Distributed Estimation in Wireless Sensor Networks with an Interference Canceling Fusion Center

Antonios Argyriou, *Senior Member, IEEE,* and Özgü Alay, *Member, IEEE*

***Abstract*—In this paper, we consider Distributed Estimation (DES) in a Wireless Sensor Network (WSN) and assume that the number of sensors in the WSN is larger than the available number of transmission slots. With classic DES, the sensors independently transmit the sampled digitized data. However, the WSN is an uplink multi-user channel where multiple sources share the channel for communicating data to a Fusion Center (FC). To this aim, we adopt the optimal communication scheme for this setup that suggests interfering transmissions and the use of Successive Interference Cancellation (SIC) at the FC. We propose a joint SIC decoder and linear Minimum Mean Square Error (MMSE) estimator for digital interfering transmission of correlated data. We further introduce an optimization framework that schedules and allocates power to the sensors optimally. We formulate the problem in two ways: an expected distortion minimization problem under a total power budget, and a transmission power minimization problem under a distortion constraint. For both cases, we consider the system performance under different operating conditions, and we demonstrate the efficiency of the proposed scheme compared to a system that employs optimized sensor selection under orthogonal transmissions.**

Keywords—Wireless sensor networks, linear distributed estimation, correlated data, interference cancellation.

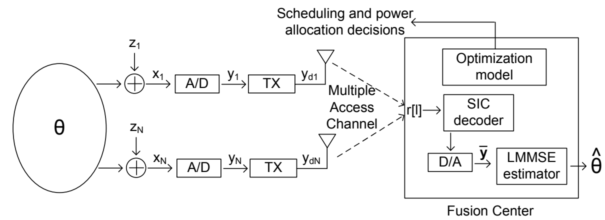


Fig. 1. WSN model for estimating the random signal θ. Each sensor i transmits y_{d_i}, the digitized and modulated version of the analog sample. Simultaneously transmitting sensors are decoded by the SIC decoder.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) have a wide range of applications including environmental monitoring, battlefield surveillance, smart grid monitoring, health care monitoring, home automation, farming, inventory tracking, etc. All these applications are based on the same fundamental task of sampling a random parameter and estimating it. Hence, it is important to keep track of the quality of the estimation accuracy through the Mean Square Error (MSE) distortion. Furthermore, estimation has to be power-efficient since in these applications the sensors are typically battery-operated. Improving power efficiency for a given MSE can be accomplished by exploiting the data correlation with Distributed Estimation (DES) algorithms that process the collected data jointly at a Fusion Center (FC). The dual problem of reducing the MSE subject to a power constraint can be more appropriate for certain applications. One class of DES algorithms that solve the previous problems and offer excellent performance, use only low-complexity linear processing [1]–[7].

The challenge today is that linear DES algorithms [1]–[7] have to operate in an environment where massive Machine Type Communication (MTC) and Internet of Things (IoT) applications require the deployment of large numbers of sensor nodes that all communicate to a FC. The problem is that even though the amount of data available increases as the number of sensors increases, it may not be possible to communicate them to the FC. Hence, it is critical to investigate how to optimize DES as the number of WSN nodes is increased.

This emerging DES scenario with multiple sensors communicating to a single FC could be modeled as a DES problem subject to a constraint on the communication rate. A subset of the literature on DES has studied this problem by employing compression [5]–[7], or with the equivalent solution of sensor selection/scheduling [1], [3], [8], which means that in both cases *the volume of the transmitted data is reduced*. A common feature of these works is that they assume orthogonal channel access from the sensors using Time Division Multiple Access (TDMA). TDMA is well-known to be suboptimal for achieving the capacity of the Multiple Access Channel (MAC) [9]. An alternative approach is to maximize the communication rate leading to *the transmission of more data* to the FC. To this aim, the authors in [10] considered transmit power allocation across the wireless sensors. The final solution was a protocol that schedules transmissions so that they interfere minimally. However, this is not a capacity-achieving scheme for the MAC. The optimal strategy that achieves the capacity of the fading MAC when the data sources are uncorrelated, is Successive Interference Cancellation (SIC) [9]. SIC decodes the received signal with the highest power first, while treating the remaining interfering signals as noise. SIC can also be optimized either by selecting the transmission power of the simultaneously transmitting sensors (power allocation problem), or by selecting the sensors that will transmit simultaneously (scheduling problem). The power allocation problem has been studied in the context of CDMA systems [11], while the optimal scheduling of sources for improving SIC has also been considered in the context of ad-hoc networks [12].

A different class of research works has also investigated DES in a MAC with non-orthogonal transmissions. Type-

Antonios Argyriou is with the Department of Electrical and Computer Engineering, University of Thessaly, Volos, Greece (Email: anargyr@ieee.org). Özgü Alay is with the Simula Research Laboratory, Oslo, Norway (Email: ozgu@simula.no).

Based Multiple Access (TBMA) has been proposed by Mergen and Tong [13], as well as by Liu and Sayeed [14], as a method to utilize the MAC in order to perform distributed detection or estimation. With TBMA, each sensor transmits over the MAC a different waveform depending on the type of the quantized observation. The authors assumed i.i.d. and uncorrelated measurements across the sensors. The detection problem for a deterministic signal in a MAC was considered by Li and Dai in [15] under the the same channel gain from the sensors towards the FC.

In this paper, we improve linear Minimum Mean Square Error (MMSE) estimation for correlated sensors in a WSN as their number is increased. To accomplish our goal, we first investigate whether the capacity of the MAC can be increased by exploiting data correlation. We show that unlike the Additive White Gaussian Noise (AWGN) channel, data correlation at the sources cannot be exploited to improve the capacity of the fading MAC. This first result drives the first system design choice in our paper: the separation of the linear MMSE signal estimation at the FC from the optimal capacity-achieving SIC-based digital receiver that does need to exploit correlation. The second design choice is on the system optimization where a subset of the sensors with the most valuable information is selected for transmission under the presence of bandwidth limitations. In our system, this calls for a novel cross-layer optimization of SIC and linear MMSE estimation. This concept drives the formulation of a novel problem formulation that considers jointly SIC and linear MMSE estimation. The formulation is enabled by the developed analytical expressions for the instantaneous packet loss probability of sensor transmissions, and the average MSE. The optimization problem is formulated and carried out in two forms: distortion minimization under a power constraint, and power minimization under a distortion constraint. For both problems, we provide the results under the optimal scheme and results for a low-complexity polynomial-time heuristic.

The contributions and main results of this paper are:

1) We show that for correlated sources the ergodic capacity Upper Bound (UB) of the fading MAC cannot be more that the capacity UB for uncorrelated sources. This result motivates our approach for not exploiting the signal correlation at the demodulator/decoder but at the signal estimator. This result has a direct impact on the design of potentially different demodulation/decoding and estimation algorithms.
2) We propose an algorithm for joint SIC decoding and MMSE estimation for correlated data in a WSN. The algorithm can operate as a stand-alone system without any sensor coordination (scheduling or power allocation). Interestingly, for low data correlation across the sensors, this system performs similarly to a state-of-the-art sensor selection algorithm that uses TDMA.
3) We propose a sensor scheduling and power allocation framework for minimizing the distortion or power. The framework is accompanied by a low complexity heuristic algorithm. Our optimization is enabled by an analytical MSE model of our joint SIC decoder and MMSE estimator.
4) Our results indicate that for lower correlation across the sensor data, the sub-optimality of TDMA increases rapidly as the number of WSN nodes increases. Furthermore, the MSE and power benefits of our scheme are also increased when the variance of the random signal is increased, i.e., when the signal is more random.

