Robust Decentralized and Distributed Estimation of a Correlated Parameter Vector in MIMO-OFDM Wireless Sensor Networks

Kunwar Pritiraj Rajput, *Graduate Student Member, IEEE,* Mohammad Faisal Ahmed, *Graduate Student Member, IEEE,* Naveen K. D. Venkategowda, *Member, IEEE,* Aditya K. Jagannatham, *Member, IEEE,* Govind Sharma, *Member, IEEE,* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—An optimal precoder design is conceived for the decentralized estimation of an unknown spatially as well as temporally correlated parameter vector in a multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) based wireless sensor network (WSN). Furthermore, exploiting the temporal correlation present in the parameter vector, a rate-distortion theory based framework is developed for the optimal quantization of the sensor observations so that the resultant distortion is minimized for a given bit-budget. Subsequently, optimal precoders are also developed that minimize the sum-MSE (SMSE) for the scenario of transmitting quantized observations. In order to reduce the computational complexity of the decentralized framework, distributed precoder design algorithms are also developed which design precoders using the consensus based alternating direction method of multipliers (ADMM), wherein each SN determines its precoders without any central coordination by the fusion center. Finally, new robust MIMO precoder designs are proposed for practical scenarios operating in the face of channel state information (CSI) uncertainty. Our simulation results demonstrate the improved performance of the proposed schemes and corroborate our analytical formulations.**

***Index Terms*—Decentralized estimation, wireless sensor network (WSN), orthogonal frequency division multiplexing (OFDM), multiple access channel (MAC), alternating direction method of multipliers (ADMM), quantization, rate-distortion theory.**

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). A. K. Jagannatham would like to acknowledge that his work was supported in part by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, in part by the Space Technology Cell, IIT Kanpur, in part by the IIMA IDEA Telecom Centre of Excellence, in part by the Qualcomm Innovation Fellowship, and in part by the Arun Kumar Chair Professorship.

K. P. Rajput, A. K. Jagannatham and G. Sharma are with the Department of Electrical Engineering, Indian Institute of Technology, Kanpur, Kanpur, 208016, India (e-mail: pratiraj@iitk.ac.in, adityaj@iitk.ac.in, govind@iitk.ac.in.)

M. F. Ahmed is with Cisco Systems India Pvt. Ltd., Bengaluru, India (e-mail: mdfaisal165@gmail.com)

N. K. D. Venkategowda is with the Department of Science and Technology, Linköping University, 60174 Norrköping, Sweden (e-mail: naveen.venkategowda@liu.se.)

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk)

I. Introduction

Wireless sensor networks (WSNs), in which small battery powered sensor nodes (SNs) are deployed across a large geographical area, have facilitated a variety of applications such as environmental monitoring, smart healthcare, surveillance and others. In such a WSN, the SNs typically collect observations to monitor an event/ phenomenon of interest, followed by their transmission over wireless links to a fusion center for further processing. Since the SNs in a WSN are both power-and bandwidth-constrained, it is essential to design schemes that efficiently pre-process the SN observations prior to transmission in order to enhance the accuracy of parameter estimation. The sensors are typically miniature and have relatively modest computational capabilities. Furthermore, they can simultaneously sense multiple parameters that can be transmitted over a MIMO channel [1]. For instance, in a typical environmental monitoring application, the different elements of the parameter vector can be pressure, temperature and moisture etc. Further, another significant challenge is the frequency selective fading channel between each sensor and the fusion center, which results in inter symbol interference (ISI). To overcome this, one can employ OFDM. Therefore, MIMO-aided OFDM enables the sensing and simultaneous transmission of multiple parameters over a wideband channel. Such a system is well-suited for application in low-mobility scenarios. A brief review of the state-of-the-art along with the portrayal of the salient solutions is presented next.

A. Related Contributions on Spatial and Temporal Correlation

Starting with the treatises that have proposed schemes to exploit the temporal correlation for improved parameter estimation, [2] considers the problem of estimation of a temporally correlated parameter and derives the optimal sensor collaboration strategy. Their scheme requires inter-sensor communication and optimization to be performed prior to transmission, which leads to a higher computational overhead. Dong *et. al.* [3] have derived both offline and online strategies for optimal power allocation that minimize the weighted sum distortion arising in the estimation of two temporally correlated Gaussian sources in a WSN having a pair of energy harvesting SNs and a fusion center. Li and Parker [4] have proposed a novel method based on exploiting the spatio-temporal correlation for the estimation of missing observations in WSNs. Das *et*

TABLE I: A comparative summary of contributions of the salient existing works

Feature	[2]	[12]	[14]	[16]	[18]	[19]	[20]	[21]	[22]	[25]	Proposed
Spatial correlation	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Temporal correlation	✓	×	×	×	×	×	×	×	×	×	✓
Analog observations	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Quantized observations	×	✓	✓	✓	×	×	×	×	×	×	✓
Stochastic CSI uncertainty	×	×	×	×	×	×	×	✓	×	×	✓
Norm ball CSI uncertainty	×	×	×	×	✓	✓	✓	×	×	×	✓
MIMO-OFDM sensor network	×	×	×	×	×	×	×	×	×	×	✓
Coherent MAC	✓	×	✓	×	✓	✓	✓	✓	✓	✓	✓
Vector estimation	×	×	×	✓	✓	×	✓	✓	✓	✓	✓
Decentralized estimation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Distributed estimation	✓	×	×	×	×	×	✓	×	×	×	✓
Frequency selective channel	×	×	×	×	×	×	×	×	×	×	✓
Total power constraint	×	✓	×	✓	×	✓	×	✓	×	✓	✓
Individual power constraints	✓	×	×	×	✓	×	✓	×	✓	✓	✓

al. [5] have proposed a scheme for wireless sensor networks wherein the minimum number of event monitoring SNs is found in the face of spatial correlation. Later, exploiting the temporal correlation present in the parameter of interest, each SN adjusts its sleep cycle for effectively monitoring the phenomenon of interest, while minimizing the energy consumption. Recently, an environmental monitoring scheme was developed by Ko *et al.* [6], that also exploits the spatio-temporal correlation for optimal transmission in an internet of things (IoT)-based wireless powered sensor network. The SNs therein exploit the spatio-temporal correlation and energy level to determine whether or not to transmit the observed data to the IoT gateway. Özçelikkale *et al.* [7] proposed a scheme for the remote estimation of a temporally correlated field using an energy harvesting SN. The optimal energy and data buffer sizes derived therein depend on the degrees of freedom of the signal. As a further development, various schemes have also been proposed for quantizing the observations of the SNs in a WSN, while achieving different objectives. A brief review is presented next.

B. Related Works on Quantization of SN Observations

Sun and Goyal [8] have proposed a scheme for distributed optimal scalar quantizer design when inter-sensor communication is performed prior to transmitting the observations to the fusion center. Their theoretical as well as practical analysis shows that the distortion at the fusion center can be minimized significantly with inter-sensor communication even when the sensor communicates at a very low rate. Nevat *et al.* [9] have proposed spatial field resonstruction schemes based on quanitized SN observations with power and bandwidth constraints. Msechu and Giannakis [10] have proposed a novel data reduction technique based on interval censoring followed by quantization of SN observations for deterministic as well as random parameter estimation. Li and Al-Regib in [11] have proposed an optimal scheme, which strikes a tradeoff between the number of active SNs and the resultant bit-rate of each active SN, to minimize the estimation MSE at the fusion center. Each active SN's observation is quantized using a single bit, thus making it well suited for scenarios having stringent bandwidth constraints. A block coordinate descent based iterative algorithm has been proposed by Chaudhary and Vandendorpe [12] for joint quantization of the observations and SN power allocation, toward minimizing the reconstruction error at the fusion center. An adaptive quantization scheme has been conceived in [13], wherein each SN transmits only 1-bit quantized observations after comparing the observed value to a quantization threshold. The quantization threshold in their work is chosen adaptively based on the previous transmissions of other SNs. A different 1-bit quantization scheme is proposed by Ribeiro and Giannakis [14] for the maximum likelihood estimation of a deterministic parameter. A non-uniform quantization and power allocation scheme has been proposed by Zhou *et al.* [15] for energy efficient transmission in a WSN. Sani and Vosoughi [16] have proposed a joint power allocation and quantization scheme such that the resulting distortion is minimized at the fusion center. However, they do not exploit any form of correlation in their scheme and also use an orthogonal MAC that is bandwidth inefficient. Leinonen *et al.* [17] conceived a low complexity quantized compressive sensing algorithm for the estimation of a correlated sparse source and derived the optimality bound for the scenario with two sensors and a decoder. Naturally, due to limited feedback, quantization error etc., perfect CSI is never available either at the fusion center or the SNs in a practical WSN. Hence, some works have also proposed schemes for robust parameter estimation in the presence of imperfect CSI in WSNs, which are reviewed next.

C. Related Treatises on Robust Designs for Imperfect CSI

Zhu *et al.* [18] have proposed robust precoder and linear combiner designs for the estimation of a scalar parameter considering the bounded CSI uncertainty model. Venkategowda *et al.* [19] have proposed robust precoder designs for the estimation of a scalar parameter considering the ellipsoidal and norm ball CSI uncertainty models, using the minimum variance distortionless precoding (MVDP) framework. Liu *et al.* [20] have proposed robust centralized and distributed schemes for scalar parameter estimation with ellipsoidal CSI uncertainty. Furthermore, a robust precoder design has been proposed by Rostami and Falahati [21] for the estimation of a vector parameter considering the stochastic CSI uncertainty model. However, [20] and [21] neither consider the temporal correlation or quantization of observations and optimal bit allocation subject to a given bit budget.

D. Research Gap and Motivation

Since temporal correlation results in multiple closely related observations of a single parameter, this can naturally be exploited for improved estimation performance. Furthermore, quantization is of key importance in modern digital wireless systems, since it limits the transmission to a few bits over each subcarrier, thus leading to improved spectral efficiency. However, quantization also leads to quantization error, which has to be accounted for optimal transmission. This has been accomplished in this paper by leveraging the results from rate-distortion theory and optimal bit allocation. Finally, since the CSI is typically estimated using a finite number of pilot symbols in a practical wireless system, the channel estimation error leads to CSI uncertainty, which impacts the precoder design and resulting performance. Therefore, the robust precoder/ combiner design framework that takes the CSI uncertainty into account is key toward obtaining enhanced parameter estimates. To the best of our knowledge, the existing works have not considered the problem of estimation of a temporally correlated parameter incorporating also the quantization and CSI uncertainty effects. Therefore, various schemes are proposed in this work to fill this void. A brief summary of the contributions of this paper is presented next.

E. Our Contributions

Two different estimation frameworks, based on decentralized and distributed approaches, have been proposed for the estimation of a correlated vector parameter in a MIMO-OFDM WSN. In the decentralized setting, the various computational tasks are carried out by the fusion center and the individual precoder matrices are subsequently fed back to each SN. This requires that each SN provides the information about its observation matrix, observation noise statistics etc. to the fusion center. The centralized scenario refers to the setting, where the observations corresponding to all the sensors are directly available at the fusion center without any degradation. By contrast, in the distributed setting, the sensors design their respective precoders relying on their mutual exchange of messages. This significantly reduces the computational complexity and communication overhead required for precoder design. The various contributions of the proposed work are as follows.

- First, optimal precoders are developed for the scenario of analog sensor observation transmission over a coherent MAC, for minimizing the sum-MSE (SMSE) at the fusion center assuming the availability of perfect CSI.
- Subsequently, exploiting the temporal correlation present in the parameter vector, a rate-distortion theory based framework is developed for optimal bit-allocation to quantize the SN's observations to minimize the sum distortion corresponding to all the subcarriers, given a total bit-rate constraint for each SN. The optimal precoders are also determined for this scenario with the aid of quantized observations, once again, considering the availability of perfect CSI.
- In order to further reduce the computational complexity and communication overheads, dual consensus ADMM-based distributed schemes are also proposed for precoder design relying on both analog as well as quantized observations.
- Robust precoder designs are next proposed for practical scenarios under CSI uncertainty, once again for scenarios with/without quantization.
- Simulation results demonstrate that the proposed designs approach the corresponding MSE benchmark at high SNR, and are also resilient to the degradation arising due to CSI uncertainty.

The remainder of the paper is organized as follows. Section II describes the scheme proposed for spatio-temporally correlated parameter estimation. Section III extends the same to a scenario of quantized observations, while Section IV presents the distributed versions of both the schemes, based on ADMM. Section V presents robust precoder designs in the face of CSI uncertainty, first for scenarios with analog observations, followed by scenarios with quantized observations. Simulation results are provided in Section VI to illustrate the attainable performance, followed by our conclusions in Section VIII.

