as the correlation radius increases (number of representative nodes decreases). The performance of ES-MDCE performs is identical to the basic ES-C encoding scheme.

first node in the sequence is allocated $R_1 = H(X_1)$; the second is allocated $R_2 = H(X_2|X_1)$, etc.; the last node in the neighborhood is allocated rate: $R_K = H(X_K|X_{K-1},\ldots, X_1)$, assuming $|C_i| = K$. We set $r_i = r_{corr}$ around each representative node.

- The *BF* scheme evaluated here is based on the cooperative distributed transmit beamforming scheme suggested in [Mudumbai et al. 2010]. An assumption in [Mudumbai et al. 2010] is that upon each transmission, cooperating nodes are synchronized in frequency with each other. Furthermore they need to be synchronized in phase so that the gains of beam forming are realized due to constructive interference at the sink. Satisfying the latter assumption, a simple randomized algorithm is offered in [Mudumbai et al. 2010] to achieve phase coherency of the transmitted bits at the sink. Similarly to the above schemes, the beamforming groups G_i are formed within correlation radius r_{corr} around the k representative nodes.

5.3.3. Energy Efficiency and Estimate Inaccuracy. All schemes are simulated and compared in terms of two metrics: *energy consumption* and *estimate* inaccuracy at the sink. *Energy consumption* is given by the average energy consumption E per node in the network for τ reporting rounds. *Estimate inaccuracy* is given by the average inaccuracy of the phenomenon point source estimate at the sink over τ rounds:

$$\overline{V} = (1/\tau)\sum_{i=1}^{\tau} V_i$$

where at round i, $V_i = 100|s_i - s'_i|/M$. s_i is the true value of $\boldsymbol{S}$ at (x_i, y_i), and s'_i is the estimated measurement at the sink given the report of representative group G_i, or in the case of *NCD* the representative node i.

The average energy consumption per network node E for ES tends to be less than that of NCD and DSC as shown on Figure 9 and comparable to the performance of distributed beamforming with $K = 5$ cooperative beamforming nodes. As expected, both ES-MCDE and ES-C perform identically in terms of energy efficiency. When more nodes are added to the network (in effect increasing its density), as expected E decreases for all schemes. However, both ES schemes benefit substantially more from the network density increase, compared to DSC and NCD. Larger network density allows for smaller values of K^* leading to a lower number of actively transmitting nodes in $|A_x|$ required to convey measurement x. At network densities of 1.5 nodes/m^2, ES consumes 2 times less the transmission energy of NCD; at 1.9 nodes/m^2, the energy is 3 times less; and *nearly 5 times* less at densities around 3.5 nodes/m^2 depending on the correlation radii. $|G_i|$ grows from 15 to 33 nodes as density increases (notice that $|G_i|$ is always under acquisition capacity) and $|A_x|$ decreases from 7 to 3 nodes.

BF also benefits from the addition of cooperative nodes. Evaluating the performance of BF, the phase synchronization feedback algorithm is successfully run at each beamforming node to achieve 70% phase coherency at the receiver (which is equivalent to achieving about 80% of beamforming gains). Operating below its theoretical efficiency, BF outperforms the ES for the case of 10 cooperating nodes in the region of lower network densities. However, the synchronization assumptions of ES are much more relaxed than those of the BF scheme. In reality, cooperating more than 10 nodes would pose a challenge to any practical distributed transmit beamforming system.

Interestingly, DSC compares rather poorly with the rest of the schemes in terms of energy efficiency. This is so since all nodes within a given neighborhood with radius r_{corr} transmit with positive rates, allocated to account for nodes observations' correlation. The total rate equals at least the joint entropy of the observed measurements and is higher than the rest of the schemes.

Figure 10 demonstrates that the inaccuracies in the source estimates are rather similar for ES, NCD, and BF. This is expected in the case of ES and NCD since as shown in the Appendix the distortion of the two schemes numerically converges. All four schemes report information only about the value of the phenomenon at the representative nodes i. However, ES-MCDE outperforms slightly ES-C, achieving lower estimate inaccuracy. DSC obtains slightly lower inaccuracy, since all nodes in each r_{corr} neighborhood around a representative node send information regarding the sensed value at their position.