The rest of the paper is organized as follows. The system model and an overview of the proposed scheme is described in Section II. Section III serves as a detailed motivation of this paper. The proposed algorithm for joint SIC decoding and MMSE estimation is presented in Section IV. We jointly consider SIC and MMSE and derive closed-form expressions for the MSE that can be used for the optimization in Section V. The problem formulation and the implementation of the solution are presented in Sections VI and Section VII respectively. Finally, Section VIII provides the performance results and Section IX concludes this paper.

II. System Setup

We consider a WSN that consists of a set of nodes $\mathcal{N}$ with $|\mathcal{N}|$=N. Each sensor is making observations, x_i, on a random source signal θ with zero mean and variance σ_θ^2. The analog observation x_i is then digitized and transmitted to the FC as illustrated in Figure 1. Upon collecting all the digitized observations, FC's mission is to estimate θ.

Observation Model and Signal Compression. We assume that the sensors observe θ with different correlation ρ_i as $x_i = \rho_i\theta + z_i$. The sampling noise z_i is AWGN with zero mean, variance $\sigma_{z_i}^2$, and is uncorrelated across the sensors. The observations form the random vector $\mathbf{x} = [x_1 \ ...x_i... \ x_N]^T$, and if we similarly define the vectors $\boldsymbol{\rho}$ and $\mathbf{z}$ that contain the ρ_i's and z_i's respectively, then we can write $\mathbf{x} = \boldsymbol{\rho}\theta + \mathbf{z}$. The observations are then quantized. The input signal to the quantizer of sensor i is the analog sample x_i and after quantization the resulting signal is $y_i = \rho_i\theta + z_i + q_i$. In the above, q_i is the quantization noise and is assumed independent across sensors because it is performed locally at each sensor without coordination[1]. In the quantizer, $2^{\overline{R}_i}$ representation levels are used per source sample (or $\overline{R}_i$ bits). With a uniform probabilistic quantizer the upper bound of the variance of the quantization noise at sensor i is $\sigma_{q_i}^2 = \frac{W^2}{(2^{\overline{R}_i}-1)^2}$, where $2W$ is the range of the sensed signal.

Communication Model. With K source samples, the total number of bits that must be communicated is $K\overline{R}_i$ over the L time-domain samples (see Figure 2). These bits are coded with a capacity-achieving channel code and modulated with a PSK constellation that have a combined spectral efficiency of R_i bits/symbol. So the previous discussion leads to $K\overline{R}_i$=LR_i. The combined effect of channel coding and digital modulation on the bits of the digitized samples is formally expressed for a specific symbol transmitted during time-domain sample l as:

$$y_{d_i}[l] = \text{CC-PSK}(L, R_i, \text{input bits}) \tag{1}$$

[1]This approximation, that is followed in the literature [1], [7] for tractability, becomes more accurate for smaller quantization steps.

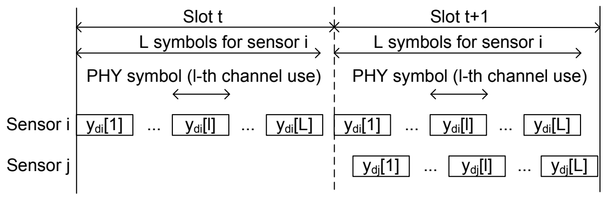


Fig. 2. Modeling packet transmissions in the WSN under different channel access assumptions: orthogonal (left), and asynchronous interfering (right).

This function expresses the channel coding and the modulation in a compact form. As an example, for a sample of 8 bits and uncoded QPSK, 4 time domain samples/symbols will be produced each containing 2 bits. Similary, other PHYs could be modeled with this approach (e.g., CDMA).

Channel Model. The transmission of a packet takes place over a wireless link with slow flat Rayleigh fading. Hence, $h_i[l]$=h_i for every time-domain sample/symbol during the transmission of a packet, and $|h_i|$∼Ray($\mathbb{E}[|h_i|^2]$). The Rayleigh fading channel is characterized by the average received power that is defined as $\mathbb{E}[|h_i|^2]$=$1/dist^a$ where $dist$ is the distance between the sensors and the FC, and a is the path loss exponent set to 3. The fading levels are accurately measured at the FC, while the sensors/transmitters do not require any channel knowledge. The channel is constant for multiple packet transmissions [1]–[3], [16]. Finally, the vector $\mathbf{h}$ packs the channel gains from all the sensors.

Channel Access Schemes. We assume that a set network-to-FC slots $\mathcal{T}$ with $|\mathcal{T}|$=T are available for transmitting packets, each consisting of L time-domain samples. The notion of a slot in this paper represents a modeling tool that allows us to capture asynchrony between several packet transmissions, and does not correspond to the complete duration of a packet transmission at the PHY (Figure 2). Two channel access schemes are examined. First, TDMA with orthogonal transmissions where sensors access the channel sequentially. For TDMA, the received signal from sensor i at the FC for the l-th time-domain sample is

$$r[l] = \sqrt{P_i} h_i y_{d_i}[l] + w[l], \tag{2}$$

where $w[l]$ is the noise sample at the FC that is AWGN with zero mean and variance σ_w^2. Also P_i is the transmit power at sensor i. Since the power of the PSK symbol is equal to P_i, we set $\sigma_{y_{d_i}}^2$=1 to avoid complications in the derived expressions. The average Signal-to-Noise Ratio (SNR) per symbol is

$$\text{SNR}_i = P_i \, \mathbb{E}[|h_i|^2]/\sigma_w^2 \tag{3}$$

With TDMA, a packet transmission requires L time-domain samples and so T sensors can transmit. TDMA communication is typically used by the vast majority of today's WSN systems while any signal processing algorithms are developed on top of the TDMA scheme. This is consistent with slotted channel access schemes for WSNs. Formally, in this case the set of sensors that transmit is $\mathcal{S}$ while for any given slot t, a subset of sensors $\mathcal{S}_t$ can transmit. The cardinality of this set is one ($|\mathcal{S}_t|$=1) since only one sensor can transmits in that slot.

The second scheme is our proposed approach that adopts interfering transmissions within each slot as illustrated in Figure 2 (right). At any given slot, a subset of sensors $\mathcal{S}_t$ can transmit. If we focus on a tagged sensor i, then the received time-domain sample l at the FC is:

$$r[l] = h_i \sqrt{P_i} y_{d_i}[l] + \sum_{j \in \mathcal{S}_t / \{i\}} h_j \sqrt{P_j} s_{d_j}[l] + w[l] \tag{4}$$

Here, $s_{d_j}[l]$ is the signal contribution of a sensor j during time-domain sample l, that can be asynchronous with the packet transmitted from sensor i (as illustrated in Figure 2 (right)). As we will later see, this potentially asynchronous situation is irrelevant for decoding the tagged sensor i since the contribution of all the other sensors will be treated as noise (i.e., the symbols of multiple sensors are not decoded jointly). To ensure a fair comparison with the TDMA, we impose a sum transmit power constraint P that is enforced over each time-domain sample as a system parameter.