Notation: Bold lower case letters $\tilde{\mathbf{x}}$ and $\mathbf{x}$ are used to represent the time and frequency domain vectors, respectively. Bold capital letters $\tilde{\mathbf{X}}$ and $\mathbf{X}$ are used to represent time and frequency domain matrices, respectively. Furthermore, $(.)^T$ and $(.)^H$ stand for transposition and Hermitian transposition, respectively. $\text{Tr}(\mathbf{X})$ denotes the trace of a matrix $\mathbf{X}$. $\mathbf{X} = \text{diag}(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_L)$ denotes a block diagonal matrix $\mathbf{X}$ that has matrices $\mathbf{X}_i$'s, $1 \leq i \leq L$ on its principal diagonal. $\mathbb{E}\{.\}$ denotes the expectation operator and $||\mathbf{x}||$ denotes the Euclidean norm of vector $\mathbf{x}$. For a complex quantity x, $\mathcal{R}(x)$ denotes its real part. The operator $\text{vec}(\mathbf{X})$ creates a column vector of size mn obtained by stacking the column vectors of a matrix $\mathbf{X}$. The operator $\text{vec}_m^{-1}(\mathbf{x})$ rearranges the vector $\mathbf{x} \in \mathbb{C}^{mn\times 1}$ into a matrix with m rows and n columns. The symbol $\otimes$ denotes the matrix Kronecker product. $\mathbf{I}_N$ denotes a $N \times N$ identity matrix.

II. WSN System Model for Parameter Sensing and Estimation

A typical WSN deployment comprises of a large number of sensors that sense/monitor an event or multiple events of interest and subsequently transmit their observations to the fusion center for final estimation of the underlying quantity of interest. Since the wireless channel is fading in nature, for accurate estimation, it is essential to design the precoders at the sensors considering also the power and bandwidth constraints at the sensors. This is the goal of the decentralized estimation schemes proposed in this section. Consider a WSN having L SNs, with the observation vector $\widetilde{\mathbf{x}}_l(n) \in \mathbb{C}^{q\times 1}$ of the lth SN at time instant n modeled as

$$\widetilde{\mathbf{x}}_l(n) = \mathbf{A}_l\tilde{\boldsymbol{\theta}}(n) + \widetilde{\mathbf{v}}_l(n), \tag{1}$$

where $\mathbf{A}_l \in \mathbb{C}^{q\times K}$ represents the observation matrix for the lth SN and $\widetilde{\mathbf{v}}_l(n) \in \mathbb{C}^{q\times 1}$ denotes the corresponding observation noise that is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_l)$. The quantity $\tilde{\boldsymbol{\theta}}(n) \in \mathbb{C}^{K\times 1}$ denotes the unknown temporally correlated parameter vector at time instant n, which is to be estimated.

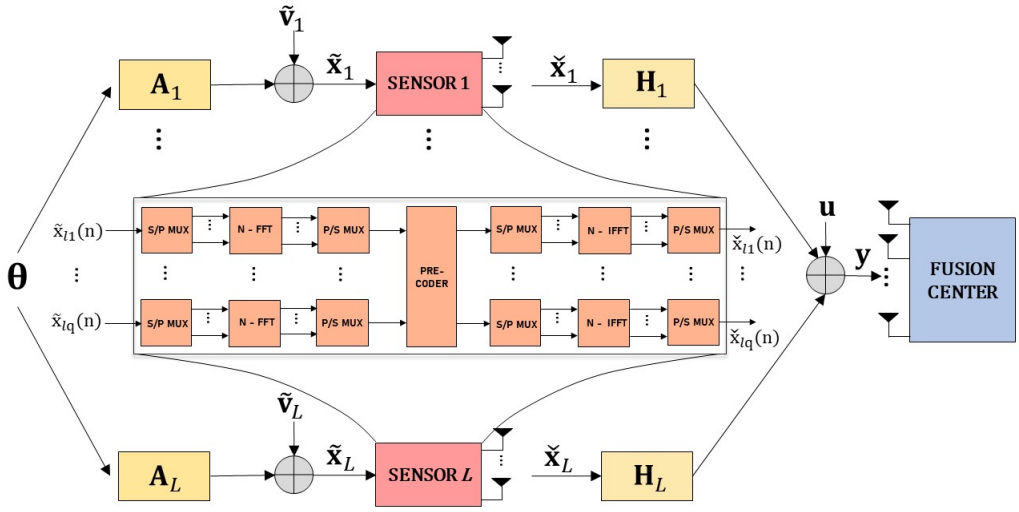


Fig. 1: System model for vector parameter estimation in an OFDM based WSN

It can be readily observed that since each sensor l observes the same parameter vector $\tilde{\boldsymbol{\theta}}$ modulated by the observation matrix, the spatial correlation across sensors is captured by the observation matrices $\mathbf{A}_l$, $1 \leq l \leq L$, as also noted in [1], [22], [23]. The kth element of $\tilde{\boldsymbol{\theta}}(n)$, denoted by $\tilde{\theta}_k(n)$, $1 \leq k \leq K$, is a sample of a zero-mean temporally correlated wide sense stationary (WSS) Gaussian random process that has the power spectral density of $S_k(f)$. Concatenating the observations $\widetilde{\mathbf{x}}_l(n)$ over N time instants at the lth SN, one obtains the model

$$\underbrace{[\widetilde{\mathbf{x}}_l(0), \widetilde{\mathbf{x}}_l(1) \ldots, \widetilde{\mathbf{x}}_l(N-1)]}_{\widetilde{\mathbf{X}}_l} = \mathbf{A}_l \underbrace{[\tilde{\boldsymbol{\theta}}(0), \tilde{\boldsymbol{\theta}}(1), \ldots, \tilde{\boldsymbol{\theta}}(N-1)]}_{\widetilde{\boldsymbol{\Theta}}} + \underbrace{[\widetilde{\mathbf{v}}_l(0), \widetilde{\mathbf{v}}_l(1), \ldots, \widetilde{\mathbf{v}}_l(N-1)]}_{\widetilde{\mathbf{V}}_l}. \tag{2}$$

Performing the N-point row-wise fast Fourier transform (FFT) of the observation matrix $\widetilde{\mathbf{X}}_l$ yields

$$\mathbf{X}_l = \widetilde{\mathbf{X}}_l \boldsymbol{\Phi} = \mathbf{A}_l \widetilde{\boldsymbol{\Theta}} \boldsymbol{\Phi} + \widetilde{\mathbf{V}}_l \boldsymbol{\Phi} = \mathbf{A}_l \boldsymbol{\Theta} + \mathbf{V}_l, \tag{3}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{N \times N}$ denotes the FFT matrix with its (r,t)th element equal to $\frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}rt}$. The above model after the FFT operation at the lth SN be expanded as

$$[\mathbf{x}_l(0), \mathbf{x}_l(1), \ldots, \mathbf{x}_l(N-1)] = \mathbf{A}_l [\boldsymbol{\theta}(0), \boldsymbol{\theta}(1), \ldots, \boldsymbol{\theta}(N-1)] + [\mathbf{v}_l(0), \mathbf{v}_l(1), \ldots, \mathbf{v}_l(N-1)], \tag{4}$$

where $\mathbf{x}_l(m)$, $0 \leq m \leq N-1$, represents the frequency domain observation vector corresponding to the mth subcarrier at the lth SN. The corresponding model can be extracted from the above equation as

$$\mathbf{x}_l(m) = \mathbf{A}_l \boldsymbol{\theta}(m) + \mathbf{v}_l(m). \tag{5}$$

The quantities $\boldsymbol{\theta}(m)$ and $\mathbf{v}_l(m)$ denote the corresponding frequency domain components of the unknown parameter vector and observation noise, where both have zero mean and their covariance matrices are defined as $\mathbf{R}_\theta(m) \in \mathbb{C}^{K \times K}$ and $\mathbf{R}_{v,l}(m) \in \mathbb{C}^{q \times q}$, respectively. As given in [24, Sec. 2.4], for large N, the kth component of $\boldsymbol{\theta}(m)$, denoted as $\theta_k(m)$, is a zero-mean random variable with variance [1]

$$\mathbb{E}\left[|\theta_k(m)|^2\right] = S_k(f_m), \tag{6}$$

where $f_m = \frac{m}{N}$, $0 \leq m \leq N-1$.

Let $\mathbf{F}_l(m) \in \mathbb{C}^{N_t \times q}$, where N_t denotes the number of transmit antennas at each SN, represent the precoder corresponding to subcarrier m at the lth SN. Therefore, the precoded outputs $\check{\mathbf{x}}_l(m)$ at the lth SN are obtained as

$$\check{\mathbf{x}}_l(m) = \mathbf{F}_l(m)\mathbf{x}_l(m) = \mathbf{F}_l(m)\mathbf{A}_l\boldsymbol{\theta}(m) + \mathbf{F}_l(m)\mathbf{v}_l(m), \tag{7}$$

which are subsequently loaded on subcarriers $m = 0, 1, \ldots, N-1$, followed by transmission to the fusion center over a coherent MAC in the MIMO-OFDM WSN. Post the FFT operation at the receiver, the output vector $\mathbf{y}(m) \in \mathbb{C}^{N_r \times 1}$, for subcarrier m, is given as

$$\mathbf{y}(m) = \sum_{i=l}^{L} \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{A}_l\boldsymbol{\theta}(m) + \sum_{l=1}^{L} \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{v}_l(m) + \mathbf{u}(m), \tag{8}$$

where $\mathbf{H}_l(m) \in \mathbb{C}^{N_r \times N_t}$ denotes the MIMO channel between the lth SN and the fusion center for the mth subcarrier, $\mathbf{u}(m) \in \mathbb{C}^{N_r \times 1}$ represents the receiver noise that has zero mean with its covariance matrix defined as $\mathbf{R}_u(m) \in \mathbb{C}^{N_r \times N_r}$

[1]Let $x(n)$ denote a sample of a zero-mean temporally correlated WSS Gaussian random process at time instant n. Consider now a block comprising N such samples from 0 to $N-1$. Performing the N-point fast Fourier transform (FFT) results in the N-point frequency domain samples $X(m)$, where $0 \leq m \leq N-1$. Then it follows from Section 2.4 of [24], for large N, the mth frequency domain sample $X(m)$ is distributed as $\mathcal{CN}(0, S_x(f_m))$, where $f_m = \frac{m}{N}$ and $S_x(f_m)$ denotes the power spectral density (PSD) of x at frequency f_m. Thus, the time domain correlation of the random observation process is exploited using the PSD in the frequency domain.

and N_r denotes the number of antennas at the fusion center. Motivated by the MVDP framework [25], the output vector $\mathbf{y}(m)$ in (8) will be a distortionless estimate of the parameter vector $\boldsymbol{\theta}(m)$ at the fusion center under the condition

$$\sum_{l=1}^{L} \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{A}_l = \mathbf{I}_K. \tag{9}$$

The distortionless criterion above eliminates inter parameter interference and is similar to the constant gain condition employed in a standard Capon beamformer [26]. Furthermore, it also renders the optimization tractable since the joint optimization of the combiner and precoder is non-convex.

The optimal precoders designs for the estimation of the parameter $\boldsymbol{\theta}(m)$ at the fusion center for various scenarios are determined in the subsequent sections.

III. Parameter Estimation with Analog Observations

Considering the transmission of analog observations, the corresponding MSE of the estimate $\widehat{\boldsymbol{\theta}}(m)$ for the mth subcarrier is easily deduced as shown in (10). Taking the sum across all the N subcarriers, the SMSE becomes

$$\begin{aligned}\text{SMSE} &= \sum_{m=0}^{N-1} \text{Tr}\left[\sum_{l=1}^{L} \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)\right] \\ &\quad + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right], \end{aligned} \tag{11}$$

using the properties $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$ and $\text{Tr}\left(\mathbf{A}^H\mathbf{BCD}\right) = \text{vec}(\mathbf{A})^H\left(\mathbf{D}^T \otimes \mathbf{B}\right)\text{vec}(\mathbf{C})$, this can be further simplified to the compact form shown below

$$\begin{aligned}\text{SMSE} &= \sum_{l=1}^{L}\sum_{m=0}^{N-1} \mathbf{f}_l^H(m)\mathbf{Q}_l(m)\mathbf{f}_l(m) + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right] \\ &= \sum_{l=1}^{L} \mathbf{f}_l^H\mathbf{Q}_i\mathbf{f}_l + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right], \end{aligned} \tag{12}$$

where $\mathbf{f}_l(m) = \text{vec}\left[\mathbf{F}_l(m)\right] \in \mathbb{C}^{N_tq\times q}$ and $\mathbf{Q}_l(m) = \left[\mathbf{R}_{v,l}(m) \otimes \mathbf{H}_l^H(m)\mathbf{H}_l(m)\right] \in \mathbb{C}^{N_tq\times N_tq}$. Furthermore, $\mathbf{f}_l \in \mathbb{C}^{NN_tq\times 1}$ is obtained by stacking the vectors $\mathbf{f}_l(m)$ for all the subcarriers $m = 0, 1, .., N-1$, and $\mathbf{Q}_l \in \mathbb{C}^{NN_tq\times NN_tq}$ is a block-diagonal matrix with $\mathbf{Q}_l(m)$ as its block diagonal elements. The above expression of the SMSE can be further streamlined to the form

$$\text{SMSE} = \mathbf{f}^H\mathbf{Q}\mathbf{f} + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right], \tag{13}$$