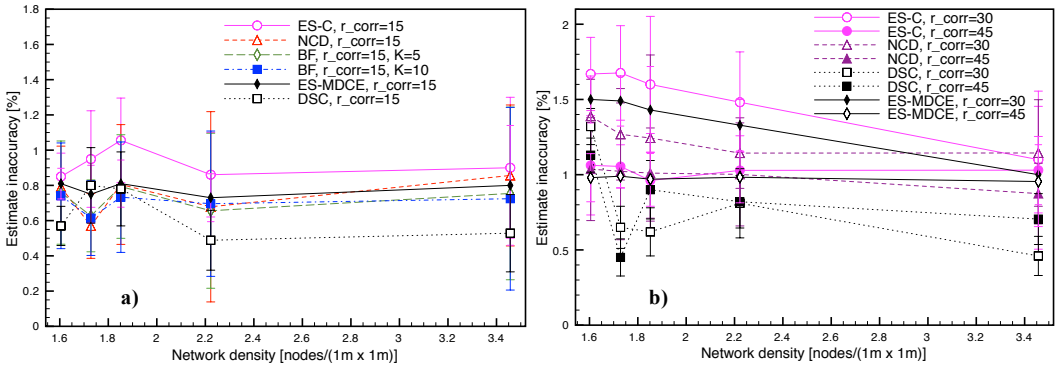

Figure 10. Estimate inaccuracy; a) r_{corr} = 15m; b) r_{corr} = 30; 45m. ES-MDCE achieves lower inaccuracy and lower variance than the basic ES-C encoding scheme. However, all four scheme estimate the source relatively accurately. The measurements at the representative sites found via vector quantization model well the phenomenon.

6. ENCODING SENSING: NUMBER OF TRANSMISSIONS AND ENERGY EFFICIENCY

In this section, we derive a bound on the minimum required number of nodes in G, for the ES operation. The analysis below is valid for both ES-C and ES-MCDE. This also allows us to obtain a bound on the number of transmitting nodes in A_x for the worst practical case of ES operation, providing a quantitative argument for at least twice better energy efficient performance of ES vs. NCD in the simulation evaluation.

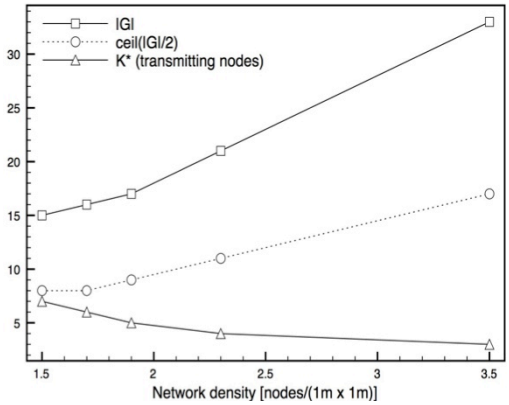

Figure 11. Number of nodes in a group G_i and number of transmitting nodes K^* for different network densities. Even at lower densities, K^* is less than $|G|/2$, the worst-case of ES operation, and drops further with density. K^* is the same for both ES-MDCE and ES-C.

In (2.2) we require that ${}^{|G|}C_K \geq \mathrm{M}$, where $K = |A_x|$. If the group size $|G|$ is fixed, ${}^{|G|}C_K$ is maximized for $K = |G|/2$. What is the least value of $|G|$ that guarantees ${}^{|G|}C_{|G|/2} \geq \mathrm{M}$? Using Sterling's approximation we get

$$ {}^{|G|}C_{|G|/2} \geq 2^{|G|-1} / \sqrt{|G|/2} $$

$2^{|G|-1} / \sqrt{|G|/2} \geq M$ implies ${}^{|G|}C_{|G|/2} M$. Solving for $|G|$:

$$ 0.5|G| \geq \log(0.5|G|)/4 + \log(M)/2 + 0.5 $$

We also notice $\log(0.5|G|)/4 \leq 1$ for practical values of $|G|$ (up to $|G| = 32$ nodes[7]). Then, we obtain that

$$ |G| \geq 2\lceil \log(M)/2+1.5 \rceil \qquad (6.1) $$

Satisfying (6.1) ensures there are enough codewords for each message in M. Here, $K^* = |G|/2$. This is the maximum possible number of nodes in the set A_x, largest codeword weight, and the worst case of ES operation. Allowing larger group sizes $|G|$ substantially reduces the value of K^* utilized in ES as shown in Figure 11. In practice, $K^* << |G|/2$ is sufficient to encode M messages, as we observe in the example WSN application of section 5.

