Robust Hybrid Transceiver Designs for Linear Decentralized Estimation in mmWave MIMO IoT Networks in the Face of Imperfect CSI

Priyanka Maity, *Graduate Student Member, IEEE,* Kunwar Pritiraj Rajput, *Member, IEEE,* Suraj Srivastava, *Member, IEEE,* Naveen K. D. Venkategowda, *Member, IEEE,* Aditya K. Jagannatham, *Senior Member, IEEE* and Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Hybrid transceivers are designed for linear decentralized estimation (LDE) in a mmWave multiple-input multiple-output (MIMO) IoT network (IoTNe). For a noiseless fusion center (FC), it is demonstrated that the MSE performance is determined by the number of RF chains used at each IoT node (IoTNo). Next, the minimum-MSE RF transmit precoders (TPCs) and receiver combiner (RC) matrices are designed for this setup using the dominant array response vectors, and subsequently, a closed-form expression is obtained for the baseband (BB) TPC at each IoTNo using Cauchy's interlacing theorem. For a realistic noisy FC, it is shown that the resultant mean squared error (MSE) minimization problem is non-convex. To address this challenge, a block-coordinate descent-based iterative scheme is proposed to obtain the fully digital TPC and RC matrices followed by the simultaneous orthogonal matching pursuit (SOMP) technique for decomposing the fully-digital transceiver into its corresponding RF and BB components. A theoretical proof of the convergence is also presented for the proposed iterative design procedure. Furthermore, robust hybrid transceiver designs are also derived for a practical scenario in the face of channel state information (CSI) uncertainty. The centralized MMSE lower bound has also been derived that benchmarks the performance of the proposed LDE schemes. Finally, our numerical results characterize the performance of the proposed transceivers as well as corroborate our various analytical propositions.

Index Terms—Hybrid transceiver design, Internet of things (IoT), mmWave communication, linear decentralized estimation, wireless sensor networks.

I. INTRODUCTION

Internet of Things networks (IoTNes) constitute a key component of the Internet of Things (IoT) [1], which has the potential of supporting cutting-edge applications spanning diverse areas such as precision agriculture [2], remote healthcare [3], environmental monitoring [4], and smart cities [5] etc.. IoTNes are comprised of miniature IoT nodes (IoTNos) dispersed over a wide geographical area for the transmission of their suitably pre-processed observations to a fusion centre (FC) over a wireless channel for further processing and estimation. The ever-increasing mobile device density coupled with the large number of data-intensive applications has led to the impending spectrum crunch in the sub 6 GHz band. This is further aggrevated by the deployment of a large number of IoT systems, each comprised of a multitude of IoTNos.

The abundant spectrum in the mmWave (30-300 GHz) band, which has hitherto only been partially explored, has the potential of supporting high data rates coupled with massive connectivity in the IoTNes. However, communication in the mmWave regime encounters several obstacles such as severe signal blockage and debilitating propagation losses [6]. Fortunately, the low wavelength of mmWave signals renders it possible to integrate a large number of antennas into the otherwise small wireless devices. This in turn facilitates beamforming, which is capable of overcoming the aforementioned challenges in mmWave communication. The higher cost and power consumption due to the large number of RF chains required to support the increased number of antennas in a conventional architecture can be circumvented by employing innovative hybrid transceiver architectures [7]–[10]. Such a transceiver jointly performs analog TPC/ RC in the RF domain, together with digital TPC/ RC in the baseband, which minimizes the number of RF chains required.

Coming next to parameter estimation in IoTNes, linear decentralized estimation (LDE), constitutes an excellent strategy [11], [12] where multiple IoTNos precode their observations and transmit them to the FC using a multiple access channel (MAC) for receive combining. This in turn necessitates the design of optimal TPCs and RCs that minimize the MSE of parameter estimation at the FC. An overview of the various related works in this area is presented next.

A. Literature Review

The most popular model of LDE has been developed in the seminal works [11] and [12]. The authors of [13] have proposed novel zero-forcing based TPC designs for sub-6 GHz IoTNes for the LDE of a vector parameter. An

P. Maity, S. Srivastava and A. K. Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology, Kanpur, Kanpur, 208016, India (e-mail: pmaity@iitk.ac.in, ssrivast@iitk.ac.in, adityaj@iitk.ac.in.)

