Encoded Sensing for Energy Efficient Wireless Sensor Networks

Milen Nikolov and Zygmunt J. Haas, *Fellow, IEEE*

***Abstract*— Energy efficient communication is a fundamental design problem in wireless networks significantly affecting network performance and the lifetime of wireless sensor networks (WSNs). We introduce encoded sensing—an approach for collaborative encoding and transmission of sensors data—that drastically reduces communication energy expenditure in WSN. Encoded sensing exploits the inherent spatial structure in sensed data to adaptively partition a WSN into groups of sensor nodes, so that nodes in each group sense highly correlated values. Each group encodes all individual measurements sensed by its nodes at time t into a single binary sparse codeword via novel minimum distance combinatorial encoding local algorithm. When the codeword's Hamming weight equals w, a subset of w nodes in the group cooperatively transmits a single binary symbol each. Upon receiving the w bits, the sink has enough information to decode a measurement estimate, which is within a small error from each of the group nodes' individual measurements. The error is bounded and guaranteed to satisfy a priori QoS accuracy requirements. We compare encoded sensing to non-cooperative state-of-the-art transmission protocols and demonstrate at least a factor of two in energy savings, without significant loss of measurement quality. Encoded sensing achieves at least 80% the energy savings of theoretically optimal cooperative transmission distributed beamforming architectures. We show by simulations and theoretical derivations that as the size of a node group grows the performance of encoded sensing converges to the optimal transmission energy efficiency.**

***Index Terms*— Energy efficiency, cooperative transmissions, distributed beam forming, spatial correlation, reporting, encoded sensing, compressive sensing, minimum distance combinatorial encoding.**

I. Introduction

COMMUNICATING data is among the most energy expensive operations across different applications of WSN; receiving and transmitting data constitutes more than 60% of the total energy consumption on networks of small wireless devices running over popular interfaces including Zigbee/IEEE 802.15.4 or IEEE 802.11a/b (e.g. [1], [2], [3]). Reducing the amount of data transmitted and/or energy consumed per transmission could lead to significantly longer network lifetime and reduced costs of nodes redeployment, maintenance, and network outages. Accordingly, there is a significant body of technical literature that covers different aspects of the topic. As noted in a comprehensive survey on energy efficiency in WSN ([4]), there are three major types of techniques aiming to decrease network radio transmissions and save energy in WSN: *sampling compression*, *data compression*, and *communication compression*.

Sampling compression schemes for WSN exploit inherent structure of the observed phenomenon to *reduce the number of measurements* required for its accurate description. The reduction of measurements results in reduction of transmitted data and allows for effective duty-cycling. Examples of sampling compression include schemes eliminating measurements and hence transmissions of nodes with highly correlated observations due to sensors' spatial proximity ([5], [6]). Elimination of such correlated nodes' transmission would not distort (within a bound) the overall phenomenon estimate at the sink. A more recent development of sampling compression schemes utilizes the theory of compressive sensing. Those schemes rely on the fact that often physical phenomena can be accurately described by signals possessing sparse representation in a certain basis. In the context of WSN the signals may have sparse representation over time at a given sensor node (intra-signal sparse structure, e.g. [7]–[9]) and/or over space across the measurements of different sensor nodes at any fixed time (inter-signal sparse structure, e.g. [10]). In either case, $O(K \log(n))$ measurements and transmissions received at the sink are sufficient to reconstruct a set of n phenomenon realizations comprising $K \ll n$ non-zero values in a certain projection.

Data compression for WSN is subtly different: given all available (potentially redundant) sensors measurements aggregated over some time window, nodes transmit a *data set of size commensurate with the compression entropy of the measurements*. Distributed Source Coding (DSC) schemes leverage the Slepian-Wolf source-coding theorem and allow compression of correlated sources without communication between the sensor nodes and without loss of their measurements' joint entropy. The latter is advantageous in distributed systems like WSN, and Cristescu *et al.* [11] and Yuen *et al.* [12] analyze the application of DSC to WSN, showing that each sensor node in the network transmits optimally less bits the more similar its measurement is to that of other transmitting reference nodes. Using such schemes, there is no loss of measurement accuracy.

In parallel to solutions exploiting the redundant structure of observed phenomena to maximize reduction in measurements/transmissions, communication protocols with the objective of transmission energy cost minimization have been investigated as well. For instance, transmission of longer

Manuscript received March 5, 2017; revised July 3, 2017; accepted September 10, 2017. Date of publication September 26, 2017; date of current version December 21, 2017. This work was supported by the National Science Foundation under Grant CNS-1040689 and Grant ECCS-1308208. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dario Pompili. *(Corresponding author: Milen Nikolov.)*

The authors are with the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: mvn22@cornell.edu; haas@ece.cornell.edu).

Digital Object Identifier 10.1109/JSEN.2017.2756865

packets requires longer radio on-time, which inherently depletes more energy. Furthermore, longer packets are prone to larger packet error rate, which in turn leads to larger communication cost due to packet re-transmissions. Reduced packet size, i.e. *communication compression*, leads to shorter radio on-time and lower communication cost ([5]).

Finally, energy reduction schemes that use data compression, sampling compression, and communication compression have been supplemented by a fourth orthogonal set of *distributed cooperative transmission* solutions, recently becoming increasingly feasible in the context of WSN. Unlike the three former compression types, cooperative transmission schemes do not explicitly rely on reducing the redundancy or number of transmitted bits. On the contrary, inspired by the gains of spatial diversity realized in MIMO[1] systems ([13]), cooperative communication in the context of WSN utilizes a set of nodes each equipped with a single antenna and each transmitting 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, resulting in transmission diversity gains first explored in [14]–[16]. The signal gains translate directly into energy savings per message transmission. Various schemes for optimal selection of cooperating nodes ([17]); algorithms for cooperative space-time coding of the transmitted messages ([18]); and cooperative game theoretic frameworks optimizing network quality of service per given energy budget ([19]) have provided further practical and theoretical underpinnings for cooperative transmission applications in WSN. The theoretical promise of cooperative transmission exploiting spatial diversity is significant. For instance, distributed transmit beam forming (BF) architectures (e.g. [20], [21]) employing n cooperative beamforming nodes theoretically result in a fundamentally optimal factor of n in energy savings, as compared to non-cooperative transmission.

Combining cooperative transmission gains with sampling compression, data compression, and communication compression would potentially lead to drastically lower overall communication energy consumption. However, data− and sampling−compression techniques operate on minimizing the transmission of redundant information, while redundancy is at the core of cooperative transmission's spatial diversity gains.

Is there a collaborative signaling architecture that can achieve the gains of cooperative transmission, but without relying on redundant messages/spatial diversity and remaining compatible with compression schemes? In this work we answer this question in the affirmative, by introducing the Encoded Sensing (ES) scheme.

A. The Basic Idea Behind Encoded Sensing

Suppose a WSN measures a natural phenomenon like temperature, soil moisture, humidity, solar radiation, etc., over a range of values $[a, b]$, e.g. [−30 °C, 50 °C]. Per encoded sensing (ES), each node quantizes the range $[a, b]$ in equal intervals of length ε. Each interval is uniquely labeled and mapped to an index $x = 1, 2, \ldots, M$ (Fig. 1).

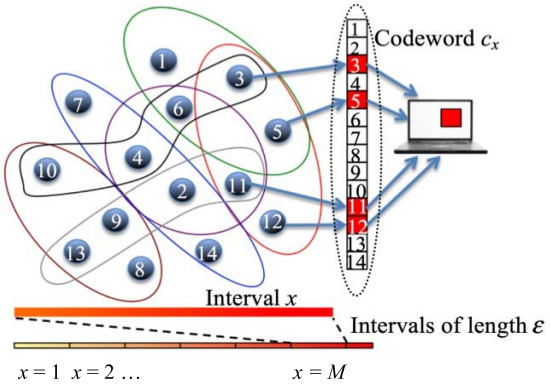

Fig. 1. Encoded sensing (ES) simplified schematic - single measurement transmission: at time t all 14 nodes in the group G sense a value in the same unique small interval x (in solid red gradient). Only the 4 nodes (circled in red line) that are "assigned" to the red interval x transmit to the sink; the rest 10 nodes assigned to different intervals remain silent, resulting in a sparse codeword c_x. The sink decodes c_x and recovers the unique interval x. Since all the node measurements in the group G fall within interval x, measurement uncertainty after decoding is bounded by x's length. Given the assumed spatial correlation of the 14 nodes in G, only 4 binary symbols are sufficient to encode 1001 distinct intervals using MDCE. To achieve similar measurement accuracy a non-cooperative scheme relying on optimal spatial decorrelation would require transmission of 10 symbols at the same energy per symbol. This results in ∼2.5 times less energy consumption exploiting ES *coding diversity*.

As we will discuss in the sequel, ES leverages sampling compression, data compression, and communication compression, while introducing a novel collaborative transmission scheme based on *coding diversity*, in lieu of spatial diversity previously exploited in cooperative transmission schemes.

ES exploits *data compression and sampling compression* by adaptively partitioning a WSN into groups of sensor nodes, with highly correlated measurements, so that in each group G the maximum difference between any two nodes' measurements cannot be larger than ε. Nodes in G are assigned unique IDs: $0, 1, \ldots, |G| - 1$ as shown in Fig. 1. As part of the ES collaborative transmission scheme, each group G encodes all individual measurements sensed by its nodes at time t into a single binary sparse codeword via novel Minimum Distance Combinatorial Encoding (MDCE) local algorithm, running at each node. MDCE assigns each interval x to a unique set of nodes $A_x \subset G$. Given that nodes in G measure values falling within interval x at time t, each of the nodes in $A_x \subset G$ assigned to x contributes a "1" at a specific location in the codeword c_x; each of the nodes in G/A_x contributes a "0" at a specific location in c_x as shown in Fig. 1 and according to the MDCE. The codeword's Hamming weight equals w. We show that upon receiving the right set of w binary symbols, the sink has enough information to decode the measurement estimate within a small absolute error, less than 2ε away from each of the group nodes' individual measurements. Inherently, ES MDCE-based collaborative transmission employs lossy *data compression* by minimizing w as a function of the number of nodes in a group and ε: as the number of nodes in a group G

[1]These systems need to be differentiated from distributed multi-user multiple-input, multiple-output (MU-MIMO) beamforming systems such as in [22] where a sink transmits to many devices.

increases, w decreases, thus reducing the amount of required sent/received bits per measurement, for a fixed ε.