Synchronization. We assume that all sensors have a local oscillator synchronized to the receiver carrier frequency. On the other hand, our system does not require time synchronization since the packets are decoded with SIC at the level of complete packets, i.e., there is no symbol-level decoding with SIC.

Transmission/Estimation Schemes. To illustrate the effects of interference on correlated sources, we study the following schemes in this paper:

1) ORTH-MMSE-OPT: Each sensor compresses and transmits its own signal directly to the destination in its own time slot (Figure 2 (left)). The optimum set of sensors are selected for transmission given the bandwidth constraint expressed through the number of available slots T. The FC then uses linear MMSE estimation.

2) SIC-MMSE: Each sensor compresses its own signal, but in this mode, interfering transmissions are allowed. The sensors transmit and interfere in an uncoordinated fashion and the result is an arbitrary interference pattern (Figure 2 (right)). Next, SIC decoding is applied in the digital packets at the FC and as a next step, the FC uses MMSE estimation. This scheme offers significant practical implementation advantages: Any sensor is free to transmit during any time slot it desires, and the recovery from the implications of such an approach are in full responsibility of the receiver/FC.

3) SIC-MMSE-OPT: Compression, and correlation exploitation with MMSE is exercised as in the last scheme. However, the optimum set of interfering sensors in specific slots (scheduling) and their transmission power (power control) are selected under the scheme in Figure 2 (right). In particular the FC selects the sensors that will transmit in such a way that they are decodable with SIC, but also the decoded packets have the highest impact on the distortion minimization.

III. Motivation

In this section, we review first the limitations of linear MMSE estimation under orthogonal transmissions in multiple access WSNs. Next, we discuss the limitations of employing capacity-achieving multi-user communication (e.g., interfering transmissions under SIC decoding [9]). Two critical observations motivate the need for our proposed SIC-MMSE-OPT scheme.

A. Limitations of Orthogonal Transmissions

Consider a WSN with T available slots and a set $\mathcal{N}$ of $|\mathcal{N}|$ sensors, where $|\mathcal{N}|=N > T$. For any set of sensors $\mathcal{S} \subset \mathcal{N}$ that transmit, we clearly need $|\mathcal{S}|=T$. The baseband received signal model is that of (2). Under linear MMSE estimation, and if all the sensor observations are available at the FC, the MSE can be easily proven to be a fractional expression that is inversely proportional to the number of available observations. In particular the MSE under the linear MMSE estimator is can be calculated for our data model as [17]:

$$\begin{aligned} \text{MSE} &= \text{Tr}(\boldsymbol{\rho}^H(\boldsymbol{\Sigma}_{\mathbf{z}} + \boldsymbol{\Sigma}_{\mathbf{q}})^{-1}\boldsymbol{\rho} + \boldsymbol{\Sigma}_{\theta}^{-1})^{-1} \\ &= \frac{\sigma_\theta^2}{\sigma_\theta^2 \sum_{i=1}^N \frac{\rho_i^2}{\sigma_{z_i}^2 + \sigma_{q_i}^2} + 1} \end{aligned} \tag{5}$$

With this orthogonal transmission scheme (denoted as ORTH), and by extending the previous result to accommodate packet losses, the result the MSE becomes:

$$\text{MSE}_\theta^{\text{ORTH}}(\mathcal{S}) = \frac{\sigma_\theta^2}{\sigma_\theta^2 \sum_{i\in\mathcal{S}}(1-\pi_i^{\text{ORTH}})\frac{\rho_i^2}{(\sigma_{z_i}^2+\sigma_{q_i}^2)} + 1} \tag{6}$$

In the above, π_i^{ORTH} is the *outage probability* for packets transmitted from sensor i and is a metric that is used for characterizing slow fading channels [9]. In the denominator, the fraction in the summation essentially corresponds to the *SNR contribution* of each sensor. Note also that the MSE is a function of the set $\mathcal{S}$ of transmitting sensors.

By further elaborating on the outage probability, it can actually lead to a closed-form result:

$$\begin{aligned} \pi_i^{\text{ORTH}} &= \Pr\left\{\log_2\left(1+\frac{P_i|h_i|^2}{\sigma_w^2} < R_i\right)\right\} \\ &= 1-\exp\left(\frac{-(2^{R_i}-1)}{P_i/\sigma_w^2}\right) \end{aligned} \tag{7}$$

The last equation follows from $|h_i|$ being Rayleigh.

By considering that only T sensors can transmit, the optimal scheduling policy is to select the best sensors from the set $\mathcal{N}$ given the limited number of slots. The term "best" can be translated in terms of the ratio $(1-\pi_i^{\text{ORTH}})\frac{\rho_i^2}{\sigma_{z_i}^2+\sigma_{q_i}^2}$. The linear MMSE estimator collects these contributions and combines them optimally. Thus, the sensor that has the highest ratio, has the highest *SNR contribution* which means higher MSE reduction in (6). Consequently, from (6) we can see that an ordering of the sensors in terms of their *SNR contribution*, and then the selection of the first T, is the optimal strategy in this case. This approach also serves as our baseline scheme.

B. Limitations of Multi-User Communication

Our intention is to calculate the capacity of the AWGN and fading MAC channels in order to determine the maximum transmission rate in the WSN when the data is correlated.

Theorem 1. *The capacity UB of the fading MAC for correlated sources is equal to the capacity UB for uncorrelated sources.*

Proof: To prove this result, we reuse our existing WSN model and focus on a slot t where the set of sensors that transmit simultaneously is $\mathcal{S}_t$. Also assume symbol-level synchronization just for this proof to illustrate the idea. When the correlation of simultaneously transmitting sources is captured with $\boldsymbol{\rho}$, the AWGN MAC capacity $C_{\text{AWGN-MAC}}(\boldsymbol{\rho})$, calculated in [9], is a function of the correlation vector $\boldsymbol{\rho}$. Hence, the multi-user capacity for correlated data transmission in the AWGN MAC depends on the data correlation vector and is higher than the case of uncorrelated data, i.e., $C_{\text{AWGN-MAC}}(\boldsymbol{\rho}) \geq C_{\text{AWGN-MAC}}(\mathbf{0})$.