We can now compare the worst-case energy efficiency of ES vs. the average energy efficiency of NCD. Since each node sending a bit per ES needs to be identified by the sink, we consider a DSSS physical layer as outlined above. The energy E_b consumed per bit transmission is spread over a PN sequence.

In the case of ES, we do not need to detect the sign of the signal (i.e. we are not interested whether the transmitted bit is 0 or 1 as long as some signal is there). Suppose $K^* = |G|/2 = \lceil \log(M)/2+1.5 \rceil = |c_x|$ bits are transmitted per message, as in the worst case of ES operation from (6.1).

In contrast, the NCD scheme over the same underlying DSSS physical layer, would require transmission of a single representative node from G, sending $\log(M)$ bits. Each of these bits, as standard in single user SS systems, is spread over a PN sequence and modulated antipodally. The value of each bit, 0 or 1, has to be detected.

[7] If $|G| = 32$, ${}^{|G|}C_{|G|/2} \approx 6*10^8$ distinct messages can be mapped, which is sufficient for many applications in WSN, for instance.

Given a transmission energy budget of E each, ES and the NCD schemes can send respectively d_{ES} and d_{NCD} messages in total: $d_{ES} = E/(|c_x|E_b)$ and $d_{NCD} = E/[\log(M)E_b]$. In the worst-case for ES we have $|c_x| = \lceil \log(M)/2+1.5 \rceil$ and

$$d_{ES} = 2d_{NCD}/(1+3/\log(M)) \quad (6.2)$$

Thus, ES in its worst regime of operation is close to 2 times as energy efficient as NCD or a corresponding duty-cycling scheme. Correspondingly, the ES scheme achieves at least a factor of 2 better energy efficiency than NCD in simulations. Suppose the NCD scheme utilizes different, for instance non-antipodal, modulation so that in effect only $\lceil \log(M)/2 \rceil$ signals are transmitted. Still, the energy per bit required to achieve the bit error rate obtained via the former antipodal modulation would need to at least double (for instance, if orthogonal instead of antipodal modulation is used).

7. RELATED WORK

One of the first and most prominent examples of collaboration across multiple wireless sensor nodes is cooperative communication [Nosrania et al. 2004]. Inspired by the gains of spatial diversity realized in MIMO systems [Cui et al. 2006], cooperative communication in the context of WSN utilizes a set of nodes each equipped with a single antenna. The nodes transmit cooperatively (portions of) the same message to a sink. Since different nodes have different spatial coordinates, the sink receives multiple instances of the same message that have suffered fading over statistically independent spatial paths. The earliest realizations of cooperative transmission utilize the decode-and-forward [Sendonaris et al. 2003] and the amplify-and-forward mechanisms [Laneman et al. 2001]. Various schemes for optimal selection of cooperating nodes [Simic et al. 2008]; algorithms for cooperative space-time coding of the transmitted messages [Li et al. 2005]; and cooperative game theoretic frameworks optimizing network quality of service per given energy budget [Wu et al. 2012] have been further investigated. Encoded sensing is distinct from such cooperative communication strategies. Per ES, each node in the subset A_x of G transmits a single bit directly to the sink, without additional communication across collaborating nodes.

ES does not rely on spatial diversity in order to achieve energy efficient communication. This is in contrast to more recent distributed transmit beam forming systems (BF) [Mudumbai et al. 2010]. Per BF, a set of cooperating nodes simultaneously transmit copies of the same message to a single sink[8]. The signals of the cooperating nodes are specifically synchronized in phase and frequency resulting in constructive interference at the sink. Theoretically, employing n cooperative beam-forming nodes could lead up to a factor of n energy savings compared to non-cooperative signaling. However, in the most recent proof-of-concept beamforming systems, noiseless 1-bit feedback from the sink is assumed so that all cooperatively transmitting nodes are approximately synchronized in phase. This feedback is provided via a separate cable connection [Mudumbai et al. 2007; Mudumbai et al. 2009] to the sink; this is often not feasible in practice. Although ES also requires that cooperating nodes have access to the same message, it does not require as fine synchronization of message transmissions as beamforming.

[8] These systems need to be differentiated from distributed multi-user multiple-input, multiple-output (MU-MIMO) beamforming systems such as [29] where a si transmits to many devices.