where $\mathbf{f} \in \mathbb{C}^{LNN_tq\times 1}$ denotes a similar stacking of $\mathbf{f}_l$ over all the L SNs and $\mathbf{Q} = \text{diag}\Big[\mathbf{Q}_1, \mathbf{Q}_2, \cdots, \mathbf{Q}_L\Big] \in \mathbb{C}^{LNN_tq\times LNN_tq}$. The distortionless constraint in (9) can be simplified by applying the vec operator to both sides and exploiting the property $\text{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\text{vec}(\mathbf{B})$, for any compatible matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, as $\sum_{l=1}^{L} \mathbf{W}_l(m)\mathbf{f}_l(m) = \text{vec}\left[\mathbf{I}_K\right]$, where $\mathbf{W}_l(m) = \left[\mathbf{A}_l^T \otimes \mathbf{H}_l(m)\right] \in \mathbb{C}^{N_rK\times N_tq}$. Hence, the aggregated distortionless constraint across all the N subcarriers can be represented as

$$\mathbf{W}\mathbf{f} = \mathbf{g}, \tag{14}$$

where $\mathbf{W}_l \in \mathbb{C}^{NN_rK\times NN_tq}$ and in turn $\mathbf{W} \in \mathbb{C}^{NN_rK\times LNN_tq}$ are defined as $\mathbf{W}_l = \text{diag}\Big[\mathbf{W}_l(0), \mathbf{W}_l(1), \ldots, \mathbf{W}_l(N-1)\Big]$, $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_L]$ and $\mathbf{g} = \text{vec}(\mathbf{I}_K) \otimes \mathbf{1}_N \in \mathbb{C}^{NK^2\times 1}$, with $\mathbf{1}_N$ denoting the N- dimensional vector of all ones. The optimization problem minimizing the SMSE across all the SNs in (13) subject to the distortionless constraint in (14), can be succinctly represented as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{Q}\mathbf{f} \\ \text{subject to} \quad & \mathbf{W}\mathbf{f} = \mathbf{g}, \end{aligned} \tag{15}$$

where the constant term corresponding to the noise covariance has been ignored in the objective function of (13), since it does not affect the minimization procedure. Using the Karush-Kuhn-Tucker (KKT) framework [27, Sec. 5.5.3], the closed form solution for the optimal vector $\mathbf{f}^*$ is readily-obtained as

$$\mathbf{f}^* = \mathbf{Q}^{-1}\mathbf{W}^H\left[\mathbf{W}\mathbf{Q}^{-1}\mathbf{W}^H\right]^{-1}\mathbf{g}. \tag{16}$$

The optimal precoding vector $\mathbf{f}_l^*(m)$ corresponding to the mth subcarrier at the lth SN can be obtained from $\mathbf{f}^*$ by extracting the subvector corresponding to the indices $[(l-1)N + (m-1)]\,N_tq+1$ to $[(l-1)N + m]\,N_tq$. Finally, the corresponding optimal precoding matrix $\mathbf{F}_l^*(m)$ can be retrieved using the relation $\text{vec}_{N_t}^{-1}[\mathbf{f}_l^*(m)]$.

A. Parameter Estimation with Transmit Power Constraint

One can also additionally incorporate a total transmit power constraint in the optimization procedure to derive the optimal precoders as follows. From (7), the average transmit power $P_l(m) = \text{Tr}\left[\mathbb{E}\left[\left[\mathbf{F}_l(m)\mathbf{x}_l(m)\right]\left[\mathbf{F}_l(m)\mathbf{x}_l(m)\right]^H\right]\right]$ of the lth SN for the mth subcarrier is given by

$$P_l(m) = \text{Tr}\Big[\mathbf{F}_l^H(m)\boldsymbol{\Gamma}_l(m)\mathbf{F}_l(m)\Big] = \mathbf{f}_l^H(m)\boldsymbol{\Psi}_l(m)\mathbf{f}_l(m), \tag{17}$$

where $\boldsymbol{\Gamma}_l(m) = \left[\mathbf{A}_l\mathbf{R}_\theta(m)\mathbf{A}_l^H + \mathbf{R}_{v,l}(m)\right] \in \mathbb{C}^{q\times q}$ and $\boldsymbol{\Psi}_l(m) = [\boldsymbol{\Gamma}_l(m) \otimes \mathbf{I}_K] \in \mathbb{C}^{qK\times qK}$. Hence, the total transmit power considering all the N subcarriers for all the SNs in the system is readily obtained as

$$\sum_{l=1}^{L}\sum_{m=0}^{N-1} \mathbf{f}_l^H(m)\boldsymbol{\Psi}_l(m)\mathbf{f}_l(m) = \mathbf{f}^H\boldsymbol{\Psi}\mathbf{f}, \tag{18}$$

where $\boldsymbol{\Psi}_l \in \mathbb{C}^{NqK\times NqK}$ is a block diagonal matrix with $\boldsymbol{\Psi}_l(m), 0 \le m \le N-1$ as its block diagonal elements, and $\boldsymbol{\Psi} \in \mathbb{C}^{LNqK\times LNqK}$ in turn contains matrices $\boldsymbol{\Psi}_l, 1 \le l \le L$ on its principal diagonal. Hence, the optimization problem for determining the optimal precoders subject to the total SN transmit power being limited to P_T can be formulated as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{Q}\mathbf{f} \\ \text{subject to} \quad & \mathbf{W}\mathbf{f} = \mathbf{g} \\ & \mathbf{f}^H\boldsymbol{\Psi}\mathbf{f} \le P_T. \end{aligned} \tag{19}$$

$$\text{MSE} = \mathbb{E}\left[\text{Tr}\left[\left[\widehat{\boldsymbol{\theta}}(m)-\boldsymbol{\theta}(m)\right]\left[\widehat{\boldsymbol{\theta}}(m)-\boldsymbol{\theta}(m)\right]^H\right]\right] = \text{Tr}\left[\sum_{l=1}^{L}\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)+\mathbf{R}_u(m)\right].$$
$$= \text{Tr}\left[\sum_{l=1}^{L}\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)\right] + \text{Tr}\left[\mathbf{R}_u(m)\right], \tag{10}$$

Once again, invoking the KKT framework [27, Sec. 5.5.3], the above optimization problem can be solved as shown in the technical report [28, Sec.-III], and the optimal vector $\mathbf{f}^*$ is determined as

$$\mathbf{f}^* = \left[\mathbf{Q}+\nu\boldsymbol{\Psi}\right]^{-1}\mathbf{W}^H\left[\mathbf{W}\left[\mathbf{Q}+\nu\boldsymbol{\Psi}\right]^{-1}\mathbf{W}^H\right]^{-1}\mathbf{g}, \tag{20}$$

where ν denotes the Lagrangian dual variable corresponding to the inequality constraint. One can now extract the vectors $\mathbf{f}_l^*(m)$ and in turn the precoding matrices $\mathbf{F}_l^*(m)$ using the procedure described below (16). One can also modify the above problem to restrain the power of each SN l to P_l by including the constraint $\mathbf{f}_l^H\boldsymbol{\Psi}_l\mathbf{f}_l \leq P_l, \quad 1 \leq l \leq L$. The resulting optimization will be similar to (19) with the total power constraint replaced by L such power constraints and can be solved efficiently using optimization tools such as CVX [29].

B. Centralized MMSE Error Bound

The minimum MSE (MMSE) of a centralized estimator, which serves as a lower bound for the above estimation framework, is obtained next. The best performance is achieved when all the observations across all the subcarriers of each SN are directly available to the fusion center. The centralized MMSE estimate of the parameter thus obtained serves as a valuable benchmark for the performance of the proposed decentralized estimation schemes. Therefore, stacking all the observations collected by all the SNs over all the subcarriers, one obtains

$$\underbrace{\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \vdots\\ \mathbf{x}_L\end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix}\mathbf{I}_N\otimes\mathbf{A}_1\\ \mathbf{I}_N\otimes\mathbf{A}_2\\ \vdots\\ \mathbf{I}_N\otimes\mathbf{A}_L\end{bmatrix}}_{\mathbf{C}}\underbrace{\begin{bmatrix}\boldsymbol{\theta}(0)\\ \boldsymbol{\theta}(1)\\ \vdots\\ \boldsymbol{\theta}(N-1)\end{bmatrix}}_{\boldsymbol{\theta}} + \underbrace{\begin{bmatrix}\mathbf{v}_1\\ \mathbf{v}_2\\ \vdots\\ \mathbf{v}_L\end{bmatrix}}_{\mathbf{v}}. \tag{21}$$

For the MMSE fusion rule, the SMSE bound at the fusion center is obtained as

$$\text{SMSE}_{\text{MMSE}} = \text{Tr}\left[\left[\mathbf{R}_\theta^{-1}+\mathbf{C}^H\mathbf{R}_v^{-1}\mathbf{C}\right]^{-1}\right], \tag{22}$$

where the quantities $\mathbf{R}_\theta \in \mathbb{C}^{NK\times NK}$ and $\mathbf{R}_v \in \mathbb{C}^{LNq\times LNq}$ denote the covariance matrices of the parameter vector $\boldsymbol{\theta}$ and stacked observation noise vector $\mathbf{v}$, respectively, that are defined as $\mathbf{R}_\theta = \text{diag}\left[\mathbf{R}_\theta(0), \mathbf{R}_\theta(1), \ldots, \mathbf{R}_\theta(N-1)\right]$ and $\mathbf{R}_v = \text{diag}\left[\mathbf{R}_{v,1}(0), \mathbf{R}_{v,1}(1), \ldots, \mathbf{R}_{v,L}(N-1)\right]$. While analog observations enhance the accuracy of estimation, their transmission is unrealistic for digital modulation based modern WSNs. In view of this, the next section presents a similar framework for the optimal quantization of SN observations and their pre-processing for digital transmission.

IV. Parameter Estimation with Quantized Observations

A rate-distortion theory based framework is now developed for optimal quantization of the observations. This is followed by the development of the optimization problem toward optimal distribution of the bits across the subcarriers, subject to a bit budget, with the aim of minimizing the net distortion corresponding to parameter estimation. This process is aided by the temporal correlation, leveraging which leads to a lower bit load and SMSE minimization at the fusion center for the given bit budget. To suppress the inter-parameter interference, the ZF estimator is employed at the sensor followed by the quantization of the individual components of the parameter vector $\boldsymbol{\theta}(m)$. Consider the frequency domain observation vector $\mathbf{x}_l(m)$ corresponding to the lth SN on the mth subcarrier as modeled in (5). Prior to quantization, the zero forcing (ZF) estimate of the parameter vector is initially obtained at each SN as follows

$$\widehat{\mathbf{x}}_l(m) = \boldsymbol{\theta}(m) + \left[\mathbf{A}_l^H\mathbf{A}_l\right]^{-1}\mathbf{A}_l^H\mathbf{v}_l(m) = \boldsymbol{\theta}(m) + \mathbf{z}_l(m), \tag{23}$$

where $\mathbf{z}_l(m) \in \mathbb{C}^{q\times 1}$ is the effective noise term. Setting $\mathbf{R}_{v,l}(m) = \mathbf{I}_q$, without loss of generality, it follows that $\mathbf{z}_l(m)$ has a mean of zero and a covariance matrix equal to $\left(\mathbf{A}_l^H\mathbf{A}_l\right)^{-1}$. Let, $\widehat{x}_{l,k}(m)$ denote the kth element of the ZF-estimate of $\widehat{\mathbf{x}}_l(m)$, which can be written as

$$\widehat{x}_{l,k}(m) = \theta_k(m) + z_{l,k}(m). \tag{24}$$

Its variance can be evaluated as

$$\sigma_{l,k}^2(m) = \mathbb{E}\left[|\widehat{x}_{l,k}(m)|^2\right]$$
$$= S_k(f_m) + \mathbb{E}\left[|z_{l,k}(m)|^2\right] = S_k(f_m) + \sigma_{z_{l,k}}^2(m), \tag{25}$$

where $\sigma_{z_{l,k}}^2(m) = \left[\left[\mathbf{A}_l^H\mathbf{A}_l\right]^{-1}\right]_{kk}$. Based on rate-distortion theory [30, Th. 10.3.2] for the quantization of Gaussian samples[2], the minimum achievable distortion $D_{l,k}(m)$, using $b_{l,k}(m)$ bits for the quantization of $\widehat{x}_{l,k}(m)$, is given by

$$D_{l,k}(m) = \sigma_{l,k}^2(m)2^{-2b_{l,k}(m)} = \left[S_k(f_m) + \sigma_{z_{l,k}}^2(m)\right]2^{-2b_{l,k}(m)}. \tag{26}$$

Therefore, the sum distortion resulting from the quantization of the kth element of the vector $\widehat{x}_{l,k}(m)$ for all the subcarriers at the lth SN can be expressed as

$$D_{l,k} = \sum_{m=0}^{N-1} D_{l,k}(m) = \sum_{m=0}^{N-1}\sigma_{l,k}^2(m)2^{-2b_{l,k}(m)}. \tag{27}$$

[2]It follows from [24, Sec. 2.4], for large N, the elements of the frequency domain parameter vector $\boldsymbol{\theta}$ are uncorrelated. Furthermore, since they are Gaussian, it follows that they are independent.