Regarding the fading MAC, and due to the random nature of the channel, there is the notion of the *ergodic capacity* that is the capacity averaged over several channel realizations. To achieve this capacity one must apply a channel coding across all these realizations. We can calculate the UB as follows:

$$\begin{aligned} C_{\text{Fading-MAC}}(\boldsymbol{\rho}) &= \mathbb{E}[\log_2(1+SNR)] \leq \log_2(1+\mathbb{E}[SNR])] \\ &= C_{\text{Fading-MAC}}^{\text{UB}}(\boldsymbol{\rho}) \\ &= \log_2(1+\frac{P_{\text{Fading-MAC}}^{\text{avg}}(\boldsymbol{\rho})}{\sigma_w^2}) \end{aligned} \tag{8}$$

The average power $P_{\text{Fading-MAC}}^{\text{avg}}$ of the *aggregate useful received signal* determines the capacity UB of the MAC [9]. Interfering transmissions and SIC decoding can achieve the capacity of the AWGN MAC and the ergodic capacity of the fading MAC [9]. To calculate this UB under Rayleigh fading in (8), recall that the channel model for multi-user transmission is that of (4). Hence:

$$\begin{aligned} &P_{\text{Fading-MAC}}^{\text{avg}}(\boldsymbol{\rho}) = \mathbb{E}\left[\{\sum_{i\in\mathcal{S}_t}\sqrt{P_i}h_i y_{d_i}[l]\}\{\sum_{i\in\mathcal{S}_t}\sqrt{P_i}h_i y_{d_i}[l]\}^*\right] \\ &= \sum_{i\in\mathcal{S}_t} P_i\,\mathbb{E}[|h_i|^2]\sigma_{d_i}^2 + \mathbb{E}\left[\sum_{i\in\mathcal{S}_t}\sum_{j\in\mathcal{S}_t/\{i\}}\sqrt{P_iP_j}h_ih_j^* y_{d_i}[l]y_{d_j}^*[l]\right] \\ &= \sum_{i\in\mathcal{S}_t} P_i\,\mathbb{E}[|h_i|^2]\sigma_{d_i}^2 = \sum_{i\in\mathcal{S}_t} P_i\,\mathbb{E}[|h_i|^2] \end{aligned} \tag{9}$$

The last follows from the uncorrelated Rayleigh channels and is independent of $\boldsymbol{\rho}$. Thus, the corresponding UB expression in (8) becomes:

$$\begin{aligned} C_{\text{Fading-MAC}}^{\text{UB}}(\boldsymbol{\rho}) &= \log_2(1+\frac{\sum_{i\in\mathcal{S}_t}P_i\,\mathbb{E}[|h_i|^2]}{\sigma_w^2}) \\ &= C_{\text{Fading-MAC}}^{\text{UB}}(\boldsymbol{\rho}=\mathbf{0}) \\ &= C_{\text{AWGN-MAC}}(\boldsymbol{\rho}=\mathbf{0}) \end{aligned} \tag{10}$$

We observe that the ergodic capacity UB of the fading MAC with correlated sources is equal to the capacity of both the fading MAC UB and AWGN MAC but with uncorrelated signals. Or we can condense these results as follows:

$$C_{\text{Fading-MAC}}(\boldsymbol{\rho}) \leq C_{\text{Fading-MAC}}^{\text{UB}}(\mathbf{0}) \leq C_{\text{AWGN-MAC}}(\boldsymbol{\rho}) \tag{11}$$

Therefore, we cannot exploit the correlation between the source signals in a Rayleigh fading channel as effectively as in AWGN channel where we would have a capacity gain. ■

The previous results highlight the problems with the transmission of correlated data from many sources in uncorrelated Rayleigh fading channels. First, improving the MSE depends on our ability to collect more data and to combine them optimally with linear MMSE. Second, interference cannot be exploited for increasing the multi-user capacity, i.e., we cannot transmit more data. Motivated by these observations, *our objective in this paper is to use interference for increasing the number of* high value *transmissions within T slots. This is accomplished by optimizing SIC in such a way that the receiver decodes the packets from the sensors that contribute more to the MSE reduction.*

IV. Joint SIC Decoding and MMSE Estimation

Based on the ideas developed in the previous section, we now present the joint SIC decoding and MMSE estimation algorithm. For all the expressions we derive, the channel access scheme we presented in Figure 2 (right) is used.

A. Successive Decoding Algorithm for Interfering Sensors

Let us first describe the algorithm operations during the t-th slot where the set of the sensors that transmit is $\mathcal{S}_t$. After the sensors transmit, the FC attempts to decode the transmitted packets by employing ordered SIC (OSIC). That is, the packet with the highest energy/bit is decoded first while the other packets are treated as noise regardless of their content [9]. For exposition purposes, let us assume that the packet transmitted from sensor i and received at the FC has the highest power. The instantaneous Signal to Interference plus Noise Ratio (SINR) for the packet from sensor i during slot t, and for a given channel vector $\mathbf{h}$, can be derived from (4):

$$\text{SINR}_i(\mathbf{h}, \mathcal{S}_t) = \frac{P_i|h_i|^2}{\sum_{j\in\mathcal{S}_t/\{i\}} P_j|h_j|^2 + \sigma_w^2} \quad (12)$$

The SINR is calculated only for sensor i while the remaining sensor transmissions are considered as interference. Also it is important to recall that P_i, P_j are the transmission power per time-domain sample and so $\text{SINR}_i(\mathbf{h}, \mathcal{S}_t)$ is normalized per time-domain sample.

Regarding the actual packet-level decoding it proceeds as follows: The start of each packet can be identified with a preamble correlation operation [18]. After the packet with the highest energy/bit is identified, a Matched Filter (MF) for the specific sensor i with its known channel gain h_i is used. Then, Hard Decision Decoding (HDD) is used for detection. After decoding the LR_i bits of the packet correctly, then the packet is removed from the aggregate signal in (4). Note that according to the theoretical description of SIC in §6.1 of [9], and a trace-driven study that quantifies the prospective gains of SIC [12], OSIC operates precisely as described above, i.e., at the packet-level and without requiring symbol-level synchronization. The same process is continued until it is completed for all the transmitted packets in slot t. If a packet is not decoded successfully, it cannot be removed from the aggregate signal and the SIC decoding chain fails. The above may of course be true even if the instantaneous SINR in (12) is more than the required packet decoding threshold [19]. This is something that is considered in our simulations, i.e., we do not assume ideal SIC decoding. The final result will be that several packets will be decoded at the receiver depending on the performance of SIC over the T available slots. Hence, the FC will have available multiple digital observations that are mapped to their corresponding quantization level through D/A conversion (also illustrated in Figure 1). We denote the available quantized signals from this step as the vector $\bar{\mathbf{y}}$. This mapping is necessary in order to proceed to the estimation.

B. Correlation Exploitation with Linear MMSE Estimation

The next step in our proposed joint decoding/estimation algorithm is to exploit the knowledge of the data model and the correlation that exists in all the decoded quantized signals $\bar{\mathbf{y}}$. Recall that our final goal is to estimate the random variable θ from the several digitized decoded signals that are available in $\bar{\mathbf{y}}$. Since we have a number of observations equal to the number of decoded packets, the data model becomes:

$$\bar{\mathbf{y}} = \bar{\boldsymbol{\rho}}\theta + \bar{\mathbf{z}} + \bar{\mathbf{q}} \quad (13)$$

Similarly with $\bar{\mathbf{y}}$, the bar in all the vectors denotes the subset of the data model that corresponds to decoded packets (e.g., $\bar{\mathbf{q}}$ contains the q_i's of the decoded sensors). Next, we employ a linear MMSE estimator for the received digital signals. So the proposed estimator is:

$$\hat{\theta} = (\boldsymbol{\rho}^H(\boldsymbol{\Sigma}_\mathbf{z} + \boldsymbol{\Sigma}_\mathbf{q})^{-1}\boldsymbol{\rho} + \boldsymbol{\Sigma}_\theta^{-1})^{-1}\boldsymbol{\rho}^H(\boldsymbol{\Sigma}_\mathbf{z} + \boldsymbol{\Sigma}_\mathbf{q})^{-1}\bar{\mathbf{y}} \quad (14)$$

The covariance matrix of the source signal vector (denoted as $\boldsymbol{\Sigma}_\theta$) is actually a scalar, i.e., $\boldsymbol{\Sigma}_\theta$=$\sigma_\theta^2$. This covariance can be known, or it can be calculated online as we do in this paper.