A somewhat separate stream of technical literature discusses the collaborative operation of nodes in WSN taking into account correlations in measurements. For example, a scheme is introduced in [Vuran and Akyildiz 2006] eliminating the transmission of measurements that are highly correlated. As shown in [Vuran and Akyildiz 2006], the elimination of certain highly correlated messages would not distort (within a bound) the overall phenomenon estimate at the sink. Thus network energy is conserved. The outcome of the scheme is that within a group of nodes measuring the same (or very similar) values, only a single node transmits an entire measurement value at every timeslot.

In a similar spirit, techniques relying on distributed source coding have been suggested for WSN. The authors of [Critescu et al.] and [Yuan et al. 2008] leverage results based on the Slepian-Wolf source-coding theorem. Namely, correlated sources can be compressed separately (e.g. without communication between the source nodes) and without loss to the level of their joint entropy. Consequently, according to a pre-specified scheme, each sensor node in the network transmits fewer bits the more correlated its measurement is to reference nodes.

8. DISCUSSION

The encoded sensing scheme analyzed in this paper employs a specific combinatorial assignment of nodes to an index of quantized sensory input. The resulting mechanism for distributed and sparse representation of sensory data allows for guaranteed minimal energy expended on transmissions, while maintaining source fidelity at the sink over time. The ES sparse representation of sensory data allowed us to design a novel sparse DSSS receiver significantly reducing the circuit complexity (and hence cost) of the standard DSSS receiver. In a number of non-ES networks engineered around DSSS, only a small fraction of the potential system users are active at a time. This generates analogous signal sparseness at the sink/base station. The challenge, unlike in an ES system, is that the signal is not guaranteed to be K-sparse but only approximately so. Engineering compressive sensing-based DSSS system in the face of sparseness uncertainty may present avenues for further research.

Due to space limitations we did not consider the use of ES in a multihop network setting in this work. However, ES can easily be adjusted for use as a cooperative transmission layer of various cooperative routing schemes such as the one in [Elhawary and Haas 2011]. The broadcast nature of wireless transmissions causes a number of nodes to receive the same message x after a neighboring transmission. These nodes can cooperate using ES to forward x onto the next hop of a routing path. We believe coordinating such clusters of forwarding nodes and ensuring all nodes in a cluster receive the same message w.h.p. are feasible and deserve further investigation.

Finally, codes different then the combinatorial one presented in this paper can be used in ES. An intriguing code would be one that provides graceful degradation of message fidelity. I.e. given a node in A_x erroneously transmits a bit (or remains silent), the received codeword should be guaranteed to yield only a slightly degraded version of x when decoded.

APPENDIX

Distortion of measurement per ES and NCD

The estimate of the source **S** at the sink produced utilizing either ES or NCD is distorted since 1) only information regarding the measurements of k representative out of the N nodes in the network is taken into account; 2) there is channel noise and sensing imprecision noise present. The CC-MAC algorithm in [Vuran and Akyildiz 2006] utilized for representative node selection both by NCD and ES ensures that the number of selected k nodes is minimized so that the distortion $D(k)$ is within a quality of service constraint D_{QoS}. We demonstrate that ES achieves similar level of distortion. For both schemes the distortion is given by standard Minimum Square Error (MSE) metric:

$$D(k)= E[(\mathbf{S} - \mathbf{S}')^2] = E[\mathbf{S}^2] - 2\,E[\mathbf{S}\,\mathbf{S}'] + E[\mathbf{S}'^2] \quad \text{(A.1)}$$

where **S** is the value of the point source; **S**' is the estimate of the point source at the sink given the k reports. Suppose the channel noise is AWGN, $Z_j \sim N(0, \sigma_Z^2)$.

For a non-cooperative transmission scheme such as NCD, the measurement X_j of a representative node j is transmitted most efficiently using uncoded transmission [Gastpar and Vetterly 2003] subject to power constraint P per node, per measurement. The received signal at the sink is

$$Y_j=\sqrt{P/(\sigma_S^2+\sigma_W^2)}*(S_j+W_j)+Z_j .$$

The optimal decoder at the sink is given by the standard MMSE estimator [Gastpar and Vetterly 2003]:

$$S'_j = Y_j(E[S_j\ Y_j]/E[Y_j])$$

and

$$\boldsymbol{S}' = (1/k)\sum_{j=1}^{k} S'_j.$$

Evaluating (A.1) term by term and simplifying yields

$$D_{NCD}(k)=D(k,P)=\sigma_S^2-\sigma_S^4[(\sigma_S^2+\sigma_W^2)(1+\sigma_Z^2/P)]^{-1}\times \left[\frac{1}{k}\left(2\sum_{i=1}^{k}\rho_{S,i}-1\right)-\frac{\sigma_S^2}{k^2}\sum_{i=1}^{k}\sum_{j=1,j\neq i}^{k}\frac{\rho_{i,j}}{\sigma_S^2+\sigma_W^2}-\frac{(1-1/k)\sigma_Z^2}{P}\right] \quad \text{(A.2)}$$

Per ES, $|A_x| = K$ nodes transmit for each of the k representative groups. The nodes transmit in a multi-access interference limited system utilizing DSSS. The channel noise is negligible in comparison to nodes' mutual interference. However, since the system operates at acquisition capacity, the probability of transmission error due to interference is close to 0 ([Madhow and Pursley 1993]) and does not incur distortion of measurement. Here, we have assumed the range of values a phenomenon can take is split in intervals of length $\varepsilon = 2\beta\sigma_W$, $\beta > 2$. Hence, per each group G_j, the estimation of measurement at the sink is $S'_j = S_j + 2\beta\sigma_W$.

As before, $\mathbf{S}' = (1/k)\ \sum_{j=1}^{k} S'_j$ and plugging in (A.1):

$$D_{ES}(k)=\sigma_S^2\left[1-\frac{1}{k}\left(2\sum_{i=1}^{k}\rho_{S,i}-1\right)+\frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\rho_{i,j}+\left(\frac{2\beta\sigma_W}{\sigma_S}\right)^2\right] \quad \text{(A.3)}$$

Aside from the term $\left(2\beta\sigma_W/\sigma_S\right)^2$, as shown in Figure 7, ES achieves similar distortion to NCD as the number of representative groups/nodes increases and $\sigma_W < \sigma_S$.

REFERENCES

Victor Shnayder, Mark Hempstead, Bor-rong Chen, Geoff Werner Allen, and Matt Welsh. 2004. Simulating the power consumption of large-scale sensor network applications. In *Proceedings of the 2nd international conference on Embedded networked sensor systems* (SenSys '04). ACM, New York, NY, USA, 188-200. DOI=10.1145/1031495.1031518 http://doi.acm.org/10.1145/1031495.1031518

O. Landsiedel, K. Wehrle, and S. Gotz. 2005. Accurate prediction of power consumption in sensor networks. In *Proceedings of the 2nd IEEE workshop on Embedded Networked Sensors* (EmNets '05). IEEE Computer Society, Washington, DC, USA, 37-44.

Trevor Pering, Yuvraj Agarwal, Rajesh Gupta, and Roy Want. 2006. *CoolSpots*: reducing the power consumption of wireless mobile devices with multiple radio interfaces. In *Proceedings of the 4th international conference on Mobile systems, applications and services* (MobiSys '06). ACM, New York, NY, USA, 220-232. DOI=10.1145/1134680.1134704 http://doi.acm.org/10.1145/1134680.1134704

Bhaskar Krishnamachari, Deborah Estrin, and Stephen B. Wicker. 2002. The Impact of Data Aggregation in Wireless Sensor Networks. In *Proceedings of the 22nd International Conference on Distributed Computing Systems* (ICDCSW '02). IEEE Computer Society, Washington, DC, USA, 575-578.

Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. 2000. Directed diffusion: a scalable and robust communication paradigm for sensor networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking* (MobiCom '00). ACM, New York, NY, USA, 56-67. DOI=10.1145/345910.345920 http://doi.acm.org/10.1145/345910.345920

Mehmet C. Vuran and Ian F. Akyildiz. 2006. Spatial correlation-based collaborative medium access control in wireless sensor networks. *IEEE/ACM Trans. Netw.* 14, 2 (April 2006), 316-329. DOI=10.1109/TENT.2006.872544 http://dx.doi.org/10.1109/TENT.2006.872544

Raghuraman Mudumbai, Joao Hespanha, Upamanyu Madhow, and Gwen Barriac. 2010. Distributed transmit beamforming using feedback control. *IEEE Trans. Inf. Theor.* 56, 1 (January 2010), 411-426. DOI=10.1109/TIT.2009.2034786 http://dx.doi.org/10.1109/TIT.2009.2034786