Since only one codeword c_x per group is transmitted at time t to convey WSN's measurements to the sink node, the amount of data transmitted is reduced as the number of groups in the WSN decreases. We show that the number of groups transmitting measurements per ES is minimized, resulting in *sampling compression* similar to [5].

To transmit the codeword c_x, the w nodes that have contributed a "1" to c_x transmit one binary symbol each, per transmitted measurement; the remaining nodes remain silent as shown in Fig. 1. Since only one symbol is transmitted per node, *communication compression* in terms of packet lengths/transmission errors is optimized, too.

ES collaborative transmission is inherently coupled with compression schemes, and the achieved energy savings compared to state-of-the-art communication architectures relying only either on compression or on cooperative transmission are significant. We compare ES energy efficiency and measurements accuracy to three state-of-the-art WSN communication architectures representative of sampling compression ([5]), data compression ([11], [12]), and cooperative transmission ([20], [21]), respectively. We show at least a factor of 2 in energy savings for ES across all comparisons, without significant loss of measurement quality. Encoded sensing achieves at least 80% the energy savings of theoretically optimal cooperative transmission distributed beamforming architectures. As the size of a node group grows, the energy savings of ES converge to the optimum theoretically, and as verified by simulations.

II. System Model

Suppose a set of N sensor nodes is deployed within an area A, according to some spatial distribution Γ. We assume:

- the nodes measure the values of a continuous phenomenon (temperature, soil moisture, electromagnetic radiation, etc.), over A;
- time is discretized in slots; a set of active nodes report measurements to the sink during each timeslot;
- nodes transmit over multiple-access AWGN channel at the same average power per bit;
- the sink is within the transmission radius r of the nodes;
- nodes are coarse-grained synchronized on timeslot level;
- nodes within transmission radius, may exchange local information and estimate internode distances either based on control signal strengths, as in the well-established scheme in [23], or via GPS if available;
- nodes run a neighbor discovery procedure upon network deployment, and periodically henceforth, to maintain valid neighborhood knowledge.

These assumptions are typical in efficient sampling and distributed data compression schemes for WSN in various application (e.g., [5], [11], [12]).

Although measurements at different positions in A are frequently assumed to be *i.i.d.* in the research literature, we make the more realistic assumption that measurements are instead spatially-correlated and dependent on the distance between sensors (e.g., [5], [10]).

Formally, the statistical description of the measured phenomenon can be captured by a continuous source **S** giving rise to a space-time random field $s(l, x, y)$. At time l the random field is defined as $\{S_j[l] = s(l, x_j, y_j)\colon (x_j, y_j) \in A\}$ at spatial position (x_j, y_j), 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. S_j is then characterized by:

$$E\left[S_j\right] = 0, \quad Var\left[S_j\right], \quad \rho_{i,j} = E\left[S_i S_j\right]/\sigma_S^2,$$

where $\rho_{i,j}$ are the correlation coefficients of the JGRV.[2]

Phenomenon's values at different points in A are dependent on one another via correlation coefficients $\rho_{i,j}$, which represents some function of the distance between the points. Formally, this function is represented by a parameterized spatial covariance model, which reflects the nature of the specific phenomenon and determines the correlation coefficient. This statistical spatial-correlation function could be dictated by the physics of the phenomenon's nature (e.g. temperature, soil moisture, electromagnetic radiation, etc. [24]), or learnt by sensor nodes via sample training data at different locations in A ([25]). Thus, we assume that the function is known. For instance, the covariance model could be *Power Exponential*,[3] since different phenomena monitored by sensor networks are approximated this way (e.g. [5], [24]):

$$\rho_{i,j} = K_\theta\left(\|\mathbf{i} - \mathbf{j}\|\right) = e^{(-\|\mathbf{i}-\mathbf{j}\|/\theta_1)^{\theta_2}} \tag{1}$$

where $||\bullet||$ denotes Euclidean distance between sensors i and j at (x_i, y_i) and (x_j, y_j), respectively. Figure 2 illustrates a set of representative realizations of the phenomenon model and node deployments for different values of the correlation parameters θ_1 and θ_2. The parameter θ_2 determines how "grainy" is the observed phenomenon realization: lower values of θ_2 lead to more uneven and "grainier" realizations.

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 variation* W_j. The sensing variations can be due to the imprecision of the sensing equipment, or due to spatial changes in the observed values sensed by the sensors. W_j's are assumed to be *i.i.d.* Gaussian random variables, such that $E[W_j] = 0$ and $Var[S_j] = \sigma_W^2$. Consequently, node j measures $X_j = S_j + W_j$.

III. Encoded Sensing

As per above model, suppose the N nodes in the network are partitioned into groups of nodes with highly correlated measurements, so that in each group G the maximum difference between any two nodes' measurements is less than ε. Each node in the WSN quantizes the phenomenon range $[a, b]$ in equal intervals of length ε. Each interval is uniquely labeled,

[2]We assume memoryless source and (similarly to [5]) do not account for temporal correlations. ES is readily adjustable to temporal correlations; i.e., if some readings are more likely than others, the algorithms listed below 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.

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

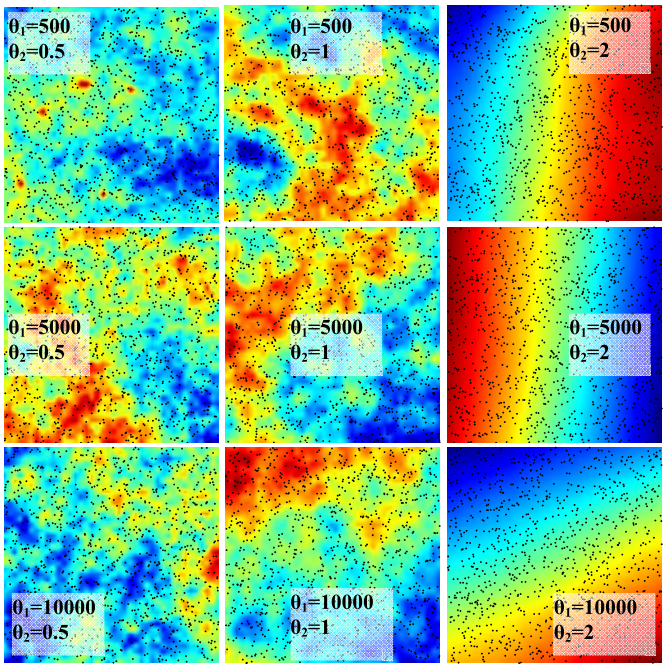

Fig. 2. Influence of parameters θ_1 and θ_2 in the correlation model of eq. (1) on the spatial distribution of phenomenon values. Notice that θ_2 affects significantly the behavior of the phenomenon. Nodes (black points) are placed uniformly at random. Network density is 2.5 nodes/m^2.

and the intervals are mapped to an index $x = 1, 2, \ldots, M$.[4] In this section, we describe encoded sensing's three stages – encoding, transmission, and decoding – within a single group of nodes $G \subset N$, for clarity. The algorithms presented in this section remain unchanged as we extend this setup in section VI to multiple groups.

Assume the nodes in group G are assigned unique IDs: $0, 1, \ldots, |G|-1$. Also, assume $|G|$ and the number of intervals M in $[a, b]$ are known by all the nodes[4]. At timeslot t, each node j in G measures value $X_j = S_j + W_j$. Since by construction of group G the maximum difference between nodes' measured values is ε and the length of each interval x is also ε, X_j falls in interval $x-1$, x, or $x+1$ for all nodes j in G. To demonstrate the encoding process, at present, suppose that for all j in G, the measured X_j's at timeslot t fall within x (e.g., all the nodes in G measure exactly the same value); we address the more general cases in section IV.

A. The Encoding Stage

At timeslot t, upon measuring a value falling in interval x, $1 \le x \le M$, during the *encoding stage*, each node in G locally encodes x into binary codeword c_x of length $|G|$ using **Algorithm 1**. **Algorithm 1** is a common decentralized encoding algorithm run at all nodes with input x and K^*, and output the set A_x containing K^* integers from 0 to at most $|G|-1$. Set A_x determines codeword c_x: we let $c_x(i)$, $0 \le i \le |G|-1$, denote the i-th most significant bit in c_x, and $c_x(i) = 1$ iff $i \in A_x$. K^* is a system parameter that minimizes both the number of transmitting nodes and transmitted bits per measurement, while ensuring reliable decoding is feasible. K^* is setup as follows. Note that $|A_x| = w(c_x)$ is the Hamming weight of codeword c_x. Suppose we impose $|A_x| = w(c_x) = K$ for all x. Notice that the number of possible distinct codewords of length $|G|$ that any algorithm can generate under this constraint is ${}^{|G|}C_K$.[5] All M possible intervals must be encoded into M unique codewords so that reliable decoding is feasible. Thus, we require ${}^{|G|}C_K \ge M$. Let

$$K^* = K \quad \text{s.t. } {}^{|G|}C_K \ge M \tag{2}$$

K^* is easily computed numerically at each node for any given $|G|$ and M. For each distinct input pair (x, K^*), **Algorithm 1** performs basic combination unranking and outputs a distinct set A_x, ensuring ${}^{|G|}C_{K^*} \ge M$. Note that if $|G| \ge M$ then $K^* = 1$.