Hence, the optimal bit allocation for the kth element of the vector $\widehat{\mathbf{x}}_l$ at each SN l, which minimizes the sum-distortion, subject to a bit-budget B, can be determined as the solution of the optimization problem

$$\begin{aligned} \underset{b_{l,k}(m)}{\text{minimize}} \quad & D_{l,k} = \sum_{m=0}^{N-1} \sigma_{l,k}^2(m)(4)^{-b_{l,k}(m)} \\ \text{subject to} \quad & \sum_{m=0}^{N-1} b_{l,k}(m) \leq B. \end{aligned} \tag{28}$$

Using the KKT framework [27, Sec. 5.5.3], as shown in technical report [28, Sec.-IV], the closed form expression for the optimal number of bits $b_{l,k}^*(m)$ at the lth SN is obtained as

$$\begin{aligned} b_{l,k}^*(m) = \frac{B}{N} - \ln(4)\Bigg[\Bigg[\frac{1}{N}\sum_{m=0}^{N-1} \ln\left[\sigma_{l,k}^2(m)\ln(4)\right]\Bigg] \\ - \ln\left[\sigma_{l,k}^2(m)\ln(4)\right]\Bigg]. \end{aligned} \tag{29}$$

The procedure to round-off $b_{l,k}^*(m)$ to an integer value, while constraining the total number of bits to B, is described in Algorithm 1.

Algorithm 1 Procedure to round off the optimal bit values $b_l^*(m)$ obtained in (29)

1: **Input** $\mathbf{b}_l^* = [b_{l,k}^*(0), b_{l,k}^*(1), \cdots, b_{l,k}^*(N-1)]^T$, B
2: **Initialize** $\widetilde{\mathbf{b}}_l = \mathbf{0} \in \mathbb{Z}_+^{N\times 1}$ and sum = 0.
3: **while** sum $<$ B **do**
4: [value,index] = max($\mathbf{b}_l^*$)
5: $\widetilde{\mathbf{b}}_l$(index) = round(value)
6: sum = sum + $\widetilde{\mathbf{b}}_l$(index)
7: $\mathbf{b}_l^*$(index) = 0
8: **end while**
9: **Output** $\widetilde{\mathbf{b}}_l$

Let $x_{l,k}^q(m)$ denote the quantized observation corresponding to $\widehat{x}_{l,k}(m)$ obtained using $\tilde{b}_{l,k}(m)$ bits as determined above. The quantization process can be modeled as

$$\begin{aligned} x_{l,k}^q(m) &= \widehat{x}_{l,k}(m) + z_{l,k}^q(m) = \theta_k(m) + z_{l,k}(m) + z_{l,k}^q(m) \\ &= \theta_k(m) + e_{l,k}(m), \end{aligned} \tag{30}$$

where $z_{l,k}^q(m)$ denotes the corresponding quantization noise that has a mean zero and variance equal to $\sigma_{z_{l,k}^q}^2(m) = \frac{\Delta_{l,k}^2}{12}$, where $\Delta_{l,k}$ denotes the quantizer's step size for the (l,k)th element [31, Sec. 5.5]. Since the input signal to the quantizer and the resulting quantization noise are uncorrelated [31, Sec. 6.3], the observation noise after the ZF operation $z_{l,k}(m)$ and the quantization noise $z_{l,k}^q(m)$ are also uncorrelated. Therefore, the variance of the effective noise $e_{l,k}(m)$ can be expressed as $\sigma_{e_{l,k}}^2(m) = \sigma_{z_{l,k}}^2(m) + \sigma_{z_{l,k}^q}^2(m)$. Let $\mathbf{e}_l(m) \in \mathbb{C}^{K\times 1}$ denote the stacked vector of effective noise terms $e_{l,k}(m)$ with the covariance matrix of $\mathbf{R}_{e,l}(m) = \text{diag}\left[\sigma_{e_{l,1}}^2(m), \sigma_{e_{l,2}}^2(m), \ldots, \sigma_{e_{l,p}}^2(m)\right] \in \mathbb{C}^{K\times K}$. The pertinent optimization problem of SMSE minimization using quantized measurements can be formulated similar to (15), as shown below

$$\begin{aligned} \underset{\mathbf{f}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{Q}'\mathbf{f} \\ \text{subject to} \quad & \mathbf{W}'\mathbf{f} = \mathbf{g}, \end{aligned} \tag{31}$$

where, the matrices $\mathbf{Q}' \in \mathbb{C}^{LNN_tq\times LNN_tq}$ and $\mathbf{W}' \in \mathbb{C}^{NN_rq\times NN_tp}$ can be obtained by stacking $\mathbf{Q}_l'(m) \in \mathbb{C}^{N_tq\times N_tq}$ and $\mathbf{W}_l'(m) \in \mathbb{C}^{N_rK\times N_tK}$, similar to (15), with $\mathbf{Q}_l'(m) = \left[\left(\mathbf{R}_{e,l}^T(m)\right) \otimes \mathbf{H}_l^H(m)\mathbf{H}_l(m)\right]$ and $\mathbf{W}_l'(m) = \left[\mathbf{I}_K^T \otimes \mathbf{H}_l(m)\right]$. The closed form expression for the optimal vector $\mathbf{f}^*$ is determined as

$$\mathbf{f}^* = \mathbf{Q}'^{-1}\mathbf{W}'^H\left[\mathbf{W}'\mathbf{Q}'^{-1}\mathbf{W}'^H\right]^{-1}\mathbf{g}, \tag{32}$$

from which the individual precoders can be extracted as described in Section-III.

A. Computational Complexity and Communication Overhead

The computational complexity for calculating the precoding vector in (16) or (32) is $\mathcal{O}\Big(\left(LNN_tq\right)^3 + \left(NN_rq\right)^3\Big)$. As it can be observed, the complexity increases as $\mathcal{O}\left(L^3\right)$, which implies that the complexity becomes prohibitively high as the number of sensors in the WSN increases. Detailed step-by-step analysis of the computational complexity is given in our technical report [28, Sec.-I-A].

In the decentralized estimation schemes, the fusion center computes the precoding vector $\mathbf{f}$ followed by feeding back the precoding matrix $\mathbf{F}_l$ to each sensor L. In order to compute $\mathbf{f}$, the fusion center requires the information pertaining to other quantities such as the observation matrices $\mathbf{A}_l$, the observation noise covariance matrices $\mathbf{R}_{v,l}$ and the channel matrices $\mathbf{H}_l(m)$. In a typical WSN deployment, since the fusion center can potentially be situated at a location far from the SNs, the overhead associated with the transmission of this information results in a significant power consumption at the SNs, negatively impacting their lifetime.

Thus, to overcome the drawbacks of increasing computational complexity and a higher overhead burden, the next section proposes distributed algorithms for computation of the optimal precoders for the analog as well as digital schemes developed in Sections-III and -IV, respectively.

V. Distributed Implementation using ADMM

In the decentralized estimation schemes described previously, the fusion center acquires the observation matrices and observation noise statistics at each sensor to determine the optimal precoders, and transmits this information to the sensors over ideal feedback links. While such a procedure is feasible in systems with a few sensor nodes, the computational and communication overheads incurred by the fusion center can be prohibitively high for a system with a large number of sensors. This motivates the distributed solutions proposed in this section, wherein the sensors design their individual precoders via purely inter-sensor communication, without burdening the fusion center. The popular alternating direction method of multipliers (ADMM) offers an excellent framework

for the development of such a distributed transceiver design scheme. Starting with the optimization problem in (15), one can rewrite the objective function and constraint for precoder computation at each SN l as shown below

$$\begin{aligned} \underset{\mathbf{f}_l}{\text{minimize}} \quad & \sum_{l=1}^{L} \mathbf{f}_l^H \mathbf{Q}_l \mathbf{f}_l \\ \text{subject to} \quad & \sum_{l=1}^{L} \mathbf{W}_l \mathbf{f}_l = \mathbf{g}. \end{aligned} \tag{33}$$

The Lagrangian function for the above optimization problem is given as

$$\mathcal{L}\left(\{\mathbf{f}_l\}, \boldsymbol{\lambda}\right) = \sum_{l=1}^{L} \mathbf{f}_l^H \mathbf{Q}_l \mathbf{f}_l + \mathcal{R}\left[\boldsymbol{\lambda}^H \left[\sum_{l=1}^{L} \mathbf{W}_l \mathbf{f}_l - \mathbf{g}\right]\right], \tag{34}$$

where $\boldsymbol{\lambda}$ is the dual variable corresponding to the equality constraint. Using the KKT conditions [27, Sec. 5.5.3] $\nabla_{\mathbf{f}_l} \mathcal{L}\left(\{\mathbf{f}_l^*\}, \boldsymbol{\lambda}^*\right) = 2\mathbf{Q}_l \mathbf{f}_l^* + \mathbf{W}_l^H \boldsymbol{\lambda}^* = \mathbf{0}$, the optimal value of $\mathbf{f}_l^*$ is obtained as

$$\mathbf{f}_l^* = -\frac{1}{2} \mathbf{Q}_l^{-1} \mathbf{W}_l^H \boldsymbol{\lambda}^*. \tag{35}$$

The dual function $g(\boldsymbol{\lambda})$ for (33) is obtained by substituting $\mathbf{f}_l^*$ obtained above in the Lagrangian in (34), yielding

$$g(\boldsymbol{\lambda}) = -\frac{1}{4} \boldsymbol{\lambda}^H \left[\sum_{l=1}^{L} \mathbf{W}_l \mathbf{Q}_l^{-1} \mathbf{W}_l^H\right] \boldsymbol{\lambda} - \mathcal{R}\left[\boldsymbol{\lambda}^H \mathbf{g}\right]. \tag{36}$$

Hence, the equivalent dual optimization problem is given as

$$\min_{\boldsymbol{\lambda}} \quad \boldsymbol{\lambda}^H \left[\sum_{l=1}^{L} \mathbf{W}_l \mathbf{Q}_l^{-1} \mathbf{W}_l^H\right] \boldsymbol{\lambda} + 4\mathcal{R}\left[\boldsymbol{\lambda}^H \mathbf{g}\right]. \tag{37}$$

The above optimization problem can be modified as shown below in order to solve it in a distributed fashion using the dual consensus ADMM framework [32], [33].

$$\begin{aligned} \min_{\boldsymbol{\lambda}_l} \quad & \sum_{l=1}^{L} \left[\boldsymbol{\lambda}_l^H \left[\mathbf{W}_l \mathbf{Q}_l^{-1} \mathbf{W}_l^H\right] \boldsymbol{\lambda}_l + \frac{4}{L} \mathcal{R}\left[\boldsymbol{\lambda}_l^H \mathbf{g}\right]\right] \\ \text{s.t} \quad & \boldsymbol{\lambda}_l = \boldsymbol{\lambda} \ \forall \ l = 1, 2, \cdots, L. \end{aligned} \tag{38}$$

It can be seen that the problems in (37) and (38) have the same solution. The constraint in (38) is termed the consensus constraint that forces the local variables $\boldsymbol{\lambda}_l$ at all SNs to be identical. Since the above optimization objective is separable, each SN l can independently determine $\boldsymbol{\lambda}_l$. The augmented Lagrangian for the problem (38) with a quadratic penalty function for constraint violation is

$$\begin{aligned} \mathcal{L}_\rho(\boldsymbol{\lambda}_l, \boldsymbol{\lambda}, \boldsymbol{\psi}_l) = & \sum_{l=1}^{L} \boldsymbol{\lambda}_l^H \left[\mathbf{W}_l \mathbf{Q}_l^{-1} \mathbf{W}_l^H\right] \boldsymbol{\lambda}_l + \frac{4}{L} \mathcal{R}\left[\boldsymbol{\lambda}^H \mathbf{g}\right] \\ & + \mathcal{R}\left[\boldsymbol{\psi}_l^H (\boldsymbol{\lambda}_l - \boldsymbol{\lambda})\right] + \frac{\rho}{2} ||\boldsymbol{\lambda}_l - \boldsymbol{\lambda}||_2^2, \end{aligned} \tag{39}$$

where the quantities ρ and $\boldsymbol{\psi}_l$ denote the penalty parameter and the dual variable corresponding to the lth sensor, respectively. Using ADMM, one obtains the following iterative steps at the lth SN

$$\boldsymbol{\lambda}_l^{(k+1)} = \underset{\{\boldsymbol{\lambda}_l\}}{\arg\min} \ \mathcal{L}_\rho(\boldsymbol{\lambda}_l, \boldsymbol{\lambda}^k, \boldsymbol{\psi}_l^k) \tag{40}$$

$$\boldsymbol{\lambda}^{(k+1)} = \underset{\boldsymbol{\lambda}}{\arg\min} \ \mathcal{L}_\rho(\boldsymbol{\lambda}_l^{(k+1)}, \boldsymbol{\lambda}, \boldsymbol{\psi}_l^k) \tag{41}$$

$$\boldsymbol{\psi}_l^{(k+1)} = \boldsymbol{\psi}_l^k + \rho(\boldsymbol{\lambda}_l^{(k+1)} - \boldsymbol{\lambda}^{(k+1)}). \tag{42}$$

It is important to note that each sensor $l \in 1, 2, \ldots, L$ performs the above steps in (40)-(42) in parallel. The problems in (40) and (41) correspond to unconstrained quadratic minimization, which can be solved by computing the gradient and equating it to zero. The optimal values of $\boldsymbol{\lambda}_l^{k+1}$ and $\boldsymbol{\lambda}^{k+1}$ can thus be derived as

$$\boldsymbol{\lambda}_l^{(k+1)} = \left[\mathbf{W}_l \mathbf{Q}_l^{-1} \mathbf{W}_l^H + \frac{\rho}{2} \mathbf{I}_{NN_rK}\right]^{-1} \left[\frac{\rho}{2} \boldsymbol{\lambda}^k - \frac{\boldsymbol{\psi}_l^k}{2} - \frac{2\mathbf{g}}{L}\right], \tag{43}$$

$$\boldsymbol{\lambda}^{(k+1)} = \frac{1}{L} \sum_{l=1}^{L} \left[\boldsymbol{\lambda}_l^{(k+1)} + \frac{1}{\rho} \boldsymbol{\psi}_l^k\right]. \tag{44}$$

Using (42) and (44), it can be shown that $\frac{1}{L} \sum_{l=1}^{L} \boldsymbol{\psi}_l^k = \mathbf{0}$. The global optimal variable $\boldsymbol{\lambda}$ can be updated as

$$\boldsymbol{\lambda}^{(k+1)} = \frac{1}{L} \sum_{l=1}^{L} \boldsymbol{\lambda}_l^{(k+1)}. \tag{45}$$

The SNs can broadcast the individual $\boldsymbol{\lambda}_i$ to enable the computation of the parameter $\boldsymbol{\lambda}$ at each SN. The optimal precoding vector $\mathbf{f}_i$ at each SN l can now be determined by substituting the value of $\boldsymbol{\lambda}$ in (35). The procedure for distributed evaluation of the precoders with quantized measurements can be obtained on similar lines by starting with the corresponding optimization problem in (31) and decoupling it similar to (33).