V. MSE under Joint SIC/MMSE Decoding

In the last section, we discussed how the decoding and estimation algorithms operate to obtain the desired estimate $\hat{\theta}$. Our goal now is to model the performance of the previous scheme, so that we can optimize it. For the optimization of the joint SIC and MMSE, in the same spirit with our previous derivations for (6), the average MSE for the complete WSN that consists of sensors that interfere can be expressed as:

$$\text{MSE}_{\theta|\mathbf{h}}^{\text{INT}}(\mathcal{S}) = \frac{\sigma_\theta^2}{\sigma_\theta^2 \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{S}_t} \frac{(1-\pi_{i|\mathbf{h}}^{\text{INT}}(\mathcal{S}_t))\rho_i^2}{\sigma_{z_i}^2+\sigma_{q_i}^2} + 1} \quad (15)$$

In the previous equation

$$\pi_{i|\mathbf{h}}^{\text{INT}}(\mathcal{S}_t) = \begin{cases} 1 \text{ if } \log_2\left(1 + \text{SINR}_i(\mathbf{h}, \mathcal{S}_t)\right) < R_i \\ 0 \text{ otherwise} \end{cases} \quad (16)$$

indicates if the instantaneous rate, expressed through the Shannon formula, can meet the desired communication rate of R_i bits/symbol for sensor i, and $\text{SINR}_i(\mathbf{h}, \mathcal{S})$ expresses the average SINR during the complete duration of the transmitted packet and was defined in (12). Note that π is a function of the *interference pattern/transmission schedule* $\mathcal{S} =$

$\{\mathcal{S}_1, ..., \mathcal{S}_l, ..., \mathcal{S}_T\}$, and the channel gain $\mathbf{h}$. The above expression essentially says that the MSE depends on the specific schedule of the sensors that belong to the WSN. Therefore, in order to minimize (15) for a given channel realization $\mathbf{h}$, there is a need to control $\pi_{i|\mathbf{h}}^{INT}(\mathcal{S})$.

For the orthogonal case, where each sensor digitizes and compresses the signal while linear MMSE is executed after all the signals are received at the FC, the MSE expression is again that of (15) but in this case the outage probability is $\pi_{i|\mathbf{h}}^{\mathrm{ORTH}}(\mathcal{S}_t)$. With orthogonal transmissions, even if the system performs optimal sensor selection it can select at most T sensors, i.e., the maximum cardinality of the set $\mathcal{S}$ is T. Next, we formulate the optimization problems.

VI. Performance Optimization of Joint SIC Decoding and MMSE Estimation

Before formulating the optimization problem, we first discuss the intuition behind our optimization. For our first objective where we consider distortion minimization under a power constraint, each sensor i is assumed to transmit the digital packet at power level P_i. To be fair, a limit on the transmit power of each time-domain sample or PHY symbol is considered. When the distortion is minimized for the TDMA scheme, this will lead to the assignment P_i=P where P is the allowed power per time-domain sample. Each sensor will transmit at the maximum allowed power, since only a single sensor is allowed to transmit during each slot[2]. However, with interfering transmissions, the available power P will be distributed among multiple sensors. With SIC, this distribution can be done in such a way that more packets can be decoded at the FC, hence increasing the transmission efficiency of the system. We will illustrate this with an example. Consider a WSN with two sensors and one transmission slot. The first option is that one sensor transmits at a power level P_i=P and enjoys a rate $\log(1+\frac{P}{\sigma_w^2})$. If this rate is higher than R_i then we have a successful transmission with high probability (precise value in (7)). However, if two sensors transmit at power levels P_1, and P_2 respectively, with $P_1 + P_2 \leq P$, with SIC they can both enjoy a sum-rate equal to $\log(1+\frac{P}{\sigma_w^2})$ [9]. This can happen if P_1, P_2 are selected such that $\log(1 + \frac{P_1}{P_2+\sigma_w^2}) > R_i$, and $\log(1 + \frac{P_2}{\sigma_w^2}) > R_i$, i.e., the FC can decode both packets with high probability. Thus, this is an optimized power allocation so that more packets are decodable according to SIC. This has been a typical optimization approach for SIC [11], [20].

For our system the question is how should the sensors be selected at a given time slot t, so that the information sent is decodable but also leads to the minimum MSE. Next, we will formulate this optimization problem. We consider two different cases. First, we minimize the MSE under a fixed power constraint and then we consider power minimization under a distortion constraint.

Minimizing Distortion under a Power Constraint We first formulate the problem of MSE distortion minimization for a fixed power consumption that is imposed per time-domain sample. The distortion minimization problem is equivalent to maximizing the MSE reduction (the denominator in (15)) by selecting the sensors to transmit in each slot t among all the available N. This decision is captured with binary variable $a_{i,t}$ that indicates if the FC is able to cancel the signal from sensor i transmitted in slot t. Sensors are scheduled to interfere only if SIC can decode them. Our goal is to increase the number of decoded packets by respecting an SNR packet decoding constraint while also ensuring that these packets/signals contribute more to minimizing the MSE. Therefore, if a packet from a certain sensor cannot be canceled with SIC, it is not scheduled in this specific slot t while it can be scheduled in another slot. We also define the continuous variable $P_{i,t}$ that indicates the transmission power of sensor i in slot t. Hence, the decision vectors are: $\mathbf{a}$=$\big(a_{it} \in \{0,1\} : i \in \mathcal{N}, t \in \mathcal{T}\big)$ and $\mathbf{P}$=$\big(P_{it} \in \{0,1\} : i \in \mathcal{N}, t \in \mathcal{T}\big)$. This problem is formulated as a Mixed Integer Linear Program (MILP):

$$\textbf{DIST:}\ \max_{\mathbf{a},\mathbf{P}} \sum_{t=1}^{T}\sum_{i=1}^{N} \frac{\rho_i^2}{\sigma_{z_i}^2 + \sigma_{q_i}^2} a_{i,t},$$

$$\text{s.t.}\ \frac{P_{i,t}|h_i|^2}{\sigma_w^2 + \sum_{k\in N, k>i} P_{k,t}|h_k|^2} \geq 2^{R_i} - 1, \forall i \in \mathcal{N}, t \in \mathcal{T}\ \text{(C1)}$$

$$\sum_{k\in N, k>i} a_{k,t} \leq c_i a_{i,t}, \forall i \in \mathcal{N}, t \in \mathcal{T}\ \text{(C2)}, \sum_{t=1}^{T} a_{i,t} \leq 1, \forall i \in \mathcal{S}\ \text{(C3)}$$

$$\sum_{i=1}^{N} P_{i,t} \leq P, \forall t \in \mathcal{T}\ \text{(C4)}, P_{i,t} \leq a_{i,t} P\ \text{(C5)}$$