Now suppose for any node with ID j in group G we have

$$D_j = \begin{cases} 1, & if\ j \in A_x \\ 0, & otherwise \end{cases} \tag{3}$$

and node j only transmits a binary symbol "1" iff $D_j = 1$. I.e., the number of transmitting nodes in timeslot t is $|A_x| = w(c_x) = K^*$, and each node transmits a single binary symbol. A_x contains the IDs of transmitting nodes in group G. In this case, by construction, K^* is the minimum value guaranteeing reliable decoding is feasible, while minimizing the number of transmitted binary symbols and nodes as $|G|$ grows. We dub this simple code *combinatorial encoded sensing* (*ES-C*).

As an example of **Algorithm 1**'s execution, suppose a group of $|G| = 6$ nodes measures values all falling within interval $x = 5$, where the total number of intervals $M = 16$. It is easy to determine that $K^* = 3$ from (2). (By inspection, if $K^* = 2$, ${}^6C_2 = 15 < M = 16$ and (2) is not satisfied.) All 6 nodes in G run **Algorithm 1** with input pair $x = 5$ and $K^* = 3$. The output of **Algorithm 1** is the set $A_5 = \{4, 2, 0\}$. From (3), we have $D_5 = 0$, $D_4 = 1$, $D_3 = 0$, $D_2 = 1$, $D_1 = 0$, $D_0 = 1$. Therefore, nodes with IDs 4, 2, and 0 transmit a binary symbol "1" each. The codeword is $c_5 = 010101$, with bits $c_5(4) = c_5(2) = c_5(0) = 1$.

Algorithm 1: Assign *Interval* 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: $A_x \leftarrow \{\emptyset\}$
2: $m \leftarrow K^*$
{At each iteration add a node ID to the assignment A_x of nodes to measurement x}
3: **while** $m \ge 1$ **do**
4: $n_m \leftarrow$ maximum integer such that $\binom{n_m}{m} \le x$
5: $x \leftarrow x - \binom{n_m}{m}$
6: $m \leftarrow m - 1$
7: $A_x \leftarrow A_x \cup n_m$

[4] In section VI, we discuss the practical configuration of the parameters $[a, b]$, ε and accordingly M, for popular WSN mote boards (e.g. TelosB).

[5] Binomial coefficients are denoted $\binom{n}{k}$ and ${}^nC_k$ throughout the text.

B. The Transmission Stage

In the above example, given **Algorithm 1**'s output, codeword c_5, and the corresponding transmission decision protocol in (3), the sink receives three binary symbols "1." The sink can only infer c_5 if it could determine the significance (i.e., position) of each "1" in c_5. Notice however that the positions of "1"s in c_5 above correspond exactly to the IDs of the transmitting nodes. This is true for any output codeword c_x of **Algorithm 1**. The sink can recover the c_x if nodes' transmissions indicate nodes' identities as well.

Node identity information however is often implicit in many standard physical layer protocols employing variants of spread spectrum wave signatures embedded in each node's signal. That is the only assumption ES makes regarding the physical layer. 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 [26]. Other off-the-shelf physical layer solutions for WSN utilizing orthogonal spreading sequences satisfying ES assumptions are described in [27]–[29]. Using a DSSS physical layer approach, in ES, node j transmits binary symbol "1" by transmitting j's unique spreading sequence. The waveform received at the sink contains j's unique signature "1_j." The received waveform is correlated at the sink, for example, via a bank of matched filters, with the set of available spreading sequences' waveform signatures for all nodes in the WSN. The sink obtains the identity of the transmitting nodes during this sequence acquisition stage.

In our example above, nodes 4, 2, and 0's binary symbols are multiplied by their respective unique spreading sequences upon transmission during timeslot t. The received waveform at the sink is the integrated waveform $y(t)$ of signals "1_4," "1_2" and "1_0." After correlating $y(t)$, during the spread sequence acquisition, the sink obtains the IDs 4, 2, and 0. The sink recovers the word $0_5 1_4 0_3 1_2 0_1 1_0$, which corresponds to $c_5 = 010101$.

Section V describes and analyzes in detail the ES physical layer properties and its fundamentally optimal communication energy consumption as the number of nodes in group G increases.

C. The Decoding Stage

After receiving and correlating nodes' transmissions according to the DSSS physical layer properties, the sink identifies the nodes j in A_x each of which has transmitted its waveform "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 based on A_x. The following *Theorem* 1 ensures that the IDs of the nodes in A_x available at the sink are sufficient to recover interval 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* $m \leq n_m$

$$x = \binom{n_m}{m} + \binom{n_{m-1}}{m-1} + \ldots + \binom{n_i}{i} = \sum_{i=1}^{m} \binom{n_i}{i} \quad (4)$$

Proof: See [30]. □

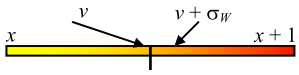

Fig. 3. The value v of the phenomenon falls close to one of the endpoints of interval x. Due to sensing instrumentation imprecision σ_W, the value is incorrectly measured to fall in interval $x + 1$.

Notice that Theorem 1 guarantees existence as well as uniqueness. Also notice that the set of IDs in A_x output by **Algorithm1** exactly matches the sequence $(n_{K*}, n_{K*-1}, \ldots, n_1)$ required by Theorem 1, where $m = K^*$. Also by **Algorithm 1**'s construction $n_{K*} \geq K^*$. Therefore

$$x = \sum_{i=1}^{m} \binom{n_i}{i} \quad (5)$$

Thus, this is the only value x to which codeword c_x can be decoded. The value x is obtained simply by *combination ranking* of A_x.

To conclude our example, after inferring the codeword $c_5 = 010101$ during the transmission stage and identifying the transmission nodes' set as $A_x = \{4, 2, 0\}$, the sink maps A_x to the sequence $(n_1, n_2, n_3) = (4, 2, 0)$. Knowing $m = K^* = 3$, the sink obtains

$$x = \sum_{i=1}^{3} \binom{n_i}{i} = \binom{4}{3} + \binom{2}{2} + \binom{0}{1} = 5,$$

which is the correct interval containing the measurements of nodes in group G, at timeslot t.

IV. Minimum Distance Combinatorial Encoding

In section III, we assumed that the measurement values X_j's during timestep t fall into the interval x for all j in G. However, by construction of group G, the maximum difference between nodes' measured values is ε and the length of each interval x is also ε, hence X_j can fall into interval $x - 1$, x, or $x + 1$ for all nodes j in G, during timeslot t.

Figure 3 illustrates an example scenario where nodes in group G observe the same value v, but report measurements falling in adjacent but different intervals due to instrumentation noise W_j with power σ_W. Any node j in G may observe value S_j of the phenomenon that is arbitrarily close to either of interval's x endpoints. Even small instrumentation noise may lead to some number k of nodes in G to measuring values in an adjacent interval to x. Notice that flipping two bits in a codeword generated by the simple *ES-C* encoding above, may render the transmission sent to the sink invalid or, worse, wrong causing large distortion in the sink's estimate. For instance, flipping two bits in $c_5 = 010101$, from the example above, may result in $c_0 = 000111$.

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 adjacent interval x, the challenge is to find an encoding scheme with the property that the correct codeword is received at the sink almost surely, given $|x - x'| = 1$.

To address the above challenge, we propose the *minimum distance combinatorial encoding* (MDCE) presented in the next section. MDCE avoids the above shortcoming

of *ES-C*, while preserving *ES-C*'s code properties w.r.t. to the code range, codeword length, and Hamming weight.

Suppose x and x' are any two intervals 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 $K^* = |A_x|$ transmitting nodes out of the $|G|$ nodes in a group, we can obtain a new valid codeword from any given valid codeword by 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 increase by 1; and if only 1 bit flips to 0, $|A_x|$ would decrease by one.) More specifically, let $\mathrm{HD}(c_x, c_{x'})$ be the Hamming distance between the codewords c_x and $c_{x'}$ assigned to measurements in intervals x and x' respectively. We require that

$$|x - x'| = 1 \Rightarrow \mathrm{HD}(c_x, c_{x'}) = 2 \tag{6}$$

Notice that similarly to *ES-C*, we still have $|c_x| = |c_{x'}| = |G|$, $w(c_x) = w(c_{x'}) = |A_x|$, and nodes transmit according to eq. (3).

We first show the benefit of such *minimum distance combinatorial encoding* (MDCE) with the latter property; and then provide the construction generating MDCE codes.

A. MDCE Probability of Erroneous Codeword

Suppose that the event in Fig. 3 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 $|x - x'| = 1$.

Theorem 2: Let F_E be the probability that a MDCE codeword transmitted by group G is invalid or wrong, then $F_E \to 0$ as $|G| \to \infty$.

Proof: See Appendix A. □

Fig. 4 (top) shows the probability that exactly k nodes are in error by varying numbers of k and $|G|$; Fig. 4 (bottom) shows the values of F_E for varying numbers of k and $|G|$, when at most k nodes are in error. The MDCE code guarantees a valid and correct codeword is transmitted by group G w.h.p., even though some nodes in G may measure different values, for instance, due to instrumentation noise.

B. 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 to encode 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 $m = w(c) = |A_x| = K^*$, so that (2) is satisfied; if true, we add c to our MDCE codebook, setting $c_x = c$ and then incrementing x by one.