A. Quantized Sensor Observations

Similarly, one can also develop a framework for distributed computation of the solution to the optimization problem in (31) for the quantized sensor observations scenario. The analysis will result in equations similar to the analog transmission scenario above at each step, with the matrices $\mathbf{Q}$ and $\mathbf{W}$ replaced by $\mathbf{Q}'$ and $\mathbf{W}'$, respectively. The optimal precoding vector $\mathbf{f}_l^*$ for each SN l is given as

$$\mathbf{f}_l^* = -\frac{1}{2} (\mathbf{Q}_l')^{-1} (\mathbf{W}_l')^H \boldsymbol{\lambda}_q^*, \tag{46}$$

where $\boldsymbol{\lambda}_q^*$ is the optimal dual variable for the quantized SN observations scenario, with its update equation given as

$$\boldsymbol{\lambda}_q^{(k+1)} = \frac{1}{L} \sum_{l=1}^{L} \boldsymbol{\lambda}_{q,l}^{(k+1)}. \tag{47}$$

The individual SN variables $\boldsymbol{\lambda}_{q,l}^{(k+1)}$ and $\mathbf{y}_{q,l}^{(k+1)}$ for each l, are updated as

$$\boldsymbol{\lambda}_{q,l}^{(k+1)} = \left[\mathbf{W}_l' (\mathbf{Q}_l')^{-1} (\mathbf{W}_l')^H + \frac{\rho}{2} \mathbf{I}_{NN_rK}\right]^{-1} \left[\frac{\rho}{2} \boldsymbol{\lambda}_q^k - \frac{\boldsymbol{\psi}_{q,l}^k}{2} - \frac{2\mathbf{g}}{L}\right] \tag{48}$$

$$\boldsymbol{\psi}_{q,l}^{(k+1)} = \boldsymbol{\psi}_{q,l}^k + \rho \left[\boldsymbol{\lambda}_{q,l}^{(k+1)} - \boldsymbol{\lambda}_q^{(k+1)}\right], \tag{49}$$

respectively.

B. Computational Complexity and Communication Overhead

In the distributed precoder design scheme, the per iteration computational complexity of evaluating the precoder vector $\mathbf{f}_l$, individual dual variable $\boldsymbol{\lambda}_l$ and global dual variable $\boldsymbol{\lambda}$ using (35), (43) and (44) is $\mathcal{O}\left(\left(NN_tq\right)^3+\left(NN_rq\right)^3\right)$, for each SN l. A detailed analysis is given in our technical report [28, Sec.-I-B]. It can be observed that the computational requirement of the distributed scheme is independent of the number of sensors L and does not grow exponentially as the number of sensors increases. Hence, this makes the distributed scheme more suitable for the implementation in large WSNs.

One can also analyze the overhead arising due to message exchanges in the distributed scheme. In each iteration every SN l has to share only its dual variable $\boldsymbol{\lambda}_l$ with the other sensors in the WSN. Naturally, the overhead arising due to such an exchange is negligible in comparison to the amount of feedback required for the decentralized scheme, wherein each sensor has to transmit information about the observation, noise covariance and channel matrices, which also increases with the number of sensors in the WSN.

C. Convergence Analysis

The convergence of the proposed ADMM based distributed scheme requires that the optimization objective of each SN l in (38), given as $g(\boldsymbol{\lambda}_l)=\boldsymbol{\lambda}_l^H\left[\mathbf{W}_l\mathbf{Q}_l^{-1}\mathbf{W}_l^H\right]\boldsymbol{\lambda}_l+\frac{4}{L}\mathcal{R}\left[\boldsymbol{\lambda}_l^H\mathbf{g}\right]$, to be strongly convex in addition to the gradient being Lipschitz continuous [34, Scenario-2,Tab-1].

For the lth SN, the condition for strong convexity requires that the Hessian matrix $\nabla^2 g(\boldsymbol{\lambda}_l)=\mathbf{W}_l\mathbf{Q}_l^{-1}\mathbf{W}_l^H\succ\mathbf{0}$, which is satisfied for the given problem. Hence, the objective function of each SN l is strongly convex [27, Sec. 3.1.4].

The condition for the gradient of a function $g(\boldsymbol{\lambda}_l)$ to be Lipschitz continuous requires that for a constant $\beta_l>0$ and $\forall\tilde{\boldsymbol{\lambda}}_l,\boldsymbol{\lambda}_l\in dom(g)$

$$\left|\left|\nabla g\left(\tilde{\boldsymbol{\lambda}}_l\right)-\nabla g(\boldsymbol{\lambda}_l)\right|\right|_2\leq\beta_l\left|\left|\tilde{\boldsymbol{\lambda}}_l-\boldsymbol{\lambda}_l\right|\right|_2.$$

One can easily verify that the above condition is also satisfied with $\beta_l=\left|\left|\mathbf{W}_l\mathbf{Q}_l^{-1}\mathbf{W}_l^H\right|\right|_F>0$. Hence, the proposed distributed scheme always converges to the globally optimum solution.

The analysis thus far considered the availability of perfect CSI at the fusion center/ SNs. However, frequently in practice, only imperfect CSI is available due to channel estimation error and limited feedback. Hence, in order to mitigate the effect of such imperfections on the system performance, the next section proposes robust precoder designs considering CSI uncertainty, modeled using both the probabilistic as well as deterministic models.

VI. Robust Precoder Designs with CSI Uncertainty

It must be noted that CSI uncertainty is inevitable in practice due to various limitations such as estimation error due to a limited pilot overhead, quantization error and feedback delay. Ignoring the CSI uncertainty, and relying purely on the available CSI estimate, as is done by the uncertainty agnostic design, leads to a degradation in the estimation performance at the fusion center. Therefore, for practical viability, this section proposes robust precoder/ combiner design procedures which account for the CSI uncertainty to achieve improved estimation performance. Furthermore, transceiver techniques are proposed considering both the popular stochastic [35] and norm-ball CSI uncertainty [36] models, to achieve average MSE as well as worst-case MSE minimization, which makes the study comprehensive in nature. The CSI uncertainty corresponding to $\mathbf{H}_l(m)$, can be modelled as

$$\mathbf{H}_l(m)=\widehat{\mathbf{H}}_l(m)+\Delta\mathbf{H}_l(m),\tag{50}$$

where $\widehat{\mathbf{H}}_l(m)\in\mathbb{C}^{N_r\times N_t}$ is the available estimate of the channel at the fusion center and $\Delta\mathbf{H}_l(m)\in\mathbb{C}^{N_r\times N_t}$ represents the error in the estimate.

A. Robust Precoder Design with Stochastic CSI Uncertainty

In the stochastic CSI uncertainty model, we consider each element of the estimation error matrix $\Delta\mathbf{H}_l(m)$ to be distributed as a zero mean i.i.d. random variable of variance σ_H^2.

1) Analog sensor observations: Substituting the expression for $\mathbf{H}_l(m)$ from (50) into (8), and using the estimation constraint $\sum_{l=1}^{L}\widehat{\mathbf{H}}_l(m)\mathbf{F}_l(m)\mathbf{A}_l=\mathbf{I}_K$, the resulting SMSE can be determined as shown in (51). The results from the following lemma are being used for the simplification of the various terms of the SMSE expression in (51).

Lemma 1. *Let $\mathbf{h}_l=vec\left[\mathbf{H}_l(m)\right]\in\mathbb{C}^{N_rN_t\times1}$ be a complex Gaussian random vector with mean $\widehat{\mathbf{h}}_l=vec\left[\widehat{\mathbf{H}}_l(m)\right]$ and covariance matrix $\sigma_H^2\mathbf{I}_{N_rN_t}$. It follows that [37]*

$$\begin{aligned}&\mathbb{E}\left[\Delta\mathbf{H}_l(m)\mathbf{X}\mathbf{X}^H\Delta\mathbf{H}_l^H(m)\right]=\sigma_H^2\mathrm{Tr}\left[\mathbf{X}\mathbf{X}^H\right]\mathbf{I}_{N_r},\\&\mathbb{E}\left[\widehat{\mathbf{H}}_l(m)\mathbf{X}\mathbf{X}^H\Delta\mathbf{H}_l^H(m)\right]=\mathbf{0},\mathbb{E}\left[\Delta\mathbf{H}_l(m)\mathbf{X}\mathbf{X}^H\widehat{\mathbf{H}}_l^H(m)\right]=\mathbf{0}.\end{aligned}\tag{52}$$

Using the first result of Lemma 1, the first term on the right hand side of (51) can be averaged as

$$\begin{aligned}&\mathbb{E}\left[\sum_{m=0}^{N-1}\sum_{l=1}^{L}\left[\mathrm{Tr}\Big[\Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{A}_l\mathbf{R}_\theta(m)\mathbf{A}_l^H\mathbf{F}_l^H(m)\Delta\mathbf{H}_l^H(m)\Big]\right]\right]\\&=\sigma_H^2N_r\sum_{l=1}^{L}\sum_{m=0}^{N-1}\mathbf{f}_l^H(m)\mathbf{T}_l(m)\mathbf{f}_l(m)=\alpha\sum_{l=1}^{L}\mathbf{f}_l^H\mathbf{T}_l\mathbf{f}_l\\&=\alpha\mathbf{f}^H\mathbf{T}\mathbf{f}.\end{aligned}\tag{53}$$

The matrix $\mathbf{T}\in\mathbb{C}^{LNN_tq\times LNN_tq}$ is a block diagonal matrix defined as $\mathbf{T}=\mathrm{diag}\left[\mathbf{T}_1,\mathbf{T}_2,\ldots,\mathbf{T}_L\right]$, where the matrices $\mathbf{T}_l\in\mathbb{C}^{NN_tq\times NN_tq}$ and $\mathbf{T}_l(m)\in\mathbb{C}^{N_tq\times N_tq}$ are defined as $\mathbf{T}_l=\mathrm{diag}\Big[\mathbf{T}_1(m),\mathbf{T}_2(m),\ldots,\mathbf{T}_L(m)\Big]$ and $\mathbf{T}_l(m)=\Big[\left[\mathbf{A}_l\mathbf{R}_\theta(m)\mathbf{A}_l^H\right]^T\otimes\mathbf{I}_{N_t}\Big]$, for $1\leq l\leq L$ and $0\leq m\leq N-1$, respectively, and the scalar quantity is $\alpha=\sigma_H^2N_r$. A detailed explanation is given in the technical report [28, Sec.

$$\text{SMSE} = \sum_{m=0}^{N-1}\sum_{l=1}^{L}\text{Tr}\Big[\Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{A}_l\mathbf{R}_\theta(m)\mathbf{A}_l^H\mathbf{F}_l^H(m)\Delta\mathbf{H}_l^H(m)\Big] + \text{Tr}\Big[\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)\Big] + \sum_{m=0}^{N-1}\text{Tr}\left[\mathbf{R}_u(m)\right]. \tag{51}$$