Fatma Bouabdallah, Nizar Bouabdallah, and Raouf Boutaba. 2008. Toward Reliable and Efficient Reporting in Wireless Sensor Networks. *IEEE Transactions on Mobile Computing* 7, 8 (August 2008), 978-994. DOI=10.1109/TMC.2008.21 http://dx.doi.org/10.1109/TMC.2008.21

A.E. Khandani, J. Abounadi, E. Modiano, L. Zheng, "Cooperative Routing in Static Wireless Networks," *IEEE Trans. Communications* 55, 11 (November 2007), 2185-2192, DOI=10.1109/TCOMM.2007.908538

Mohamed Elhawary and Zygmunt J. Haas. 2011. Energy-efficient protocol for cooperative networks. *IEEE/ACM Trans. Netw.* 19, 2 (April 2011), 561-574. DOI=10.1109/TNET.2010.2089803 http://dx.doi.org/10.1109/TNET.2010.2089803

R. Cristescu, B. Beferull-Lozano, and M. Vetterli. 2005. Networked Slepian-Wolf: theory, algorithms, and scaling laws. *IEEE Trans. Inf. Theor.* 51, 12 (December 2005), 4057-4073. DOI=10.1109/TIT.2005.858980 http://dx.doi.org/10.1109/TIT.2005.858980

K. Yuen, Ben Liang and Baochun Li. 2008. A Distributed Framework for Correlated Data Gathering in Sensor Networks. *IEEE Trans. Vehicular Tech.* 57, 1 (January 2008), 578-593, DOI=10.1109/TVT.2007.905243

A. Nosratinia, T. E. Hunter, and A. Hedayat. 2004. Cooperative communication in wireless networks. *Comm. Mag.* 42, 10 (October 2004), 74-80. DOI=10.1109/MCOM.2004.1341264 http://dx.doi.org/10.1109/MCOM.2004.1341264

Shuguang Cui, A. J. Goldsmith, and A. Bahai. 2006. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks. *IEEE J.Sel. A. Commun.* 22, 6 (September 2006), 1089-1098. DOI=10.1109/JSAC.2004.830916 http://dx.doi.org/10.1109/JSAC.2004.830916

A. Sendonaris, E. Erkip, and B. Aazhang. 2003. User cooperation diversity--part I: system description. *IEEE Trans. Comm.* 55, 11 (November 2003), 1927-1938. DOI=10.1109/TCOMM.2003.819238

J. N. Laneman, D. N.C. Tse, and G. W. Wornell. 2004. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Trans. Inf. Theor.* 50, 12 (December 2004), 3062-3080. DOI=10.1109/TIT.2004.838089 http://dx.doi.org/10.1109/TIT.2004.838089

L. Simic, S.M. Berber, K.W. Sowerby. (2008). Partner Choice and Power Allocation for Energy Efficient Cooperation in Wireless Sensor Networks. *In Proceedings of the IEEE International Conference on Communications* (ICC'08). DOI=10.1109/ICC.2008.799

Xiaohua Li, Mo Chen, Wenyu Liu. (2005). Application of STBC-encoded cooperative transmissions in wireless sensor networks. *IEEE Signal Processing Letters* 12, 2 (February 2005). 134-137. DOI=

10.1109/LSP.2004.840870(410) 12

Dan Wu, Yueming Cai, Liang Zhou, Jinlong Wang. (2012). A Cooperative Communication Scheme Based on Coalition Formation Game in Clustered Wireless Sensor Networks. *IEEE Trans. Wireless Comm.* 11, 3 (March 2012) 1190-1200, DOI=10.1109/TWC.2012.012712.111049

Raghuraman Mudumbai, D. Richard Brown, Upamanyu Madhow, and H. Vincent Poor. 2009. Distributed transmit beamforming: challenges and recent progress. *Comm. Mag.* 47, 2 (February 2009), 102-110. DOI=10.1109/MCOM.2009.4785387 http://dx.doi.org/10.1109/MCOM.2009.4785387

Charles Chien, Igor Elgorriaga, and Charles McConaghy. 2001. Low-power direct-sequence spread-spectrum modem architecture for distributed wireless sensor networks. In *Proceedings of the 2001 international symposium on Low power electronics and design* (ISLPED '01). ACM, New York, NY, USA, 251-254. DOI=10.1145/383082.383151 http://doi.acm.org/10.1145/383082.383151

L. Song and J. Shen, *Evolved Cellular Network Planning and Optimization for UMTS and LTE*. CRC Press, 2010. ISBN-13: 978-1439806494

D.H. Lehmer. The machine tools of combinatorics. In Applied Combinatorial Mathematics, E.F. Beckenback, Ed., Wiley, New York, 1964, pp. 5-31.