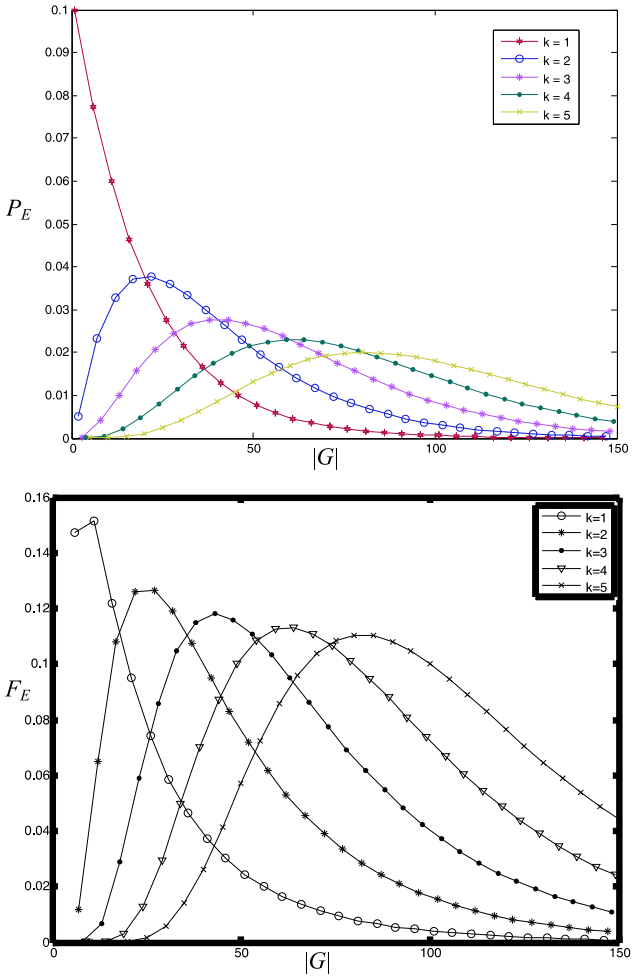

Fig. 4. MDCE probability of encoding invalid or erroneous codeword given different number k of sensors with erroneous measurements, for $p_e = 0.05$. Up: exactly k sensor nodes in error; down: *at most* k sensor nodes in error.

Theorem 3: The resulting MDCE code satisfies

$$|x - x'| = 1 \Rightarrow \mathrm{HD}(c_x, c_{x'}) = 2$$

and hence is a minimum distance combinatorial encoding.

Proof: See Appendix B. □

Table I lists the set of the MDCE and ES-C codewords for the case $|G| = 6$, $M = 20$, where $|A_x| = K^* = 3$ respectively. Notice that the MDCE codewords have weight equal to that of codewords of *ES-C* generated by **Algorithm 1** for a given M and $|G|$. In fact, *the set* of MDCE codewords and *the set* of ES-C codewords are identical, for any given M, $|G|$ and $|A_x|$. However, the *mapping* between codewords and messages is rather different. The Hamming distance between any two consecutive codewords is exactly equal to two in the case of MDCE. We denote $\mathrm{MDCE}_{m,n}$ the set of binary codewords c_x with length n and Hamming weight m; $\mathrm{MDCE}^I_{m,n}$ denotes the corresponding collection of sets $|A_x|$.

ES encoding, transmission, and decoding follow the same stages described in the previous section. As in the example above, a group of $|G| = 6$ nodes measures values all falling in interval $x = 5$, where the total number of intervals $M = 16$ and $K^* = 3$. All the nodes in G utilize the MDCE code (e.g. Table I) and look up the codeword corresponding

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

to $x = 5$: $c_5 = 011010$; correspondingly, $A_5 = \{4, 3, 1\}$. Nodes in G with IDs 4, 3 and 1 are assigned to $x = 5$. From (3), we have $D_4 = 1$, $D_3 = 1$, $D_1 = 1$, $D_5 = 0$, $D_2 = 0$, $D_0 = 0$, thus nodes with IDs 4, 3, and 1 transmit a binary symbol "1" each. The sink receives the sum of waveforms 1_j, for all j in A_x and determines the IDs of nodes in $|A_5|$. Looking up the value of x encoded by $A_5 = \{4, 3, 1\}$, the sink recovers $x = 5$. Now suppose that due to an instrumentation error node 4 measures value falling in interval $x = 6$. From Table 1, the codeword corresponding to $x = 6$, indicates $D_4 = 1$. Node 4 still transmits with nodes 3 and 1. Thus, the sink infers c_5 and the correct value of $x = 5$.

We dub the above scheme *ES-MDCE*.

V. TRANSMISSION ENERGY

We evaluate the transmission gains and the respective energy savings of the *ES-MDCE* collaborative encoding of measurements and compare those to non-cooperative (NC) transmission schemes. Suppose that all the nodes in group G measure values falling within the interval x in the timeslot t and the message encoding x is transmitted. Suppose that there are M such intervals. A NC scheme relying on spatial sampling compression would require only one representative node in the group G to transmit the message. That is sufficient, provided ε inaccuracy of measurements is within QoS guarantees, since the rest of the nodes sample values in x as well.

Since each node sending a bit per ES needs to be identified by the sink, for the analysis of ES transmission energy consumption we consider a regular DSSS physical layer, for example, specified by IEEE 802.15.4. Typically, a DSSS communication system is described both by post-acquisition-based capacity and acquisition-based capacity. The post-acquisition-based capacity characterizes the maximum number of nodes in the system so that, given processing gain F (number of PN sequence chips per bit), the probability of a bit error is less than a threshold μ'. The acquisition-based capacity characterizes the maximum number of nodes in the system, so that the probability p_a of acquiring wrong PN sequence at the sink is less than a threshold μ.

ES relies only on the correct output of the DSSS matched-filters at the sink after the acquisition stage. Hence, we are interested in the acquisition-based capacity of the system. Assume that the waveform 1_j, for all j in A_x, encoding interval x using MDCE, arrives at the sink with independent random time-delays (within any given reporting time-slot) and with independent random carrier phases.[6] The probability, p_a, of acquiring a wrong PN sequence in our setting is given by:

$$p_a \cong \frac{F-1}{2} e^{-\frac{3F}{4(|G|-1)}} \tag{7}$$

provided $F > |G|$, as in typical DSSS systems with long PN sequences (e.g. see [31]). Notice that the system is interference limited. In contrast, per NC schemes, a single representative node transmits, sending at least $\log(M)$ bits, encoding one of M intervals. The bit error rate of the NC scheme utilizing, for instance, binary amplitude shift keying OFDMA over AWGN channel[7] is given by:

$$p_b \cong \frac{1}{2} erfc\left(\sqrt{SNR}\right) \tag{8}$$

Throughout this paper, when comparing the energy performance of the ES scheme to a NC scheme, we are conservative and require not only that w.h.p. $p_a < p_b$, but also that $P_a = 1 - (1 - p_a)^m < p_b$, where $m = |A_x|$. That is, we ensure that the probability P_a of not acquiring all m signals correctly is less than the probability of error of traditionally demodulating a single bit. This ensures that the reduced energy consumption of the ES schemes is not an artifact of increased error rate vis-à-vis traditional modulation.

From (7), the acquisition capacity of the system is:

$$C_{acq} \cong \frac{3F}{4\ln(F/2\mu)} \tag{9}$$

where μ is the threshold of acceptable acquisition error rate p_a.

For example, Fig. 5 shows the bit error rate p_b of a NC system, the probability p_a of acquiring a wrong PN sequence, and the probability P_a that at least one of the received $m = |A_x|$ PN waveforms was acquired incorrectly by the ES scheme, for different group sizes $|G|$. For a standard DSSS gain $F = 512$ and system acquisition capacity $C_{acq} = 35$ in all cases $p_a < P_a < p_b < \mu$. Here SNR ≈ 5.5 dB, $p_b < \mu \approx 0.005$.

Suppose that both the NC and ES schemes operate with a budget E_b per bit, so that $p_a < P_a < p_b < \mu$, and in the case of ES the system operates under acquisition capacity C_{acq}, as discussed above. Given transmission energy budget E, the ES and the NC schemes can send respectively d_{ES} and d_{NC} messages in total: $d_{\mathrm{ES}} = E/[w(c_x)E_b]$ and

[6]Notice that these assumptions are conservative. In many systems, more timing information is available a priori, for example by virtue of a side feedback channel devoid of interference, and the probability of acquiring wrong PN sequence is lower implying larger system capacity.

[7]This choice is general, since the bit error rate expression for p_b in (8) remains the same for BPSK and *single user* DSSS systems, for instance.

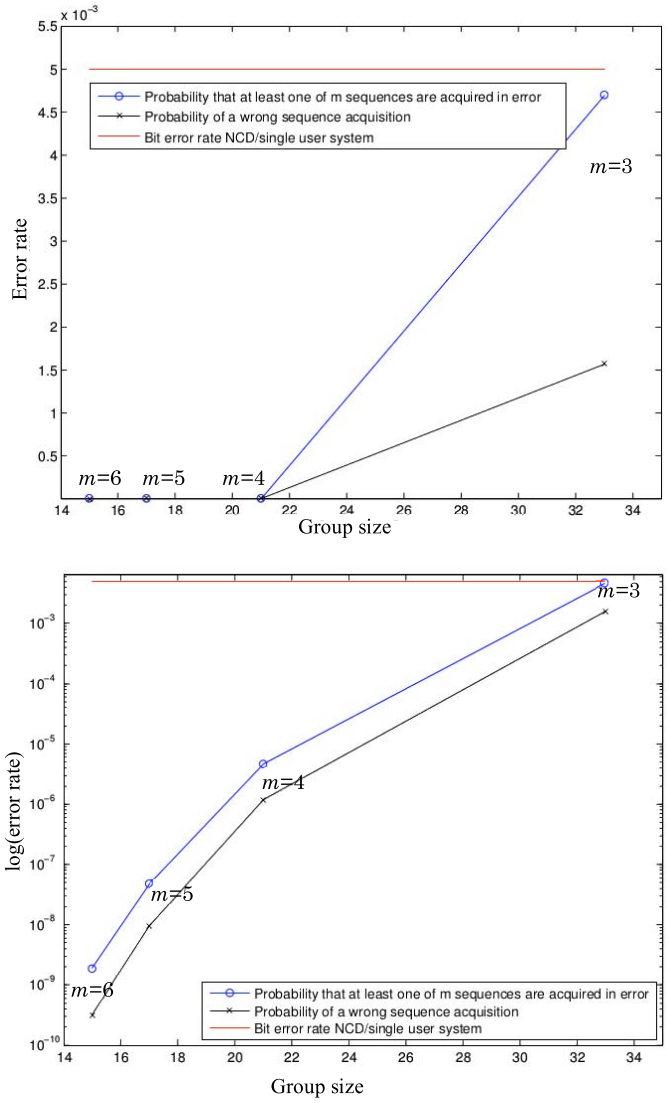

Fig. 5. Probability of wrong PN sequence acquisition at the sink (both in terms of P_a and p_a) for varying number of nodes in an ES group; and bit error rate p_b for a NC scheme. Notice that in our simulation, for all group sizes $|G|$, we require and satisfy $p_a < P_a < p_b$.