K. P. Rajput was with the Department of Electrical Engineering, Indian Institute of Technology, Kanpur, Kanpur, 208016, India. He is now with the interdisciplinary center of security reliability and trust (SnT), University of Luxembourg, 4365, Luxembourg (e-mail: kunwar.rajput@uni.lu)

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)

Feature	[6]	[13]	[14]	[21],[22]	[24]	[28]	[29]	[30]	[40]	**Our work**
mmWave WSN						✓	✓	✓	✓	✓
Vector parameter estimation		✓	✓	✓	✓	✓	✓			✓
Per sensor power constraint		✓		✓		✓	✓	✓	✓	✓
Coherent MAC		✓	✓	✓	✓		✓	✓	✓	✓
Hybrid transceiver design	✓					✓	✓		✓	✓
Noiseless FC										✓
Gaussian CSI uncertainty									✓	✓

interesting aspect of the scheme described is that it does not require any post-processing at the FC. The treatise [14] developed majorization theory-based closed-form transceiver designs considering scenarios of both noiseless as well as noisy IoTNo observations. Zhou *et al.* [15] described a learning-based iterative mixed algorithm for the detection-estimation of a scalar parameter, while also considering IoTNo defects, thus making it suitable for practical applications. The authors of [16] proposed an edge computing based scheme for decentralized sensing of a spatially correlated process. A subset of IoTNos is efficiently selected by the algorithm, which strikes an attractive trade-off between energy efficiency and sensing quality.

Leong *et al.* [17] conceived optimal power allocation schemes for estimating a time varying scalar parameter. The TPCs and RCs of [17] are designed with the goal of minimizing the resultant MSE at the FC. Furthermore, the authors of [18] analyse the MSE of estimating and tracking a time-varying parameter. The FC of [18] employs Kalman filtering to track the parameter of interest. Akhtar *et al.* [19] described TPC designs that sequentially estimate a time-varying parameter vector. Their design is iterative in nature and it is capable of operating in the face of time-varying channels while obeying specific transmit power constraints. The authors of [20] present a framework for sequentially estimating a parameter of interest, wherein only a subset of IoTNos transmit information to the FC based on prior communication among the IoTNos. A variety of schemes were proposed in [21]–[23] for LDE in a wireless powered IoTNe, where the IoTNos run on the energy harvested from the RF signals transmitted by energy access points. A set of LDE schemes have been proposed in [24]–[26] for cutting-edge massive MIMO IoTNes, which exploit the asymptotic beamforming properties of large antenna arrays. The authors of [27] propose a secure diffusion-based least-mean squares algorithm for distributed parameter estimation in the presence of an eavesdropper. However none of the above impressive contributions are directly applicable to mmWave MIMO systems relying on the bespoke signal processing employed in a hybrid transceiver. The literature specifically related to beamforming in mmWave IoTNes is reviewed next.

Liu *et al.* [28] developed novel hybrid transceiver designs for maximizing the spectral efficiency of an orthogonal MAC-based mmWave IoTNe. The cutting-edge scheme described initially derives the fully digital transceivers followed by decomposing them into their baseband and RF components. An iterative algorithm is proposed for obtaining the closed-form expressions of baseband TPCs. Then a pair of procedures are proposed for the analog TPC/ RC using the popular alternating direction method of multipliers (ADMM) and the steepest descent principles, respectively. However, a drawback of the system is that it relies on an orthogonal MAC, which is not bandwidth efficient as it requires dedicated channels between the constituent IoTNos and the FC. It must be noted at this juncture that the orthogonal MAC concept considerably differs from a coherent MAC, wherein the signal received at the FC constituted by a superposition of the signals transmitted by all the IoTNos. To further develop the latter system, the authors of [29] presented both centralized and distributed LMMSE transceiver designs for a mmWave MIMO IoTNe. A state-of-the-art hybrid transmit beamforming algorithm is proposed in [30] for EAPs and simultaneous wireless information and power transfer (SWIPT) in the support of IoT receivers . However, a limiting aspect of the above contributions, which restricts their applicability in practice, is that they assume perfect knowledge of the channel between the FC and the constituent IoTNos. Needless to say, this is hard to achieve in practice owing to several limitations, such as the limited pilot overhead and the finite precision error-prone feedback. Papers that address this critical issue of robust TPC/ RC design, in both sub-6 GHz and mmWave systems, are discussed next.