II-A]. Upon substituting $\mathbf{H}_l(m)$ from (50) in the second term of (51), the resultant terms can be simplified as

$$\mathbb{E}\left[\sum_{l=1}^{L}\sum_{m=0}^{N-1}\text{Tr}\Big[\Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\Delta\mathbf{H}_l^H(m)\Big]\right] = \alpha\sum_{l=1}^{L}\sum_{m=0}^{N-1}\mathbf{f}_l^H(m)\mathbf{P}_l(m)\mathbf{f}_l(m) = \alpha\sum_{l=1}^{L}\mathbf{f}_l^H\mathbf{P}_l\mathbf{f}_l = \alpha\mathbf{f}^H\mathbf{P}\mathbf{f}, \tag{54}$$

$$\sum_{l=1}^{L}\sum_{m=0}^{N-1}\text{Tr}\Big[\widehat{\mathbf{H}}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\widehat{\mathbf{H}}_l^H(m)\Big] = \sum_{l=1}^{L}\sum_{m=0}^{N-1}\mathbf{f}_l^H(m)\mathbf{C}_l(m)\mathbf{f}_l(m) = \sum_{l=1}^{L}\mathbf{f}_l^H\mathbf{C}_l\mathbf{f}_l = \mathbf{f}^H\mathbf{C}\mathbf{f}, \tag{55}$$

where the matrices obey $\mathbf{P}_l(m) = \left[\mathbf{R}_{v,l}^H(m)\otimes\mathbf{I}_{N_t}\right] \in \mathbb{C}^{qN_t\times qN_t}$, $\mathbf{P}_l = \text{diag}\left[\mathbf{P}_l(0),\mathbf{P}_l(1)\ldots,\mathbf{P}_l(N-1)\right] \in \mathbb{C}^{NqN_t\times NqN_t}$ and $\mathbf{P} = \text{diag}[\mathbf{P}_1,\mathbf{P}_2,\ldots,\mathbf{P}_L] \in \mathbb{C}^{LNqN_t\times LNqN_t}$. Also, the matrix $\mathbf{C}_l(m) = \left[\mathbf{R}_{v,l}^T(m)\otimes\widehat{\mathbf{H}}_l^H(m)\widehat{\mathbf{H}}_l(m)\right] \in \mathbb{C}^{qN_t\times qN_t}$ and the matrices $\mathbf{C}\in\mathbb{C}^{LNqN_t\times LNqN_t}$, $\mathbf{C}_l\in\mathbb{C}^{NqN_t\times NqN_t}$ are block-diagonal with blocks $\mathbf{S}_l$ and $\mathbf{S}_l(m)$, respectively. The simplification of (54) is given in technical report [28, Sec.-II-A]. Using these results, the average SMSE becomes

$$\text{Average SMSE} = \mathbf{f}^H\left[\mathbf{C}+\alpha\left(\mathbf{P}+\mathbf{T}\right)\right]\mathbf{f} + \sum_{m=0}^{N-1}\text{Tr}\left[\mathbf{R}_u(m)\right] = \mathbf{f}^H\mathbf{\Omega}\mathbf{f} + \sum_{m=0}^{N-1}\text{Tr}\left[\mathbf{R}_u(m)\right], \tag{56}$$

where $\mathbf{\Omega} = \left[\mathbf{C}+\alpha\left(\mathbf{P}+\mathbf{T}\right)\right] \in \mathbb{C}^{LNqN_t\times LNqN_t}$. Hence, the pertinent optimization problem that minimizes the average SMSE for this scenario with stochastic uncertainty can be framed as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{\Omega}\mathbf{f} \\ \text{subject to} \quad & \bar{\mathbf{W}}\mathbf{f} = \mathbf{g},\end{aligned} \tag{57}$$

where the matrix $\bar{\mathbf{W}}$ has a structure similar to $\mathbf{W}$ with $\mathbf{H}$ replaced by $\widehat{\mathbf{H}}$. The above optimization problem can be solved using an approach similar to the one employed in (15). The optimal precoder thus obtained will be similar to (16), with the matrices $\mathbf{Q}$ and $\mathbf{W}$ replaced by $\mathbf{\Omega}$ and $\bar{\mathbf{W}}$, respectively.

2) Quantized sensor observations: For the scenario with quantized observations and CSI uncertainty, the vector $\mathbf{y}(m)$ received at the fusion center is given in (58), where $\mathbf{e}_l(m)$ denotes the effective quantization and estimation noise as defined in Section-IV. The SMSE of robust parameter estimation for the mth subcarrier can be obtained as shown in (59). The various terms in (59) can be further simplified as shown in the technical report [28, Sec.-II-B], and the expression of SMSE, ignoring a constant term that does not affect the optimization procedure, can be stated in the final compact form of $\mathbf{f}^H\mathbf{\Omega}_q\mathbf{f}$, where $\mathbf{\Omega}_q = \mathbf{C}_q + \alpha\left(\mathbf{P}_q+\mathbf{T}_q\right) \in \mathbb{C}^{LNqN_t\times LNqN_t}$. The pertinent optimization problem that minimizes the average SMSE for this scenario with quantized sensor observations and stochastic CSI uncertainty can be formulated as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & \mathbf{f}^H\mathbf{\Omega}_q\mathbf{f} \\ \text{subject to} \quad & \bar{\mathbf{W}}_q\mathbf{f} = \mathbf{g},\end{aligned} \tag{60}$$

where the matrix $\bar{\mathbf{W}}_q$ has a similar structure to $\mathbf{W}$ with $\mathbf{H}$ replaced by $\widehat{\mathbf{H}}$ and $\mathbf{A}_l$ replaced by the identity matrix. Upon solving the above optimization problem, similar to (15), the optimal precoder is obtained as $\mathbf{f}^* = \mathbf{\Omega}_q^{-1}\bar{\mathbf{W}}_q^H\left[\bar{\mathbf{W}}_q\mathbf{\Omega}_q^{-1}\bar{\mathbf{W}}_q^H\right]^{-1}\mathbf{g}$. The next subsection deals with the scenarios when the channel estimation error is deterministic in nature and robust precoder designs are proposed which minimize the worst case SMSE for the scenarios having both analog and quantized sensor observations.

B. Robust Precoder Design with Norm Ball CSI Uncertainty

In this popular alternative model of CSI uncertainty, the estimation error $\Delta\mathbf{H}_l(m)$ is modeled using the bounded uncertainty model of

$$||\Delta\mathbf{H}_l(m)||_F \le \epsilon. \tag{61}$$

This framework is interesting and also challenging since one has to determine the expression for the worst-case SMSE followed by designing the optimal precoder that minimizes it.

1) Analog sensor observations: The first and second terms of the SMSE expressions in (51) can be simplified as

$$\text{Tr}\Big[\Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{A}_l\mathbf{R}_{\boldsymbol{\theta}}(m)\mathbf{A}_l^H\mathbf{F}_l^H(m)\Delta\mathbf{H}_l^H(m)\Big] = \mathbf{f}_l^H(m)\mathbf{J}_l(m)\mathbf{f}_l(m) \tag{62}$$

$$\text{Tr}\Big[\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{v,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)\Big] = \mathbf{f}_l^H(m)\mathbf{Q}_l(m)\mathbf{f}_l(m), \tag{63}$$

where $\mathbf{J}_l(m) = \left[\left[\mathbf{A}_l(m)\mathbf{R}_\theta(m)\mathbf{A}_l^H(m)\right]\otimes\Delta\mathbf{H}_l^H(m)\Delta\mathbf{H}_l(m)\right] = \mathbf{G}_l^H(m)\mathbf{G}_l(m) \in \mathbb{C}^{qN_t\times qN_t}$ and $\mathbf{Q}_l(m) = \mathbf{L}_l^H(m)\mathbf{L}_l(m) \in \mathbb{C}^{qN_t\times qN_t}$, with the matrices

$$\mathbf{y}(m) = \sum_{l=1}^{L} \widehat{\mathbf{H}}_l(m)\mathbf{F}_l(m)\boldsymbol{\theta}(m) + \sum_{l=1}^{L} \Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\boldsymbol{\theta}(m) + \sum_{l=1}^{L} \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{e}_l(m) + \mathbf{u}(m). \tag{58}$$

$$\text{SMSE} = \sum_{l=1}^{L}\sum_{m=0}^{N-1} \text{Tr}\left[\Delta\mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_\theta(m)\mathbf{F}_l^H(m)\Delta\mathbf{H}_l^H(m) + \mathbf{H}_l(m)\mathbf{F}_l(m)\mathbf{R}_{e,l}(m)\mathbf{F}_l^H(m)\mathbf{H}_l^H(m)\right] + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right] \tag{59}$$

$\mathbf{G}_l(m) \in \mathbb{C}^{KN_r \times KN_t}$ and $\mathbf{L}_l(m) \in \mathbb{C}^{qN_r \times qN_t}$ defined as $\mathbf{G}_l(m) = \left[\mathbf{R}_\theta^{\frac{H}{2}}(m)\mathbf{A}_l^H \otimes \Delta\mathbf{H}_l(m)\right]$ and $\mathbf{L}_l(m) = \left[\mathbf{R}_{v,l}^{\frac{1}{2}}(m) \otimes \mathbf{H}_l(m)\right]$. Substituting the expression for $\mathbf{H}_l(m)$ from the norm-ball CSI uncertainty model in (50), we can simplify the resultant expression for $\mathbf{L}_l(m) = \left[\mathbf{R}_{v,l}^{\frac{1}{2}}(m) \otimes \left[\widehat{\mathbf{H}}_l(m) + \Delta\mathbf{H}_l(m)\right]\right]$ as

$$\begin{aligned}\mathbf{L}_l(m) &= \left[\mathbf{R}_{v,l}^{\frac{1}{2}}(m) \otimes \widehat{\mathbf{H}}_l(m)\right] + \left[\mathbf{R}_{v,l}^{\frac{1}{2}}(m) \otimes \Delta\mathbf{H}_l(m)\right] \\ &= \widehat{\mathbf{L}}_l(m) + \Delta\mathbf{L}_l(m).\end{aligned} \tag{64}$$

As shown in technical report [28, Sec.-II-C], the SMSE expression, for this norm-ball uncertainty scenario can be simplified as

$$\text{SMSE} = \|\mathbf{L}\mathbf{f}\|^2 + \|\mathbf{G}\mathbf{f}\|^2 + \sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right]. \tag{65}$$

Since $\sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right]$ is a constant that does not depend on the precoding matrices, this can be ignored in the subsequent minimization. The quantity $\|\mathbf{L}\mathbf{f}\|^2$ can be further bounded, as shown in technical report [28, Sec.-II-C]

$$\|\mathbf{L}\mathbf{f}\|^2 \leq ||\widehat{\mathbf{L}}\mathbf{f}||^2 + \eta^2 \|\mathbf{f}\|^2 + 2\eta||\widehat{\mathbf{L}}\mathbf{f}|| \, \|\mathbf{f}\|, \tag{66}$$

where the constant quantity is expressed as $\eta = \sqrt{NLq\epsilon^2}$. The term $\|\mathbf{G}\mathbf{f}\|^2$ can be upper-bounded as

$$\|\mathbf{G}\mathbf{f}\|^2 \leq \|\mathbf{G}\|_F^2 \|\mathbf{f}\|^2 \leq \mu \|\mathbf{f}\|^2, \tag{67}$$

where the quantity μ is a constant defined as $\mu = \sqrt{\sum_{m=0}^{N-1}\sum_{l=1}^{L} \epsilon^2 \text{Tr}\left(\mathbf{R}_\theta(m)\right) ||\mathbf{A}_l||_F^2}$. The last inequality in (67) follows from the fact that $\|\mathbf{G}\|_F \leq \mu$ as shown in technical report [28]. Substituting the bounds for $\|\mathbf{L}\mathbf{f}\|$ and $\|\mathbf{G}\mathbf{f}\|$ from (66) and (67) into (65), yields the net bound for the SMSE as

$$\text{SMSE} \leq ||\widehat{\mathbf{L}}\mathbf{f}||^2 + \left(\eta^2 + \mu^2\right)\|\mathbf{f}\|^2 + 2\eta||\widehat{\mathbf{L}}\mathbf{f}|| \, \|\mathbf{f}\|. \tag{68}$$

For mathematical tractability, the proposed design procedure determines the precoders for this scenario via the minimization of the SMSE bound in (68). The corresponding constrained optimization problem can be formulated as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & ||\widehat{\mathbf{L}}\mathbf{f}||^2 + \left(\eta^2 + \mu^2\right)\|\mathbf{f}\|^2 + 2\eta||\widehat{\mathbf{L}}\mathbf{f}|| \, \|\mathbf{f}\| \\ \text{subject to} \quad & \bar{\mathbf{W}}\mathbf{f} = \mathbf{g}.\end{aligned} \tag{69}$$

The above optimization problem is convex in nature, since the objective is convex and the constraint is linear in terms of the optimization variable $\mathbf{f}$. It can hence be solved efficiently using convex solvers such as CVX [29].