$d_{\text{NC}} = E/[\log(M)E_b]$. The ratio $d_{\text{ES}}/d_{\text{NC}}$ determines ES's energy efficiency w.r.t. a NC scheme.

A. Encoded Sensing Worst-Case Energy Efficiency Compared to Optimal Non-Cooperative Schemes

The Hamming weight $w(c_x)$ of ES codewords equals $|A_x|$. The number of messages, d_{ES}, that the ES scheme can send decreases as $|A_x|$ increases. For a worst-case comparison we seek codeword Hamming weight resulting in maximum $|A_x|$ and minimum d_{ES}. Given fixed M and $|G|$, we require that (2) is satisfied, so that ES-MDCE is a valid code.

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

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

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

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

We also notice that since $\log(0.5|G|)/4 \leq 1$ for practical values of $|G|$ (up to $|G| = 32$ nodes),[8] then:

$$|G| \geq 2\lceil log(M)/2 + 1.5\rceil. \tag{10}$$

Satisfying (10) ensures that there are enough codewords for each message in M. Here, $K^* = |G|/2$. This is the maximum possible number of nodes in set A_x, largest codeword weight, and the worst case of ES operation for fixed $|G|$ and M.

Suppose that $K^* = |\text{G}|/2 = \lceil\log(M)/2 + 1.5\rceil = w(c_x) = |A_x|$ bits are transmitted per message, as in the worst case of the ES operation from (10). Then

$$d_{\text{ES}} = 2d_{\text{NC}}/(1 + 3/log(M)). \tag{11}$$

In its worst regime of operation, ES is close to 2 times as energy efficient as an optimal NC scheme. Notice, that an optimal NC scheme would require the transmission of at least $\log(M)$ bits. Even if we suppose that a NC scheme utilizes different modulation, for instance orthogonal non-antipodal, 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 be at least double.

Allowing larger group sizes $|G|$ substantially reduces the value of K^* utilized in ES as shown in Fig. 8. As we show in section VII, in practical scenarios $K^* \ll |G|/2$ is sufficient to generate enough codewords, leading to further energy savings.

B. Encoded Sensing Best-Case Energy Efficiency Is Optimal

T-ary orthogonal codes (e.g. PPM) achieve reliable communication over AWGN channels at the minimum possible energy per bit E_b, as T increases ([32]). There, the total power is spread over a large time interval to achieve the Shannon limit of $E_b/N_0 = -1.6$dB.

Given $|G| \geq M$, the ES scheme encodes interval x via a single binary symbol since $w(c_x) = K^* = 1$. The symbol is sent by a single source node j, whose bit is spread over j's DSSS spreading sequence. Typically, DSSS spreading sequences have low cross-correlation; ideally they are orthogonal to each other. In the latter case, node j's sequence is orthogonal to the spreading sequences of the nodes in $G/\{j\}$ and ES operates at the optimum energy per bit level, achieving the Shannon limit of $E_b/N_0 = -1.6$dB ([32]).

The ES scheme realizes a non-linear*trade-off between the number of nodes in the network and the network's energy efficiency*. As the number of nodes within a group increases ES provides optimal transmission efficiency by reducing $w(c_x)$ and inherently eliminating unnecessary nodes' transmissions of correlated measurements via MDCE codes.

[8] If $|G| = 32$, ${}^{|G|}C_{|G|/2} \approx 6{*}10^8$ distinct messages can be mapped, which is sufficient for most practical applications of WSNs.

VI. Multiple Groups

Thus far we assumed that the network is partitioned into groups of nodes with highly correlated measurements, so that in each group G the maximum difference between any two nodes' measurements is less than ε. To this end, there are two conditions we need to meet simultaneously:

- First, each group G occupies area O in A, where the maximum difference between any two realizations S_j and S_i of the phenomenon at points (x_j, y_j) and (x_i, y_i) in O sampled by any two sensors j and i is less than ε 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 $|x - x'| > 1$.

A. Source Quantization

To meet the *second condition*, we quantize the range $[a, b]$, so that the probability is low that the measurement imprecision W_j offsets the measurement of any $j \in G$ outside interval x', where $|x - x'| \leq 1$. The instrumentation imprecision W_j of sensor j is modeled statistically 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., we let $\varepsilon = 2\beta\sigma_W$, where $\beta \geq 2$. In applications, WSN nodes' (e.g. TelosB, MICA, etc.) sensors are characterized by pre-specified range and accuracy. For instance, the Sensirion temperature sensor on a TelosB TPR2420 node has range $[a, b] = [-40\ °C, 124\ °C]$ and accuracy $+/-0.5\ °C$ (at 25 °C) ([36]). By definition, the accuracy sensor parameter is the maximum difference between a true phenomenon value and sensor's output. Setting $\varepsilon = 1°C$, results in $M = 164$ intervals (i.e. $1 \leq x \leq 164$) and guarantees $|x - x'| \leq 1$ in practice.

B. 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 [5]. Given the nodes' spatial statistical distribution Γ and the spatial correlation model, e.g., eq. (1), as input, the algorithm selects k *representative* nodes out of the N nodes in the WSN. In [5], the selected k nodes are chosen so that the spatial correlation between their measurements is reduced, while k is minimized, and the distortion in the sink's estimate of the source $\mathbf{S}$ is below a given QoS threshold. As a result correlated measurements are eliminated, without sacrificing needed measurements accuracy, thus achieving optimal spatial *sampling compression,* w.r.t. QoS. Here, we exploit a related property of the algorithm in [5] to satisfy the ES requirement that w.h.p. a group G of nodes all sense values less than ε apart. Namely, 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[9] used in [5]. Consider a disk, O_j, with radius r_{corr}, centered at representative node j.

Definition 1(Representative Group): A group of nodes, G_j, situated in disk O_j is called a *Representative Group.*

Nodes within each representative group G_i utilize ES-MCDE to transmit their measurements to the sink. ES operates independently in each G_i, as described in sections III and IV: each short interval x is assigned to a distinct subset of nodes A_x in G_i according to MDCE. Upon measuring a value in interval x during time-slot t, 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 the MDCE map.

Notice that ES's cooperative transmission of nodes in each group G_i differs significantly from the scheme suggested in [5], where each of the representative k nodes transmits non-cooperatively a single measurement to the sink every time-slot. We dub the latter scheme presented in [5] *non-cooperative transmission with decorrelation* (NCD).

In Appendix C, we show that, as the sensing imprecision due to W_j decreases, for the same minimum number of representative groups/nodes k, the distortion $D_{ES}(k)$ in sink's estimates of $\mathbf{S}$ converges to the fundamentally optimal distortion, $D_{NCD}(k)$, achieved by NCD, satisfying the QoS.

VII. Performance Evaluation

We compare the energy efficiency and measurements' estimate accuracy of the ES scheme to other state-of-the-art communication schemes for WSN, employing

- data compression: *distributed source coding* (DSC) ([11], [12]);
- sampling and communication compressions: *non-cooperative transmission with decorrelation* (NCD) ([5]);
- cooperative transmission: *cooperative distributed transmit beamforming* (BF) ([20]).

A. Simulation System

The simulation environment and schemes' implementations are programmed in JAVA and utilize the BLOG Inference Engine ([33]), available online. In each simulation run, N nodes are placed within a 100[m] $\times$ 100[m] square area. S_j's are modeled as spatially correlated JGRVs with covariance model K_θ (eq. (1)). The BLOG Inference Engine is used to generate the set of JGRV S_j's.

- The phenomenon's values are assumed to be in the range $[-4\sigma_S, 4\sigma_S]$; where $\sigma_S = 625$. 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. Per *ES*, the range of phenomenon values is quantized into a sequence of intervals, where $M = 5000$. Each interval is of length $\varepsilon = 1$, with $\sigma_W = 0.25$.

[9]If the statistical properties of the phenomenon (i.e., parameters θ_1 and θ_2) do not change over time, the representative groups and correlation radius in the network needs to be determined only once. Otherwise, to satisfy QoS constraints over time, the representative groups and correlation radius have to be computed periodically as in [5], depending on the frequency of the phenomenon's changes.

• 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. An exception here is the distributed transmit beam forming, which is by itself a physical layer scheme. In all our simulation runs we require that $|G| < C_{acq} = 35$ (see Fig. 8). Thus the acquisition capacity of the system is approximately 35 users/nodes as discussed in Section V.