Constraint 1 (C1) is the SINR constraint for sensor i scheduled in slot t. For the SINR constraint to be satisfied, R_i bits/symbol need to be communicated successfully. To demodulate these symbols, the SINR in (12) must be at least equal to $2^{R_i} - 1$. This leads to the formulation of C1 in our problem formulation. In the denominator, the summation term contains all the other sensors that can potentially transmit in the same slot and have not been canceled yet. These sensors are accounted for as destructive interference (k>i accounts for these sensors that have been ordered in advance). If a sensor is not scheduled then this constraint does not need to be satisfied. That is why we have $N \times T$ of these constraints. Constraint 2 (C2) ensures that the optimal SIC decoding order is followed. Here, we define c_i as the number of links after i, in the sorted sequence of the values of $|h_i|^2/(2^{R_i} - 1)$. C2 ensures that if sensor i is not scheduled in slot t through the variable $a_{i,t}$, then none of the remaining sensors can be scheduled in that particular slot. The reason is that they have lower energy/bit ratio by our ordering requirement which means that the packet cannot be decoded. Of course if $a_{i,t}$=1 then all the remaining sensors can be potentially scheduled. The specific way of populating c_i is central in making the problem solvable: The power allocation variable $P_{i,t}$ can effectively change the energy/bit in equation (12) for a sensor i. However, if we cannot increase the power $P_{i,t}$ for a sensor i to a level that ensures that C1 is valid, then it is impossible to do that for a sensor i' that has lower energy/bit ratio $|h_{i'}|^2/(2^{R_{i'}} - 1)$. Hence, by using

[2] Even though transmitting at the maximum allowed power may not be needed for certain sensors, still it reduces the outage probability as per our previous derivations (see (7)).

R_i in this ratio, the event that power allocation decisions in our optimization change the optimal SIC ordering are avoided. Constraint 3 (C3) ensures that sensor i transmits at most once within the T slots. Constraint 4 (C4) is the power constraint per time-domain sample that cannot exceed P, while constraint 5 (C5) ensures that no power is allocated to non-scheduled sensors.

Minimizing WSN Power under a Distortion Constraint. Now we formulate the power minimization problem under a distortion constraint. To be consistent with our previous notation, the power is minimized over the whole transmission time. We define an MSE distortion threshold D that is the minimum distortion constraint as in [1]. Then the power minimization problem is formulated as:

$$\textbf{PWR:} \min_{\mathbf{a},\mathbf{P}} \sum_{t=1}^{T}\sum_{i=1}^{N} P_{i,t}, \text{ s.t. (C1), (C2), (C3)},$$
$$\sum_{t=1}^{T}\sum_{i=1}^{N} \frac{\rho_i^2}{\sigma_{z_i}^2+\sigma_{q_i}^2} a_{i,t} \geq \frac{1}{D}-\frac{1}{\sigma_\theta^2} \text{ (C4')}, \text{ (C5)}$$

Here, constraints C1-C3, and C5 are the same as in the **DIST** problem. The new fourth constraint (C4') ensures that the distortion threshold in (15) is met.

Minimizing Power for ORTH. The previously described optimization approach is also applied for the ORTH-MMSE-OPT that considers orthogonal transmission and sensor selection. In particular, we present representatively its formulation under the PWR objective:

$$\min_{\mathbf{a},\mathbf{P}} \sum_{t=1}^{T}\sum_{i=1}^{N} P_{i,t}, \text{ s.t. } \frac{P_{i,t}|h_i|^2}{(2^{R_i}-1)\sigma_w^2} \geq 1, \forall i \in \mathcal{N}, t \in \mathcal{T} \text{ (C1')},$$
$$\sum_{i=1}^{N} a_{i,t} \leq 1, \ \forall t \in \mathcal{T} \text{ (C2')}, \text{ (C3)}, \text{ (C4')}, \text{ (C5)}$$

Note that the main difference is the removal of decoding order constraints, a simpler SNR threshold constraint (C1'), while the constraint for one sensor transmission/slot (C2') was added.

VII. MILP Relaxation and Approximation

Even though designing the most efficient algorithm for solving the MILP is not the main focus of this paper (since this is a hard problem in general), we present a solution by designing an approximation algorithm. Since there are no polynomial time algorithms for solving MILPs, we first relax the optimization problems so that a Linear Program (LP) can be solved. LPs can be solved in polynomial time with interior point methods. Thus, we allow the binary variables $a_{i,t}$ to take any value between 0 and 1. After the LP is solved the results of the relaxed LP, that consists of a set of continuous values between 0 and 1, are converted to binary values. We adopt the Randomized Rounding (RR) approach that assigns the final binary values with a certain probability. Let $\tilde{a}_{i,t}$ denote the solutions of the LP, the binary values are approximated as:

$$\hat{a}_{i,t} = \begin{cases} 1 & P_r[1] = \tilde{a}_{i,t} \\ 0 & P_r[0] = 1-\tilde{a}_{i,t} \end{cases} \tag{17}$$

This rule means that the final binary solution $\hat{a}_{i,t}$ is equal to 1 with probability $\tilde{a}_{i,t}$ and equal to 0 with probability $1-\tilde{a}_{i,t}$. Values of $\tilde{a}_{i,t}$ closer to 1 increase the probability that a binary 1 is assigned. This process ensures that the cost of the MILP and LP solutions are the same. Since some constraints might be violated after (17), they are first verified.

A side-effect of the relaxed problem formulation is that the LP actually provides a solution that gives non-zero values for $a_{i,t}$ for all the sensors and all the slots. Since the channel does not change throughout the scheduling period of T slots, $a_{i,t}$ obtains fractional values that are equal across all the slots even though their sum is less than 1. So the heuristic algorithm that we propose next is based on the application of the RR procedure on a time slot basis. For each slot, we first obtain the results of the RR algorithm, we then perform a constraint check and keep the final result. Regarding the constraint check, C1 cannot be violated by RR since it only contains the continuous variable $P_{i,t}$ and so the related solutions are retained. Furthermore, when the value of $\hat{a}_{i,t}$ is equal to 0 for a certain sensor, the only constraint that must be checked is the ordering through C2, i.e., a sensor with lower ranking in the list c_i may not be scheduled. So in this case we also set $\hat{a}_{i,t}$ equal to zero for all the remaining sensors in slot t. Of course in this case the power $P_{i,t}$ is also set to zero according to C5. Finally, to ensure C3, i.e., a sensor is only activated once within T, then if $\hat{a}_{i,t}$=1 we set $\hat{a}_{i,t'}$=0,$\forall t' \in \mathcal{T}/\{t\}$. The algorithm proceeds by removing all the already scheduled sensors from the constraints. Subsequently, it repeats the same constraint checks for the remaining sensors.