2) Quantized sensor observations: For the scenario of quantized sensor observations, following a procedure similar to the previous case and ignoring the constant $\sum_{m=0}^{N-1} \text{Tr}\left[\mathbf{R}_u(m)\right]$, the SMSE can be equivalently written in a compact form as

$$\text{SMSE} \equiv \|\mathbf{L}_q\mathbf{f}\|^2 + \|\mathbf{G}_q\mathbf{f}\|^2, \tag{70}$$

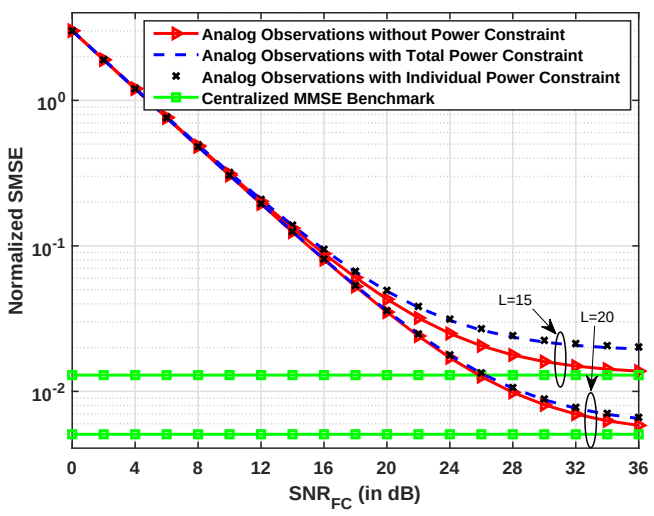
where the matrices $\mathbf{L}_q$ and $\mathbf{G}_q$ of sizes $(NLqN_r \times NLqN_t)$ and $(LNKN_r \times LNqN_t)$, respectively, are block diagonal in nature with $\mathbf{L}_{q,l} \in \mathbb{C}^{NqN_r \times NqN_t}$ and $\mathbf{G}_{q,l} \in \mathbb{C}^{NKN_r \times NqN_t}$, which are in turn block diagonal in conjunction with $\mathbf{L}_{q,l}(m) \in \mathbb{C}^{qN_r \times qN_t}$, $\mathbf{G}_{q,l}(m) \in \mathbb{C}^{KN_r \times qN_t}$, $0 \leq m \leq N-1$, where the matrices on the principal diagonal are defined as $\mathbf{L}_{q,l}(m) = \left[\left(\mathbf{R}_{e,l}^{\frac{1}{2}}(m)\right) \otimes \mathbf{H}_l(m)\right]$ and $\mathbf{G}_{q,l}(m) = \left[\mathbf{R}_\theta^{\frac{H}{2}}(m) \otimes \Delta\mathbf{H}_l(m)\right]$. Following a procedure similar to [28, Sec. II-C], one can again find the upper bound of $\|\mathbf{L}_q\mathbf{f}\|^2$, as

$$\begin{aligned}\|\mathbf{L}_q\mathbf{f}\|^2 &\leq ||\widehat{\mathbf{L}}_q\mathbf{f}||^2 + \|\Delta\mathbf{L}_q\|_F^2 \|\mathbf{f}\|^2 + 2||\widehat{\mathbf{L}}_q\mathbf{f}|| \, \|\Delta\mathbf{L}_q\|_F \|\mathbf{f}\|, \\ &\leq ||\widehat{\mathbf{L}}_q\mathbf{f}||^2 + (\eta_q)^2 \|\mathbf{f}\|^2 + 2\eta_q||\widehat{\mathbf{L}}_q\mathbf{f}|| \, \|\mathbf{f}\|,\end{aligned} \tag{71}$$

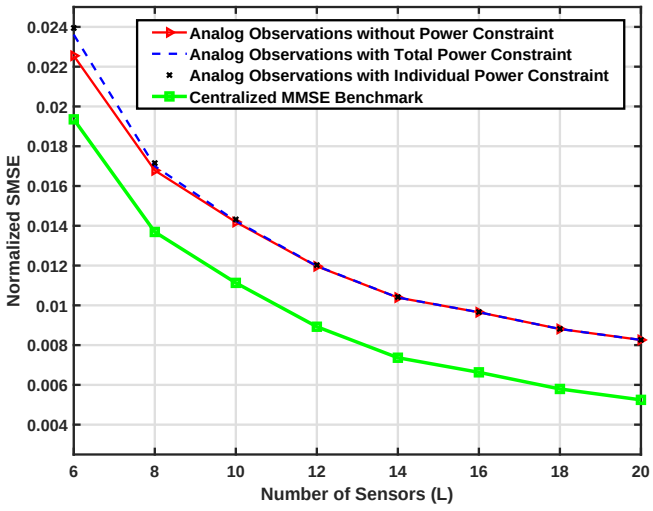
where $\widehat{\mathbf{L}}_q$ and $\Delta\mathbf{L}_q$ are block diagonal and constructed from $\widehat{\mathbf{L}}_{q,i}(m) = \left[\mathbf{R}_{e,l}^{\frac{1}{2}}(m) \otimes \widehat{\mathbf{H}}_l(m)\right] \in \mathbb{C}^{qN_r \times qN_t}$ and $\Delta\mathbf{L}_{q,l}(m) = \left[\mathbf{R}_{e,l}^{\frac{1}{2}}(m) \otimes \Delta\mathbf{H}_l(m)\right] \in \mathbb{C}^{qN_r \times qN_t}$ similar to $\mathbf{L}_q$. The last inequality in (71) follows from the fact that $\|\Delta\mathbf{L}_q\|_F \leq \sqrt{\epsilon^2 \sum_{m=0}^{N-1}\sum_{l=1}^{L} \text{Tr}\left(\mathbf{R}_{e,l}(m)\right)} = \eta_q$, which is obtained using a procedure similar to [28, Sec. II-C], where $\mathbf{R}_{v,l}(m)$ is replaced by $\mathbf{R}_{e,l}(m)$. One can also determine an upper bound for the term $\|\mathbf{G}_q\mathbf{f}\|^2$ as $\|\mathbf{G}_q\mathbf{f}\|^2 \leq \|\mathbf{G}_q\|_F^2 \|\mathbf{f}\|^2 \leq \mu_q^2 \|\mathbf{f}\|^2$, where the last inequality follows from the fact that $\|\mathbf{G}_q\mathbf{f}\| \leq \sqrt{\sum_{m=0}^{N-1}\sum_{l=1}^{L} \epsilon^2 \text{Tr}\left(\mathbf{R}_\theta(m)\right)} = \mu_q$, which can be derived following a procedure similar to the one employed in [28, Sec. II-C], with $\mathbf{A}_l\mathbf{R}_\theta(m)\mathbf{A}_l^H$ replaced by $\mathbf{R}_\theta(m)$. Hence, the optimization problem of minimizing the worst-case cumulative estimation error with CSI uncertainty and quantized measurements can be formulated as

$$\begin{aligned}\underset{\mathbf{f}}{\text{minimize}} \quad & ||\widehat{\mathbf{L}}_q\mathbf{f}||^2 + \left(\eta_q^2 + \mu_q^2\right)\|\mathbf{f}\|^2 + 2\eta_q||\widehat{\mathbf{L}}_q\mathbf{f}|| \, \|\mathbf{f}\| \\ \text{subject to} \quad & \bar{\mathbf{W}}_q\mathbf{f} = \mathbf{g},\end{aligned} \tag{72}$$

which can once again be solved using convenient tools such as CVX [29]. The next section presents the simulation results to illustrate and compare the performance of the proposed schemes.



(a)



(b)

Fig. 2: (a) Normalized SMSE versus SNR_{FC} (in dB) performance for decentralized parameter estimation. (b) Normalized SMSE performance versus the number of sensors at SNR_{FC}=20 dB.

VII. Simulation Results

This section presents our simulation results to demonstrate the performance of the proposed scheme and to verify the accuracy of our analytical formulations. The components of the temporally evolving parameter vector $\tilde{\boldsymbol{\theta}}$ of size $K = 3$ are generated using the AR-2 model, in which the kth element at time instant n is given by

$$\tilde{\theta}_k(n) = \phi_{1,k}\tilde{\theta}_k(n-1) + \phi_{2,k}\tilde{\theta}_k(n-2) + \tilde{w}(n), \qquad (73)$$

where $\tilde{w}$ is a circularly-symmetric zero-mean complex Gaussian noise process with variance $\sigma^2_{\tilde{w}}$, $\phi_{1,k}$ are set as 0.6, 0.5, 0.4 for $k = 1, 2$ and 3, respectively, and $\phi_{2,k} = 0.2\ \forall\ k$. The number of subcarriers at each SN and the total number of sensors in the WSN are $N = 64$ and $L = 10$, respectively. The elements of the observation matrix $\mathbf{A}_l$ are generated as i.i.d. complex Gaussian random variables with mean zero and unit variance. The penalty parameter ρ is set as 10. The frequency domain channel matrix $\mathbf{H}_l$ is obtained by performing N-point FFT of the 5 tap frequency selective time domain channel whose each element is generated as $\mathcal{CN}(0, 1)$. The The number of antennas at each SN and the fusion center is set as $N_t = 3$ and $N_r = 3$, respectively. The number of observations at each subcarrier of each SN is $q = 3$. Furthermore, the elements of the observation noise vectors $\tilde{\mathbf{v}}_l(n)$ and the channel noise vectors $\tilde{\mathbf{u}}_l(n)$ are generated as i.i.d. complex Gaussian random variables, distributed as $\mathcal{CN}(0, \sigma_v^2\mathbf{I}_q)$ and $\mathcal{CN}(0, \sigma_u^2\mathbf{I}_{N_r})$, respectively, where $\sigma_v^2 = \frac{1}{\text{SNR}_{\text{OB}}}$ and $\sigma_u^2 = \frac{1}{\text{SNR}_{\text{FC}}}$. The value of the observation noise SNR, denoted by SNR_{OB}, is set to 10 dB and the SNR at the fusion center, denoted by SNR_{FC}, is either varied suitably or mentioned explicitly in the caption of that particular figure. The normalized SMSE of the parameter estimate is defined as $\text{Normalized SMSE} = \frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\left|\left|\widehat{\boldsymbol{\theta}}(m) - \boldsymbol{\theta}(m)\right|\right|^2$.

Fig. 2(a) depicts the normalized SMSE performance for the decentralized parameter estimation schemes with perfect CSI and analog SN observation transmission described in Section-III, versus SNR_{FC}. It can be seen that upon increasing SNR_{FC}, the normalized SMSE performance of all the proposed decentralized estimation schemes improves, and the scheme of (15), approaches the centralized MMSE benchmark at high SNR_{FC}. Furthermore, the normalized SMSE performance of the proposed schemes subject to both total and per sensor power constraints follow closely. Fig. 2(b) plots the normalized SMSE performance of all the decentralized estimation schemes with perfect CSI and analog SN observation transmission proposed in Section-III for a varying number of SNs L in the WSN. It can be seen that as the number of SNs increases, the proposed decentralized schemes in Section-III once again yield a MSE performance that is very close to the centralized MMSE benchmark, thus demonstrating their

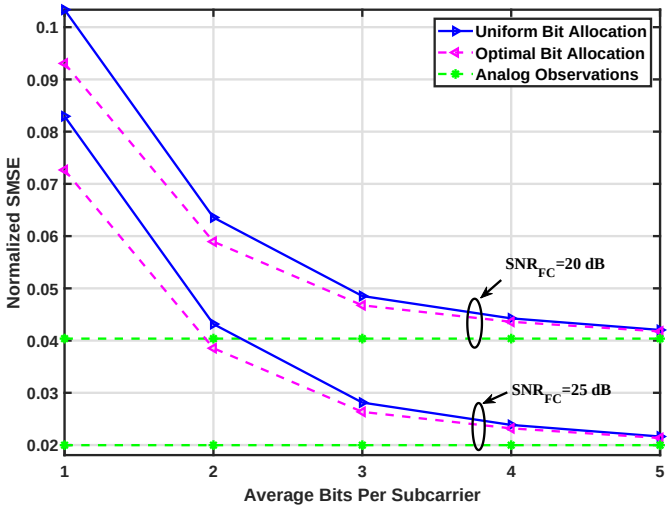
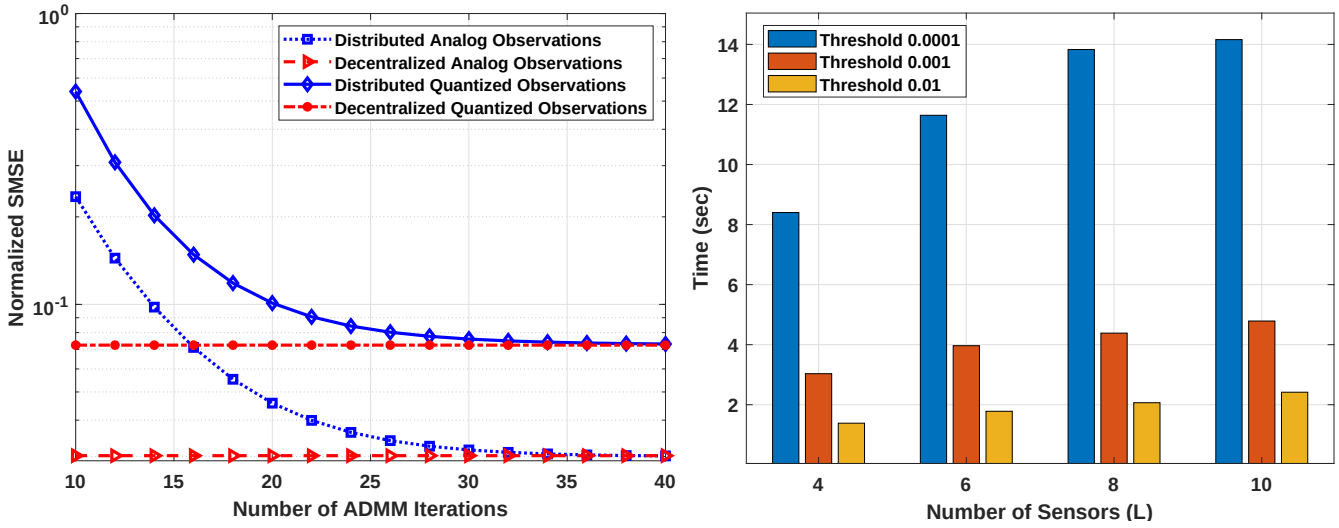
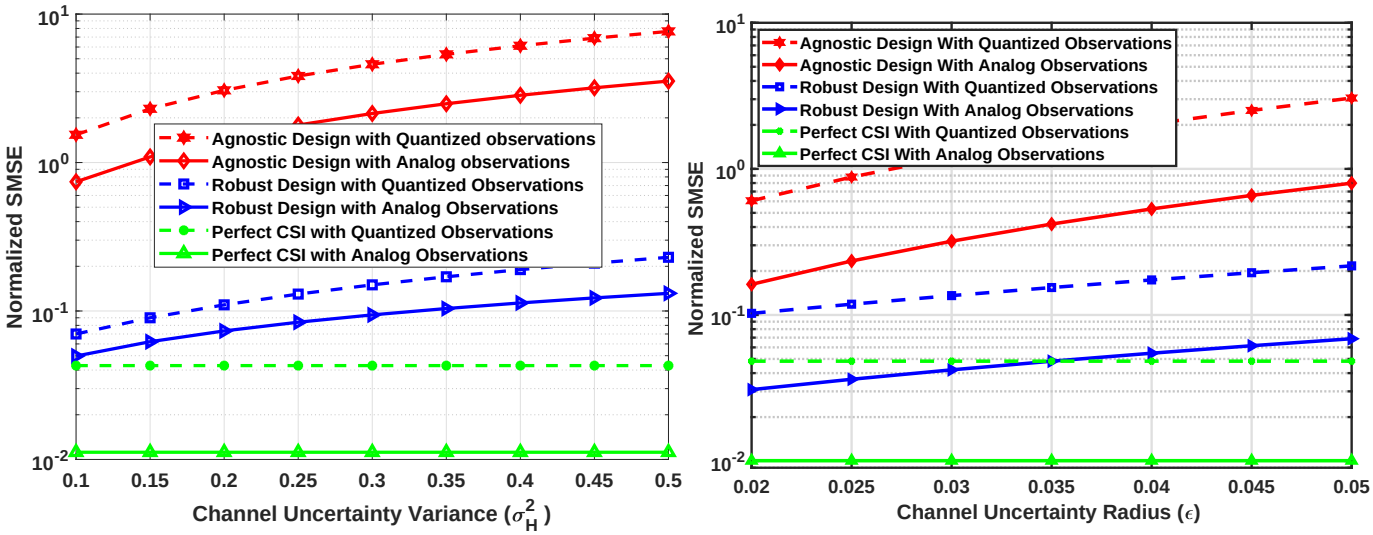