B. Simulated Schemes

• Per *NCD* [5], *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 optimal spatial sampling efficiency. Note that 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 *DSC* scheme is a single hop variation of **Algorithm 2** in [11] and **Algorithm 1** in [12]. 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 the distance r_i of i. Next, in the order of the nodes' sequence, transmission data rates are allocated as follows: the 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. Notice that the resulting DSC scheme utilizes lossless data compression over each sequence C_i of nodes.

• The *BF* scheme evaluated here is based on the cooperative distributed transmit beamforming scheme suggested in [20]. An assumption in [20] is that upon each transmission, cooperating nodes are synchronized in frequency with each other. Furthermore, the nodes need to be synchronized in phase, so that the gains of BF are realized due to constructive interference at the sink. To satisfy the latter assumption, a simple randomized algorithm is offered in [20] to achieve phase coherency of the transmitted bits at the sink. To evaluate the performance of BF, the phase synchronization feedback algorithm is successfully run at each BF node in G_j to achieve 70% phase coherency at the receiver (which is equivalent to achieving about 80% of beamforming gains) in accordance with practical BF performance. Similarly to the above schemes, the beamforming groups G_j are formed within correlation radius r_{corr} around each of k representative nodes.

C. 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 timeslots. *Estimate inaccuracy* is given by the average inaccuracy of the

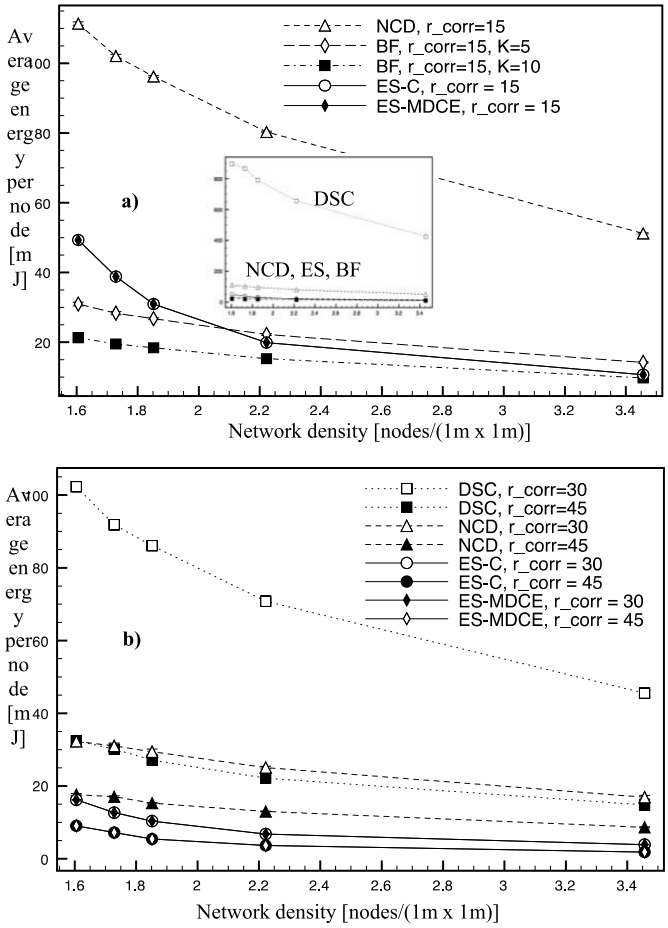

Fig. 6. Energy consumption for 150 rounds and different r_{corr} a) $r_{corr} = 15$m; inset: DSC performancs; b) $r_{corr} = 30$; 45m. Energy consumption decreases as the correlation radius increases (number of representative nodes decreases). The performance of ES-MDCE performs is identical to the basic ES-C encoding scheme.

phenomenon point source estimate at the sink over τ timeslots:

$$\overline{V} = \frac{1}{\tau} \sum_{i=1}^{\tau} V_i$$

where, at timeslot 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 in Fig. 6 and is 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. As more nodes are added to the network (i.e., increasing the node density), E decreases for all schemes. In the case of DSC and NCD, the decrease in E is due to data compression and sampling compression, respectively; in the case of BF, the decrease is due to the increase in cooperating nodes per group as the network density increases. The ES-MDCE/ES-C schemes benefit substantially more from the network density increase, compared to DSC and NCD. Larger network density allows for smaller values of K^*

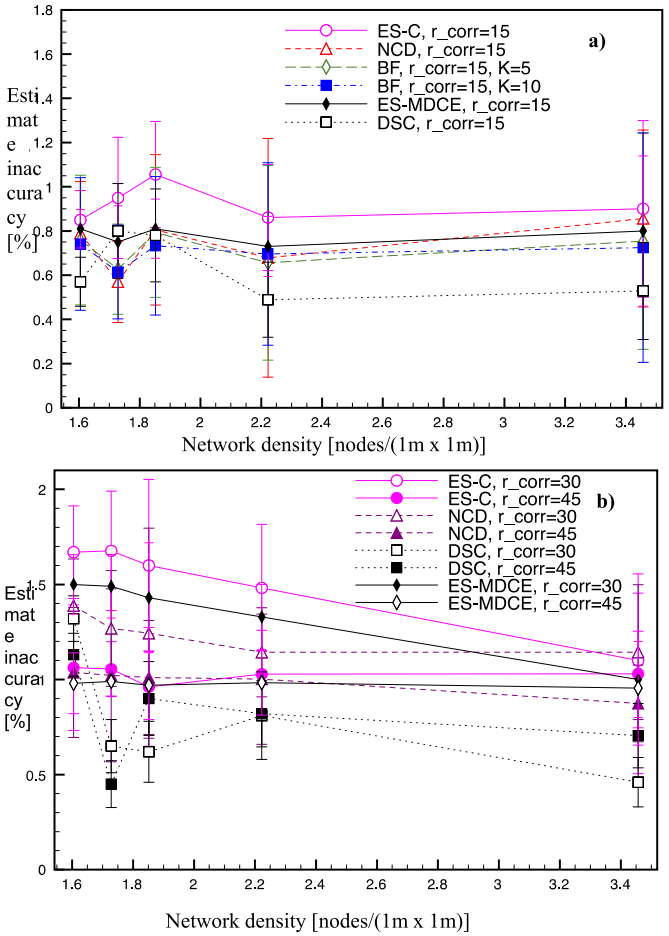

Fig. 7. Estimate inaccuracy; a) r_{corr} = 15m and; b) r_{corr} = 30; 45m. ES-MDCE achieves lower inaccuracy and lower variance than the basic ES-C encoding scheme. However, all four scheme estimate the source relatively accurately. The measurements at the representative sites found via vector quantization model well the phenomenon.

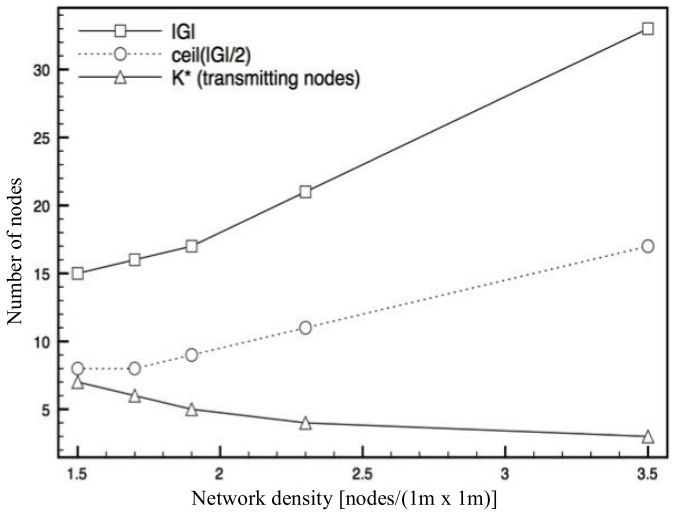

Fig. 8. Average number of nodes in a group G_i and number of transmitting nodes K^* for different network densities. Even at low densities, K^* is less than $|G|/2$, the worst-case of ES operation, and drops further with density. K^* is the same for both ES-MDCE and ES-C.

leading to a smaller number of actively transmitting nodes in $|A_x|$ required to convey the 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. As density increases, $|G_i|$ grows from 15 to 33 nodes and $|A_x|$ decreases from 7 to 3 nodes (Fig. 6).

BF edges over the performance of 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 (see [20]). Cooperating more than 10 nodes would pose a challenge to practical distributed transmit beamforming systems and in that regime ES matches BF performance at higher network density.

DSC achieves lower energy efficiency compared to the rest of the schemes, since all the nodes within a given neighborhood with radius r_{corr} transmit with positive rates, allocated to account for nodes observations' correlations. The total rate equals at least the joint entropy of the observed measurements and is higher than the rest of the schemes' rate.

Fig. 7 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 converges numerically. However, ES-MCDE outperforms slightly ES-C, achieving lower estimate inaccuracy. DSC obtains lower inaccuracy, since all nodes in each r_{corr} neighborhood around a representative node send information regarding the sensed value at their position.

VIII. Discussion

Often in WSN, groups of source nodes share common or correlated messages. Furthermore, in many such wireless networks the sink has the inherent ability to recognize the identities of transmitting nodes. The Encoded Sensing scheme offers a novel mechanism for exploiting this a priori knowledge in the system. As shown in this paper, collaborative encoding of shared measurements by a group of nodes possesses the potential of reducing drastically the amount of energy required to convey reliably these measurements. In the context of WSN, ES allows correlation patterns of information observed at sensor nodes to be accounted for both, the data modulation and the transmissions at the physical network layer, via carefully designed codewords.