Andrew J. Viterbi. 1995. *CDMA: Principles of Spread Spectrum Communication*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.

James O. Berger, Victor De Oliveira, and Bruno Sansó. (2001). Objective bayesian analysis of spatially correlated data. *Journal of the American Statistical Association*, 96, 456 (December 2001), 1361-1374. DOI = 10.1198/016214501753382282

Bayesian Logic (BLOG) Inference Engine. Retrieved June 1, 2014 from http://people.csail.mit.edu/milch/blog/index.html

Michael Gastpar and Martin Vetterli. 2003. Source-channel communication in sensor networks. In *Proceedings of the 2nd international conference on Information processing in sensor networks* (IPSN'03), Feng Zhao and Leonidas Guibas (Eds.). Springer-Verlag, Berlin, Heidelberg, 162-177.

S. Verdu. 1986. Minimum probability of error for asynchronous Gaussian multiple-access channels. *IEEE Trans. Inf. Theor.* 32, 1 (January 1986), 85-96. DOI=10.1109/TIT.1986.1057121 http://dx.doi.org/10.1109/TIT.1986.1057121.

Andrea Goldsmith. 2005. *Wireless Communications*. Cambridge University Press, New York, NY, USA.

Gordon L. Stüber. 2001. *Principles of mobile communication (2nd ed.)*. Kluwer Academic Publishers Norwell, MA, USA.

H. V. Poor and S. Verdu. 1997. Probability of error in MMSE multiuser detection. *IEEE Trans. Inf. Theor.* 43, 3 (September 1997), 858-871. DOI=10.1109/18.568697 http://dx.doi.org/10.1109/18.568697

U. Madhow and M. B. Pursley. 1993. Acquisition in direct-sequence spread-spectrum communication networks: an asymptotic analysis. *IEEE Trans. Inf. Theor.* 39, 3 (May 1993), 903-912. DOI=10.1109/18.256498 http://dx.doi.org/10.1109/18.256498

P.D. Robinson and A.J. Wathen, "Variational bounds on the entries of the inverse of a matrix," *IMA J. Numer. Anal.* **12** (4) (1992), pp. 463–486.

Gene H. Golub and Charles F. Van Loan. 1996. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA.

Clayton Shepard, Hang Yu, Narendra Anand, Erran Li, Thomas Marzetta, Richard Yang, and Lin Zhong. 2012. Argos: practical many-antenna base stations. In *Proceedings of the 18th annual international conference on Mobile computing and networking* (Mobicom '12). ACM, New York, NY, USA, 53-64. DOI=10.1145/2348543.2348553 http://doi.acm.org/10.1145/2348543.2348553

R. Baraniuk. (2007). Compressive sensing. *IEEE Signal Processing Magazine* 24, 4 (July 2007), 118-121. DOI=10.1109/MSP.2007.4286571

D. L. Donoho. 2006. Compressed sensing. *IEEE Trans. Inf. Theor.* 52, 4 (April 2006), 1289-1306. DOI=10.1109/TIT.2006.871582 http://dx.doi.org/10.1109/TIT.2006.871582

J. A. Tropp and A. C. Gilbert. 2007. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Trans. Inf. Theor.* 53, 12 (December 2007), 4655-4666. DOI=10.1109/TIT.2007.909108 http://dx.doi.org/10.1109/TIT.2007.909108

Kush R. Varshney and Peter M. van de Ven. 2013. Balancing lifetime and classification accuracy of wireless sensor networks. In *Proceedings of the fourteenth ACM international symposium on Mobile ad hoc networking and computing* (MobiHoc '13). ACM, New York, NY, USA, 31-38. DOI=10.1145/2491288.2491289 http://doi.acm.org/10.1145/2491288.2491289

David Tse and Pramod Viswanath. 2005. *Fundamentals of Wireless Communication*. Cambridge University Press, New York, NY, USA.