Correlation Estimation. Obtaining the value for the correlation coefficients of each sensor might also be a non-trivial task. To make our system fully implementable in practice, we employ an estimator of the correlation coefficient vector $\boldsymbol{\rho}$ at the FC (recall that this is a deterministic vector that contains the correlation coefficients). Based on the available measurements, we continuously update $\mathbf{\Sigma_y}$. For known AWGN and quantization variance at the sensors, this means we also know the diagonal matrices $\mathbf{\Sigma_z}, \mathbf{\Sigma_q}$. From (13) we can obtain the expression for the covariance matrix as $\mathbf{\Sigma_y} = \boldsymbol{\rho}\boldsymbol{\rho}^H\sigma_\theta^2 + \mathbf{\Sigma_z} + \mathbf{\Sigma_q}$, from which we only need σ_θ^2.

VIII. Performance Evaluation

In our setup, the sensors are spread randomly and uniformly in a disc, and in the center there is the FC. The spatial correlation model is created so that it represents reduced correlation as nodes are farther away from the center of the disc, i.e., $\rho_i = \rho^{dist(i)}$, where $dist$ is normalized in the range $[0, 1]$. Different values for ρ are tested. The other parameters are set as follows $\sigma_{z_i}^2$=σ_z^2=σ_w^2 unless otherwise noted. Also σ_w^2 is controlled through the transmit SNR given in (3). σ_θ^2 may take different values but it is equal to 10 unless otherwise specified. We also used $\bar{R}_i$= 8 bits/source sample, the range of the signal was W=1 Volt, and R_i=1, i.e., BPSK modulation. Also 100 bits are transmitted (K=12 samples per packet).

A. Results for Distortion Minimization

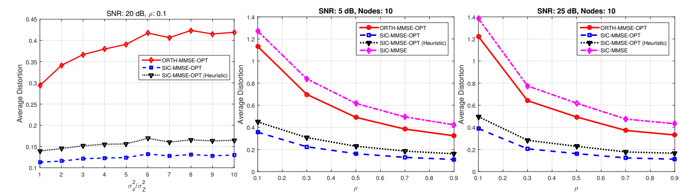
For the distortion minimization problem, we present the minimum distortion for the different schemes. We consider that T=4 slots are available, the number of available sensors vary, while the total power budget is fixed (P=2). In Figure 3, we present results for the average distortion of all the schemes. We observe that the proposed SIC-MMSE-OPT scheme outperforms SIC-MMSE and ORTH-MMSE-OPT for every SNR and ρ combination, since it ensures that SIC and MMSE are jointly optimized. The heuristic has a good performance and the gap is only increased for high values of the correlation coefficient. Recall that with OPT-based schemes, the ratio $\frac{\rho_i}{\sigma_z^2+\sigma_q^2}$ is included in the optimization objective. For the high correlation case, this means that the algorithm attempts to maximize the number of decoded sensors that have *high value* observations. So the heuristic "misses" a number of opportunities for decoding *high value* packets by not scheduling the corresponding sensors. Hence, for the high correlation case (Figure 3(c),(d)), these observations are more valuable for the performance when compared to the low correlation case (Figure 3(a),(b)), leading to an increase in the gap between the heuristic and the optimal solution. More insight into the performance of the heuristic is provided in the next subsection where we discuss power minimization.

Next we compare the performance of SIC-MMSE with ORTH-MMSE-OPT. Note that SIC-MMSE allows simultaneous transmission of many uncoordinated sensors and then applies our joint SIC-MMSE decoding algorithm without optimization. In the low ρ regime (Figure 3(a),(b)), we observe that SIC-MMSE performs almost as good as ORTH-MMSE-OPT. However, this picture changes significantly for higher ρ. For ρ=0.9 (Figure 3(c),(d)) the performance gap between ORTH-MMSE-OPT and SIC-MMSE increases, since each observation now has higher value and their optimal selection is more critical. In other words, the SIC-MMSE scheme that does not perform any optimization, but simply decodes opportunistically the observations with SIC, can reach the same performance with ORTH-MMSE-OPT for low ρ. Finally, we notice that all the schemes are more dependent on ρ rather than the transmit power expressed through the SNR.

In Figure 4(a) we present average distortion results for different $\sigma_\theta^2/\sigma_z^2$ ratios (variance of the signal to be estimated relative to the power of the sampling noise). We observe that as this ratio increases, the performance gap between the proposed scheme (SIC-MMSE-OPT) and the two other schemes increases significantly. As the signal variance is increased (the signal of interest becomes more random), the estimation accuracy suffers with ORTH-MMSE-OPT. The proposed scheme is less sensitive to σ_θ^2. We also investigated the effect of the correlation coefficient on the average distortion for N=10 sensors and we present the results in Figure 4(b,c) for different SNR levels. Here, we observe that for low data correlation, the impact of our scheme becomes more significant. But even for high correlation coefficient, the proposed scheme offers significant benefits, reducing the distortion by more than 50%.

B. Results for Power Minimization

For the power minimization, we set the MSE constraint to 0.5, and present the results in Figure 5. SIC-MMSE-OPT outperforms the other schemes for every N by ensuring



(a) (b) (c)

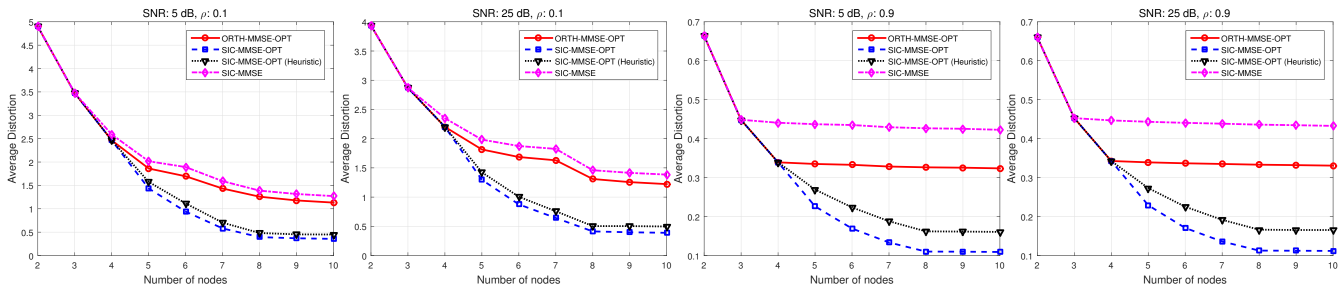
Fig. 4. The effect of the ratio $\sigma_\theta^2/\sigma_z^2$ (with $\sigma_z^2 = 1$) and ρ on average distortion.

a transmit power allocation to specific sensors so that the signal is decodable and no extra power is wasted. The SIC-MMSE-OPT power optimization algorithm identifies the set of sensors that require the minimum power for meeting the MSE constraint. For higher ρ (Figure 5(c),(d)), the value of each observation is higher since the observations across sensors are highly correlated, however, fewer packet decoding events are required. Hence, the target MSE can be achieved with few sensor observations. This also means that sensors that are physically closer the FC, hence, requiring low transmission power, are scheduled. So the heuristic works very well in this case. However, when ρ is low (Figure 5(a),(b)), more observations are needed, and more power must be spent to meet the MSE target. So "missed" scheduling opportunities from the heuristic will result in a higher performance gap from the optimal case. This result highlights the importance of using ρ in the optimization objective and the heuristic. Also note that for σ_θ^2=10, and ρ=0.9 the optimization under ORTH-MMSE fails as illustrated in Figure 5(d), i.e., there is no solution that can meet the MSE requirement.