More specifically, the number of nodes in a group G tends to increase both with growing network density and the larger level of spatial correlation across nodes. Hence, increasing the number of nodes in the network and/or the correlation of their measurements leads to smaller Hamming weight of any ES codeword and respectively to fewer transmitting nodes. Each transmitting node j sends a single binary symbol message, and the overall data that the ES transmits to convey the measurements and the energy required per transmission are guaranteed to be optimal as the size of a group increases. Hence ES utilizes data compression, sampling compression, and communication compression, coupled with collaborative transmission gains. To the best of our knowledge ES is the

first communication WSN scheme achieving this. We confirm by simulation and through analytical results that indeed ES improves energy efficiency compared to state-of-the-art communication schemes by at least a factor of two.

Per ES, the multiple access transmissions' structure *itself* represents a collaborative codeword that the sink can map back to the desired message, while the message itself is not transmitted. Other than the energy efficiency of the information transfer demonstrated in the paper, this novel paradigm may have applications in providing low-overhead transmissions security in WSN.

Codes different than the ones presented in this paper can be used in ES. An intriguing choice of a code would be one that provides graceful degradation of message fidelity. That is, given that 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. We leave this problem open for future research.

WSN control messages sent from a group of nodes to the sink can be handled efficiently as well by the ES schemes. Each type of control message that needs to be conveyed from a representative group to the sink can be assigned to a distinct subset of nodes. Since the number of control message types is typically low, the number of such required "control codewords" would be small. Control messages could be *keepalive* messages (i.e., test whether the link between the sink and a given representative group is available), power control messages (increase/decrease the transmit power of nodes within a representative group), etc. Various protocols could be designed to optimize the performance of control codewords (e.g., control codewords weight based on the number of control message types and priorities). That aspect is beyond the scope of this work.

Encoded Sensing Limitations

As presented in this paper, the ES schemes' application scenarios are subject to some limitations. We have only described a single-hop ES operation (i.e., the sink receiving sensors' measurements is within the transmission range of the nodes in the system). Due to space limitations we did not consider the use of ES in a multihop network setting in this work. However, ES can be extended for use as a cooperative transmission layer of various wireless networks cooperative routing schemes, such as the one in [34]. 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. Notice that in this setting, ES no longer relies on the amount of correlation across the measurements of sensor nodes. Coordinating such clusters of forwarding nodes and ensuring all nodes in a cluster receive the same message w.h.p. is feasible and open for further investigation.

Currently, ES requires that the sensor network possesses certain density, so that clusters of sufficiently large number of nodes (as described in section V.A) can have highly correlated measurements. This requirement can be satisfied more easily in the case where the network designer can control network topology to initially deploy sensor nodes in representative groups. However, already existing sensor networks with random nodes distribution may require large density of nodes, especially for phenomenon statistics depicted in the left most column of Fig 2. The "grainier" the phenomenon is, the smaller the correlation radius in the network. The density of nodes in each representative group and the number of representative groups should be larger. The required densities may pose practical limitations on the deployment of ES, where the statistics of the phenomenon lead to "grainier" realizations than the ones depicted in the left-most column of Fig. 2 (i.e., $\theta_2 < 0.5$). ES is most suitable for network deployment scenarios where the phenomenon spatial statistics are "smooth" ($\theta_2 > 0.5$); or, where the cost of each sensor node is low vis-à-vis feasibility of replacing sensor nodes with exhausted batteries, so that larger density of nodes is practical.

More generally, ES as presented in this work assumes statistical correlation model of the sensed phenomenon. If not available, the correlation model can be estimated as part of exploratory deployments (e.g. [37]). Given a correlation model, encoded sensing would be applicable if nodes are randomly deployed or change positions, as long as the correlation groups are updated with changes in topology. In cases where the correlation model cannot be approximated, non-parametric field estimation methods such as the ones proposed in [25] may be used instead of encoded sensing.

APPENDIX

A. Proof of Theorem 2

*Theorem 2***:** Let F_E be the probability that a MDCE codeword transmitted by group G is invalid or wrong, then $F_E \to 0$ as $|G| \to \infty$.

Proof: The probability P_k that k nodes are in error is:

$$P_k = \binom{|G|}{k} p_e^k (1-p_e)^{|G|-k}. \tag{A.1}$$

Moreover, the probability that *at most* k nodes are in error is:

$$F_k = \sum_{i=0}^{k} \binom{|G|}{k} p_e^k (1-p_e)^{|G|-k}. \tag{A.2}$$

Notice that $k \leq |G| p_e$, since $k < |G|/2$ and $p_e < 1$. Therefore using Chernoff's bound, we obtain:

$$F_k \leq exp\left\{-\frac{(|G|\, p_e - k)^2}{2|G| p_e}\right\}. \tag{A.3}$$

Suppose node j is one of the k nodes that erroneously determines the codeword to be transmitted is $c_{x'}$. Node j checks the bit at position j in $c_{x'}$ to determine its transmission decision according to eq. (3). 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 eq. (6). 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:

$$p_j = \frac{2}{|G|} \tag{A.4}$$

Given that we have *at most* k nodes in error, the probability P_R that any of the k nodes reports erroneously (causing at

least a single bit error in the codeword received at the sink) can be bounded as follows:

$$P_R = \bigcup_{j=1}^{k} p_j \leq \sum_{j=1}^{k} p_j = \sum_{j=1}^{k} \frac{2}{|G|} = \frac{2k}{|G|}, \quad \text{(A.5)}$$

where the inequality follows from the union bound. Since the probability that *at most* k nodes are in error is F_k from (A.3), we have that in this case the overall probability F_E of error in the message is given by:

$$F_E = F_K P_R \leq \frac{2k}{|G|} exp\left\{-\frac{(|G|\, p_e - k)^2}{2|G| p_e}\right\} \quad \text{(A.6)}$$

Note that if exactly k nodes are in error then

$$P_E = F_k P_R \leq \frac{2k}{|G|} \binom{|G|}{k} p_e^k (1-p_e)^{|G|-k}, \quad \text{(A.7)}$$

since the probability that exactly k nodes are in error is P_k from (A.1).

Assuming p_e is small,[10] k is a constant, and $|G|$ is large

$$P_E \leq \frac{2k}{|G|} \frac{e^{-|G|p_e} (|G|p_e)^k}{k!}. \quad \text{(A.8)}$$

Note that $P_E \to 0$ as $|G|$ increases, and the probability that a MDCE codeword transmitted by group G is invalid or wrong can be made arbitrarily low. Similarly, $F_E \to 0$ in (A.6) as $|G|$ increases. □

B. Proof of Theorem 3

Suppose that we have a set of M messages to encode 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 eq. (2) is satisfied; if true, we add c to our MDCE codebook, set $c_x = c$ and increment x by one. We stop when $x = M$.

Theorem 3: The resulting MDCE code satisfies

$$|x - x'| = 1 \Rightarrow \mathrm{HD}(c_x, c_{x'}) = 2$$

and hence is a minimum distance combinatorial encoding.

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

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

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,

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

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 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,

$$\begin{aligned} MDCE_{1,2} &= \{0MDCE_{1,1}, 1reverse(MDCE_{0,1})\} \\ &= \{0^{2-1}1^1, 10^{2-1}1^0\} = \{01, 10\}. \end{aligned}$$

[10]For instance, as the instrumentation noise decreases, likely scenario as WSN hardware solutions improve, the probability of inaccurate measurements decreases as well.

Invariantly, to obtain

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

we need to flip only 2 bits in

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

To see that, we observe that

$$0MDCE_{m,n-1} = 010^{n-m-1}1^{m-1}$$

and

$$1MDCE_{m-1,n-1} = 110^{n-m-1}01^{m-2}$$

by induction $\forall m \geq 2$. Trivially, for $m = 1$ we again only flip 2 bits to transition from

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

to

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

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. □

C. Distortion of Measurement per ES and NCD

The estimate of the source $\mathbf{S}$ at the sink produced utilizing either ES or NCD is distorted, since 1) only information regarding the measurements of the k representative out of the N nodes in the network is taken into account, and 2) there is channel noise and sensing imprecision noise present. The CC-MAC algorithm in [5] 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 optimized 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 the standard Minimum Square Error (MSE) metric:

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

where $\mathbf{S}$ is the value of the point source and $\mathbf{S}$' 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 ([35]) subject to power constraint P per node, per measurement. The received signal at the sink is:

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

The optimal decoder at the sink is given by the standard MMSE estimator ([48]):

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

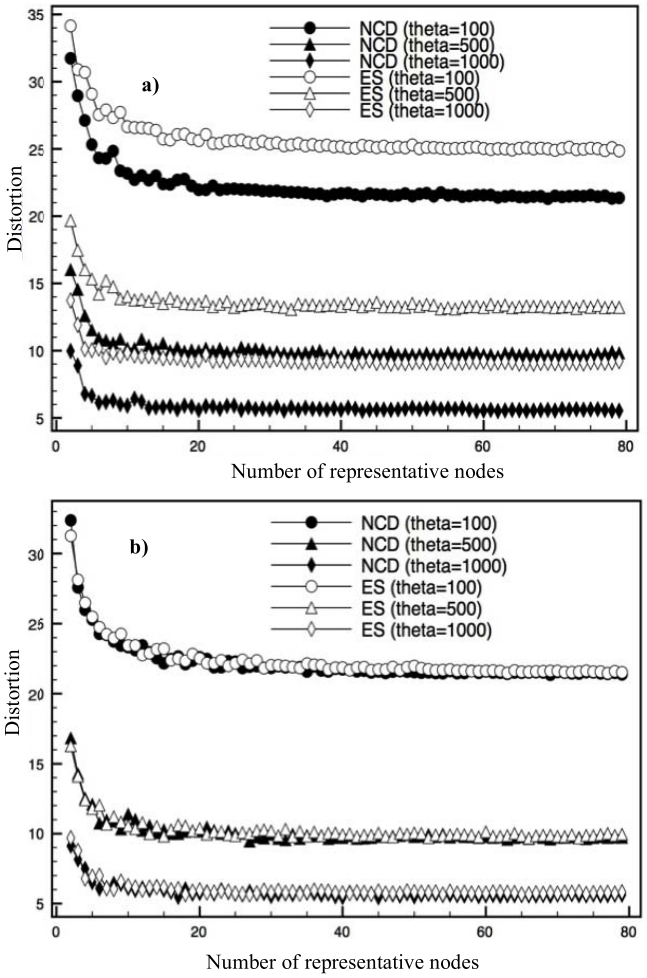

Fig. 9. Distortion of measurement for *ES* and *NCD* (as given by equations C3 and C2 respectively); $\sigma_W = 0.25$, 0.05, respectively, in a) and b); $\sigma_S = 5$ throughout. ES and NCD converge for lower values of σ_W.