Fig. 3: Normalized SMSE versus the average number of bits per subcarrier with quantized sensor observations.



(a) (b)

Fig. 4: (a) Normalized SMSE versus the number of iterations for the distributed design based on ADMM with SNR_{FC}= 20 dB (b) Convergence time versus the number of sensors L in the network for different convergence thresholds.



(a) (b)

Fig. 5: (a) Normalized SMSE versus channel uncertainty variance σ_H^2 for the scenario with stochastic CSI uncertainty. (b) Normalized SMSE versus channel error bound ϵ for the scenario with norm ball CSI uncertainty.

efficiency.

Fig. 3 shows the normalized SMSE performance of the quantized MVDP scheme in vector parameter estimation. The performance is also compared to that of a quantizer with uniform bit allocation for all the subcarriers. It can be seen that the optimal bit allocation scheme yields a significantly lower MSE than the uniform bit allocation. For instance, for an MSE of 0.0589 with $\text{SNR}_{\text{FC}} = 20$ dB, the former requires approximately 24 percent fewer bits in comparison to the latter, thus leading to a saving in terms of the bandwidth. Furthermore, as the total number of bits increases, the optimal bit allocation scheme approaches the benchmark corresponding to the transmission of analog observations that have infinite precision.

Fig. 4(a) shows the normalized SMSE performance versus iterations, i.e., convergence behaviour, of the ADMM-based distributed estimation scheme proposed in Section-IV for analog and quantized SN transmission scenarios. It can be seen that the distributed solution rapidly converges to the decentralized SMSE performance in very few iterations. This shows the convergence performance of the proposed distributed scheme. This also confirms the fact that the overhead of exchanging the dual variables among sensors is negligible in comparison to that of transmitting other quantities such as observation

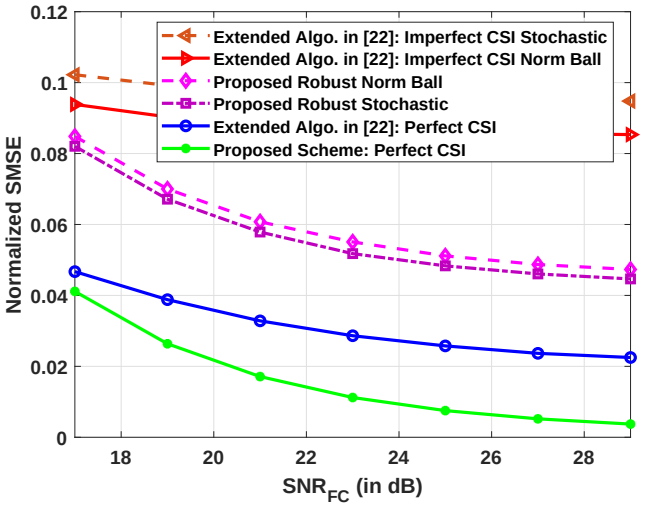


Fig. 6: MSE performance comparison with the iterative MMSE transceiver design scheme in [22].

matrices, observation noise covariances etc. for each sensor to the fusion center in order to design precoders at the fusion center and subsequently feeding back the precoders thus designed to each sensor. Fig. 4(b) depicts the convergence time of the proposed distributed scheme for different thresholds versus the number of sensors L in the network. It can be readily observed that the convergence time depends critically on the convergence threshold, while not being significantly affected by the number of sensors.

Figures 5(a) and 5(b) depict the normalized SMSE performance of the robust precoder designs developed in Section-VI for scenarios with CSI uncertainty, for the stochastic and norm ball uncertainty models, respectively. The performance of both analog as well as quantized observations is shown in the figures. It can be readily deduced that the robust precoders yield a significant performance improvement in comparison to the corresponding agnostic estimators that directly use the channel estimate. This demonstrates the suitability of the proposed framework for decentralized parameter estimation in practical scenarios in the face of CSI uncertainty. This can indeed be exploited for significantly boosting the quality of the parameter estimates computed at the fusion center.

In order to compare the MSE performance of the proposed scheme, the iterative scheme proposed in [22] has been extended to a MIMO-OFDM based WSN. Note that the model in [22] considers only a MIMO WSN system, which has otherwise no direct bearing on the schemes developed in our work. Thus, this extension is purely done for the purposes of comparison due to a dearth of directly relevant schemes in the existing literature. Their extended iterative design is run for 7 iterations and the corresponding normalized SMSE performance is plotted for comparison in Fig. 5. It can be readily observed that the proposed robust designs, for both the stochastic and norm ball CSI uncertainty models, outperform the algorithm in [22] with imperfect CSI. In addition, the proposed design with perfect CSI also outperforms its counterpart in [22]. This shows the efficacy of the proposed designs. Moreover, the proposed techniques are also non-iterative in nature, which makes them computationally tractable and hence well suited for application in WSNs.

VIII. Conclusion

This paper derived the optimal precoder designs for minimizing the SMSE of decentralized estimation of a spatially and temporally correlated parameter vector in a MIMO-OFDM WSN. The designs were initially presented for the transmission of analog sensor observations. Subsequently, a rate-distortion theory based framework has been developed for optimal quantization of the observations for minimizing the sum distortion. New precoder designs have also been developed to minimize the SMSE of parameter estimation with the transmission of quantized observations. In order to further reduce the computational complexity and overheads of precoder computation, a dual consensus ADMM-based distributed precoder design has also been derived, in which each SN designs its own precoders corresponding to the various subcarriers with the aid of minimal centralized coordination. Finally, to ensure resilience of the parameter estimate in the face of CSI uncertainty, robust precoder designs have also been derived for the transmission of both analog as well as quantized observations. Simulation results have characterized the performance of the proposed designs for various settings and also verified the various analytical propositions.

References

[1] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, 2008.

[2] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Optimized sensor collaboration for estimation of temporally correlated parameters," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6613–6626, 2016.

[3] Y. Dong, Z. Chen, J. Wang, and B. Shim, "Optimal power control for transmitting correlated sources with energy harvesting constraints," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 461–476, Jan 2018.

[4] YuanYuan Li and L. E. Parker, "A spatial-temporal imputation technique for classification with missing data in a wireless sensor network," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2008, pp. 3272–3279.

[5] S. N. Das, S. Misra, B. E. Wolfinger, and M. S. Obaidat, "Temporal-correlation-aware dynamic self-management of wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2127–2138, Dec 2016.

[6] H. Ko, S. Pack, and V. C. M. Leung, "Spatiotemporal correlation-based environmental monitoring system in energy harvesting Internet of Things (IoT)," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2958–2968, May 2019.

[7] A. Özçelikkale, T. McKelvey, and M. Viberg, "Remote estimation of correlated sources under energy harvesting constraints," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5300–5313, Aug 2018.

[8] J. Z. Sun and V. K. Goyal, "Intersensor collaboration in distributed quantization networks," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3931–3942, 2013.

[9] I. Nevat, G. W. Peters, and I. B. Collings, "Random field reconstruction with quantization in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 6020–6033, 2013.

[10] E. J. Msechu and G. B. Giannakis, "Sensor-centric data reduction for estimation with WSNs via censoring and quantization," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 400–414, 2012.

[11] J. Li and G. AlRegib, "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, 2009.

[12] M. H. Chaudhary and L. Vandendorpe, "Power constrained linear estimation in wireless sensor networks with correlated data and digital modulation," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 570–584, Feb 2012.

[13] H. Li, "Distributed adaptive quantization and estimation for wireless sensor networks," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, vol. 3, April 2007, pp. III–533–III–536.

[14] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-Part I: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, March 2006.

[15] Y. Zhou, C. Huang, T. Jiang, and S. Cui, "Wireless sensor networks and the Internet of Things: Optimal estimation with nonuniform quantization and bandwidth allocation," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3568–3574, Oct 2013.

[16] A. Sani and A. Vosoughi, "Distributed vector estimation for power- and bandwidth-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3879–3894, 2016.

[17] M. Leinonen, M. Codreanu, and M. Juntti, "Distributed distortion-rate optimized compressed sensing in wireless sensor networks," *IEEE Transactions on Communications*, vol. 66, no. 4, pp. 1609–1623, 2018.

[18] J. Zhu, R. S. Blum, X. Lin, and Y. Gu, "Robust transmit beamforming for parameter estimation using distributed sensors," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1329–1332, July 2016.

[19] N. K. D. Venkategowda, B. B. Narayana, and A. K. Jagannatham, "Precoding for robust decentralized estimation in coherent-MAC-based wireless sensor networks," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 240–244, Feb 2017.

[20] Y. Liu, J. Li, and H. Wang, "Robust linear beamforming in wireless sensor networks," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4450–4463, 2019.

[21] H. Rostami and A. Falahati, "Precoder design for decentralised estimation over MIMO-WSN based on stochastic models," *IET Communications*, vol. 12, no. 6, pp. 736–742, 2018.

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

[23] Y. Liu, J. Li, and X. Lu, "Joint transceiver design for linear MMSE data fusion in coherent MAC wireless sensor networks," *Information Fusion*, vol. 37, pp. 37–49, 2017.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. New Jersey: Prentice-Hall Inc, 1993.

[25] N. K. Venkategowda and A. K. Jagannatham, "Optimal minimum variance distortionless precoding (MVDP) for decentralized estimation in MIMO wireless sensor networks," *IEEE Signal Processing Letters*, vol. 22, no. 6, pp. 696–700, 2015.

[26] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1684–1696, 2005.

[27] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[28] K. P. Rajput, M. F. Ahmed, N. K. Venkategowda, A. K. Jagannatham, G. Sharma, and L. Hanzo, "Technical report: Robust decentralized and distributed estimation of a correlated parameter vector in MIMO-OFDM wireless sensor networks." IIT Kanpur, Tech. Rep., 2021. [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_MIMO_OFDM_WSN_21.pdf.

[29] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[30] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[31] B. Widrow and I. Kollár, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge University Press, 2008.

[32] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[33] G. Mateos, J. A. Bazerque, and G. B. Giannakis, "Distributed sparse linear regression," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5262–5276, 2010.

[34] T. Lin, S. Ma, and S. Zhang, "On the global linear convergence of the ADMM with multiblock variables," *SIAM Journal on Optimization*, vol. 25, no. 3, pp. 1478–1497, 2015.

[35] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5473–5484, 2011.

[36] Y. Guo and B. C. Levy, "Worst-case MSE precoder design for imperfectly known MIMO communications channels," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2918–2930, Aug 2005.

[37] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends® in Communications and Information Theory*, vol. 3, no. 4-5, pp. 331–551, 2007.