In Figure 6, we present the normalized *average* power per sensor for the case of simultaneously transmitting sensors under SIC-MMSE-OPT. As the signal variance increases in Figure 6(a), only then the number of simultaneously scheduled sensors increases. For σ_θ^2=1, we observe that the average number of sensors used is two. Here, the small value of the power for the third sensor means that only in a few channel realizations a third sensor was scheduled simultaneously. The number of sensors used on average increases to three for σ_θ^2=10 and four for σ_θ^2=100. For higher transmit SNR illustrated in Figure 6(b), we notice that three sensors are used nearly all the time since they all have a good channel. On the other hand for low transmit SNR most of the time one sensor is used since it requires the consumption of higher power.

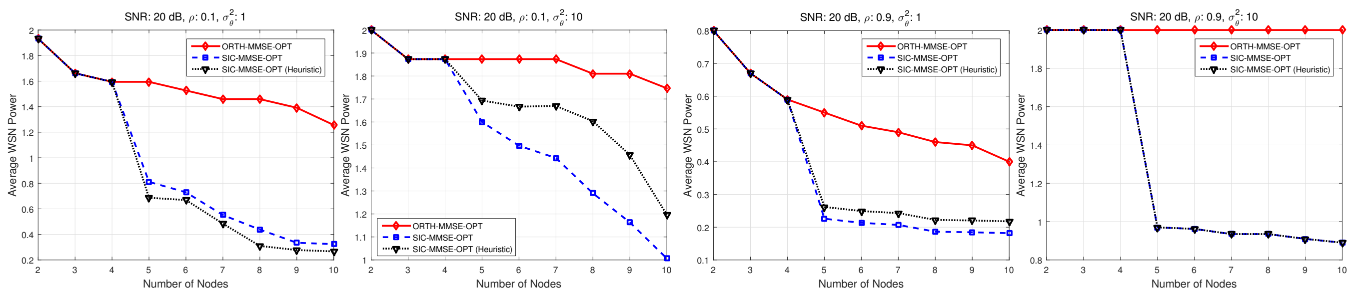
IX. Conclusions

In this paper, we studied DES of a noise-corrupted random parameter in WSN where the sensors are allowed to interfere their transmissions. We first proposed a joint SIC decoder and MMSE estimator for interfering transmissions of correlated data. Next, we modeled the MSE performance of this system and our analysis was compactly captured in a MILP optimization model. Our optimization exploits SIC by selecting sensors for simultaneous transmission so that the interfering signals are decodable while their contribution to the



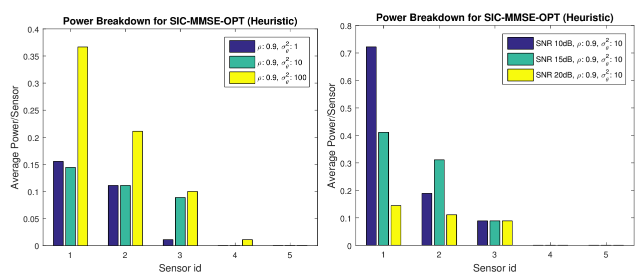
(a) (b) (c) (d)

Fig. 3. The effect of the number of sensors on the distortion of all the proposed schemes for σ_θ^2=10.



(a) (b) (c) (d)

Fig. 5. Results for power consumption minimization.



(a) For different σ_θ^2 values (b) For different SNR values

Fig. 6. Average power allocated to 10 sensors.

MSE reduction is maximized. We also proposed a heuristic that closely follows the optimal solution. The proposed framework offers faster rate of power or MSE reduction as the sensor population is increased. As our future work, we plan to study the performance of the system under average conditions in order to be able to adapt even in longer timescales. This will allow the inclusion of higher number of sensor populations at an even lower computational cost for the proposed scheme.

References

[1] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 54, no. 2, pp. 413–422, 2006.

[2] I. Bahceci and A. Khandani, "Linear estimation of correlated data in wireless sensor networks with optimum power allocation and analog modulation," *Communications, IEEE Transactions on*, vol. 56, no. 7, pp. 1146 –1156, july 2008.

[3] J. Matamoros and C. Antón-Haro, "Opportunistic power allocation and sensor selection schemes for wireless sensor networks," *Trans. Wireless. Comm.*, vol. 9, no. 2, pp. 534–539, February 2010.

[4] A. Behbahani, A. Eltawil, and H. Jafarkhani, "Linear decentralized estimation of correlated data for power-constrained wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 6003 –6016, nov. 2012.

[5] P. Ishwar, R. Puri, K. Ramchandran, and S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 765 – 775, april 2005.

[6] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part i: Gaussian case," *IEEE Trans. on Sig. Proc.*, vol. 54, no. 3, pp. 1131–1143, Oct 2006.

[7] J. Li and G. AlRegib, "Rate-constrained distributed estimation in wireless sensor networks," in *ICCCN*, 2006, pp. 317–322.

[8] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[10] M. Sha, G. Xing, G. Zhou, S. Liu, and X. Wang, "C-mac: Model-driven concurrent medium access control for wireless sensor networks," in *INFOCOM 2009, IEEE*, 2009, pp. 1845–1853.

[11] A. Agrawal *et al.*, "Iterative power control for imperfect successive interference cancellation," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 3, pp. 878–884, 2005.

[12] S. Sen, N. Santhapuri, R. Choudhury, and S. Nelakuditi, "Successive interference cancellation: Carving out mac layer opportunities," *Mobile Computing, IEEE Transactions on*, vol. 12, no. 2, pp. 346–357, 2013.

[13] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *Signal Processing, IEEE Transactions on*, vol. 54, no. 2, pp. 613–626, Feb 2006.

[14] K. Liu and A. Sayeed, "Type-based decentralized detection in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1899–1910, May 2007.

[15] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 822–833, March 2007.

[16] G. Caire and U. Mitra, "Structured multiuser channel estimation for block-synchronous ds/cdma," *Communications, IEEE Transactions on*, vol. 49, no. 9, pp. 1605–1617, Sep 2001.

[17] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.

[18] A. Argyriou and A. Pandharipande, "Collision recovery in distributed wireless networks with opportunistic cooperation," *IEEE Communications Letters*, vol. 14, no. 4, pp. 300–302, April 2010.

[19] A. Argyriou, "Link scheduling for multiple multicast sessions in distributed wireless networks," *Wireless Communications Letters, IEEE*, vol. 2, no. 3, pp. 343–346, 2013.

[20] J. Andrews and T. Meng, "Optimum power control for successive interference cancellation with imperfect channel estimation," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 2, pp. 375–383, 2003.