and

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

Evaluating (C.1) term by term and simplifying yields:

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

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 goes to 0 ([28]); also see section V) and w.h.p. does not incur distortion of measurement. Here, we have assumed the range of values a phenomenon can take is split in intervals of length $\varepsilon = 2\beta\sigma_W$, with $\beta \geq 2$. Hence, per each group G_j, the estimation of measurement at the sink is $S'_j = S_j + 2\beta\sigma_W$. As before, $\boldsymbol{S}' = (1/k)\sum_{j=1}^{k} S'_j$ and plugging in (C.1) yields

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

Aside from the term $(2\beta\sigma_W/\sigma_S)^2$, ES achieves similar distortion to NCD as the number of representative groups/nodes increases and $\sigma_W < \sigma_S$. Figure 9 illustrates the numerical evaluation of $D_{ES}(k)$ and $D_{NCD}(k)$. The behavior of the two distortion functions is evaluated for varying number of representative nodes/groups, given the correlation model (20) and its parameters: $\theta_1 \in \{100, 500, 1000\}$, and $\theta_2 = 1$. The number of representative nodes in NCD is equal to the number of representative groups in ES. For both schemes, this number is set to the minimum, while achieving the desired distortion of measurement imposed by QoS constraints.

REFERENCES

[1] V. Shnayder, M. Hempstead, B. Chen, G. W. Allen, and M. Welsh, "Simulating the power consumption of large-scale sensor network applications," in *Proc. 2nd Int. Conf. Embedded Netw. Sensor Syst. (SenSys)*, New York, NY, USA, 2004, pp. 188–200.

[2] O. Landsiedel, K. Wehrle, and S. Gotz, "Accurate prediction of power consumption in sensor networks," in *Proc. IEEE Comput. Soc., 2nd IEEE Workshop Embedded Netw. Sensors (EmNets)*, Washington, DC, USA, Feb. 2005, pp. 37–44.

[3] T. Pering, Y. Agarwal, R. Gupta, and R. Want, "CoolSpots: Reducing the power consumption of wireless mobile devices with multiple radio interfaces," in *Proc. 4th Int. Conf. Mobile Syst., Appl. Services (MobiSys)*, New York, NY, USA, 2006, pp. 220–232.

[4] M. A. Razzaque, C. Bleakley, and S. Dobson, "Compression in wireless sensor networks: A survey and comparative evaluation," *ACM Trans. Sensor Netw.*, vol. 10, no. 1, 2013, Art. no. 5.

[5] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 316–329, Apr. 2006.

[6] M. Cardei, M. T. Thai, Y. Li, and W. Wu, "Energy-efficient target coverage in wireless sensor networks," in *Proc. IEEE INFOCOM*, Mar. 2005, pp. 1976–1984.

[7] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[8] X. Wu, Q. Wang, and M. Liu, "*In-situ* soil moisture sensing: Measurement scheduling and estimation using compressive sensing," *ACM Trans. Sensor Netw.*, vol. 11, no. 2, pp. 1–6, 2014.

[9] J. Wang, S. Tang, B. Yin, and X.-Y. Li, "Data gathering in wireless sensor networks through intelligent compressive sensing," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 603–611.

[10] S. Lee, S. Pattem, M. Sathiamoorthy, B. Krishnamachari, and A. Ortega, "Spatially-localized compressed sensing and routing in multi-hop sensor networks," in *Proc. Int. Conf. GeoSensor Netw. (GSN)*, 2009, pp. 11–20.

[11] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked Slepian–Wolf: Theory, algorithms, and scaling laws," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4057–4073, Dec. 2005.

[12] K. Yuen, B. Liang, and B. Li, "A distributed framework for correlated data gathering in sensor networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 578–593, Jan. 2008.

[13] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2006.

[14] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[15] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 1927–1938, Nov. 2003.
[16] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
[17] L. Simic, S. M. Berber, and K. W. Sowerby, "Partner choice and power allocation for energy efficient cooperation in wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sep. 2008, pp. 4255–4260.
[18] X. Li, M. Chen, and W. Liu, "Application of STBC-encoded cooperative transmissions in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 12, no. 2, pp. 134–137, Feb. 2005.
[19] D. Wu, Y. Cai, L. Zhou, and J. Wang, "A cooperative communication scheme based on coalition formation game in clustered wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1190–1200, Mar. 2012.
[20] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 411–426, Jan. 2010.
[21] B. Gopalakrishnan and N. D. Sidiropoulos, "Cognitive transmit beamforming from binary CSIT," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 895–906, Feb. 2015.
[22] C. Shepard *et al.*, "Argos: Practical many-antenna base stations," in *Proc. MOBICOM*, New York, NY, USA, 2012, pp. 53–64.
[23] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in Ad-Hoc networks of sensors," in *Proc. 7th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, New York, NY, USA, 2001, pp. 166–179.
[24] J. O. Berger, V. De Oliveira, and B. Sansó, "Objective Bayesian analysis of spatially correlated data," *J. Amer. Stat. Assoc.*, vol. 96, no. 456, pp. 1361–1374, 2001.
[25] J. B. Predd, S. R. Kulkarni, and H. V. Poor, "Regression in sensor networks: Training distributively with alternating projections," *Proc. SPIE*, vol. 5910, p. 591006, 2005.
[26] C. Chien, I. Elgorriaga, and C. McConaghy, "Low-power direct-sequence spread-spectrum modem architecture for distributed wireless sensor networks," in *Proc. Int. Symp. Low Power Electron. Design (ISLPED)*, 2001, pp. 251–254.
[27] L. Song and J. Shen, *Evolved Cellular Network Planning and Optimization for UMTS and LTE*. Boca Raton, FL, USA: CRC Press, 2010.
[28] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*. Redwood City, CA, USA: Addison Wesley, 1995.
[29] U. Madhow and M. B. Pursley, "Acquisition in direct-sequence spread-spectrum communication networks: An asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 903–912, Mar. 1993.
[30] D. H. Lehmer, "The machine tools of combinatorics," in *Applied Combinatorial Mathematics*, E. F. Beckenback, Ed. New York, NY, USA: Wiley, 1964, pp. 5–31.
[31] R. R. Rick and L. B. Milstein, "Noncoherent parallel acquisition in CDMA spread spectrum systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 1994, pp. 1422–1426.
[32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge Univ. Press, 2005.
[33] *Bayesian Logic (BLOG) Inference Engine*. Accessed: Jan. 1, 2016. [Online]. Available: http://people.csail.mit.edu/milch/blog/index.html
[34] M. Elhawary and Z. J. Haas, "Energy-efficient protocol for cooperative networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 2, pp. 561–574, Feb. 2011.
[35] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. 2nd Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2003, pp. 162–177.
[36] Memsic. *TelosB Mote Platform Datasheet*. Accessed: Jul. 2, 2017. [Online]. Available: http://www.memsic.com/userfiles/files/Datasheets/WSN/telosb_dataheet.pdf
[37] A. Jindal and K. Psounis, "Modeling spatially-correlated sensor network data," *ACM Trans. Sensor Netw.*, vol. 2, no. 4, pp. 466–499, 2006.

Milen Nikolov received the bachelor's degree in computer science and mathematics from the State University of New York, Brockport, NY, USA, and the M.Sc. and Ph.D. degrees from the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. He is currently a Post-Doctoral Researcher at the Institute for Disease Modeling, Bellevue, WA, USA, where his work comprises network models for malaria epidemiology and control. His research interests are in the areas of wireless communications, information, and disease transmission networks. He has been studying algorithms for networks with dynamic topologies, collaborative communication in information networks, and the data-driven dynamic models of contagious and vector-borne diseases over spatially connected populations.

Zygmunt J. Haas (S'84–M'88–SM'90–F'07) received the Ph.D. degree in electrical and computer engineering from Stanford University in 1988. In 1988, he joined AT&T Bell Laboratories in the Network Research Area, where he pursued research in wireless communications, mobility management, fast protocols, optical networks, and optical switching. In 1995, he joined the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, where he is currently a Professor. He heads the Wireless Network Laboratory, a research group with extensive contributions and international recognition in the area of ad hoc networks and sensor networks. He is an author of over 200 technical conference and journal papers and holds 18 patents in the areas of wireless networks and wireless communications, optical switching and optical networks, and high-speed networking protocols. His interests comprise mobile and wireless communication and networks, the modeling and performance evaluation of large and complex systems, and biologically inspired networks. He received a number of awards and distinctions, including the Best Paper Awards and the 2012 IEEE ComSoc WTC Recognition Award for his outstanding achievements and contribution in the area of wireless communications systems and networks. He has organized several workshops, delivered numerous tutorials at major IEEE and ACM conferences, and served as an Editor for several journals and magazines, including the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the *IEEE Communications Magazine*, and *Wireless Networks* (Springer). He has been a Guest Editor of several IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS issues and served as a Chair of the IEEE Technical Committee on Personal Communications.