The authors of [31] and [32] propose robust TPC/ RC designs for LDE in a IoTNe, for sparse and temporally correlated parameter vectors, respectively, relying on imperfect CSI. Venkategowda *et al.* [33] designed TPCs robust to imperfect CSI for estimating a scalar parameter. Their robust designs were shown to yield an improved performance in comparison to non-robust designs that are oblivious to the CSI imperfections. The treatise [34] proposes robust closed-form TPC matrices for parameter vector estimation. The authors of [33], [34] propose robust designs employing a zero-forcing (ZF) constraint at the FC to make the post-processing simple. However, the ZF equalizer leads to noise amplification at the receiver, which degrades the signal-to-noise ratio (SNR). Moreover, they also impose a total power constraint on the MIMO IoTNe. In the context of mmWave MIMO systems, minimum-MSE robust transceivers are developed in [35] for a mmWave multi-user system. Explicitly, the authors proposed a non-linear hybrid transceiver design based on Tomlinson/Harashima precoding. Jiang *et al.* [36] have proposed robust transceiver designs for a half-duplex mmWave relay system, where the RF TPC/ RC are designed based on the strongest eigenmode of the mmWave channel. An iterative algorithm based on mutual information maximization has also been developed in their treatise for designing the baseband
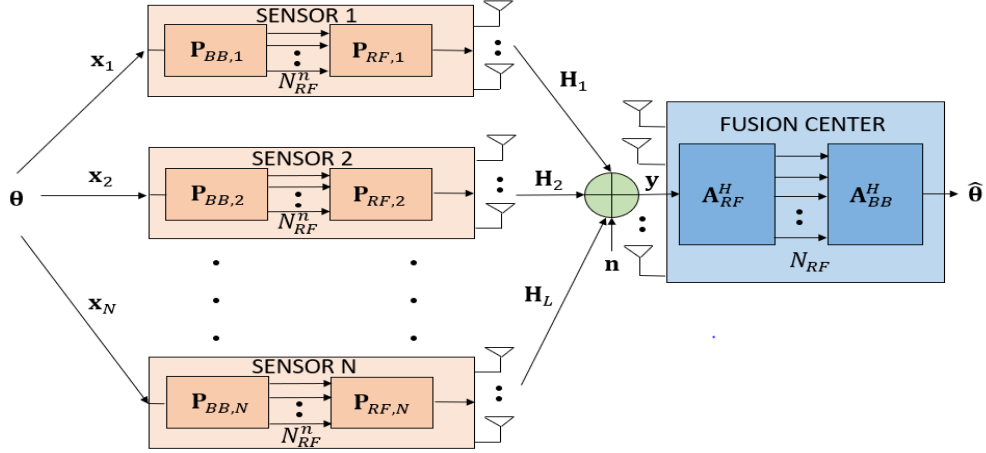
Fig. 1. Schematic diagram of hybrid MIMO signal processing for vector parameter estimation in mmWave MIMO IoTNe over a coherent MAC.

TPCs. Cai *et al.* [37] have developed a robust transceiver design based on the principle of sum-rate maximization for full-duplex (FD) MIMO relay-aided multi-user mmWave systems. Furthermore, Zhao *et al.* [38] have proposed a robust transceiver design in a full-duplex (FD) mmWave multi-cell network. In [37], [38], iterative algorithms are developed based on the classic penalty dual decomposition (PDD) technique for a scenario subject to imperfect CSI. The authors of [39] characterized the performance of a robust hybrid TPC scheme in a downlink mmWave IoTNe relying on single antenna IoTNos and a multi-antenna FC. Furthermore, a noteworthy aspect of the design procedure of [39] is that it maximizes the minimum secrecy at each IoTNo, in the presence of an eavesdropper (ED), which guarantees privacy of the transmitted information. Moreover, in their work [39], the CSI between the ED and the FC is assumed to be imperfectly known, whereas that between each IoTNo and the FC is assumed to be perfectly known. In the pioneering paper in [40], the authors describe a strategy to design the hybrid TPC in a mmWave-based IoTNe with multiple-antenna IoTNos and a single antenna FC. Hybrid baseband and RF TPCs were designed that attained the minimum-MSE for LDE of the parameter. Robust hybrid TPCs were also designed accounting for the CSI uncertainty. However, the results and framework of [40] are restricted to scalar parameter estimation due to having a single antenna FC in the system. To the best of our knowledge, there is no literature on the design of the robust hybrid transceivers for the LDE of a parameter vector in a coherent MAC-based mmWave IoTNe, accounting also for realistic CSI uncertainty. We fill the knowledge-gap in the present work. Table I presents a comparison of the contributions of this paper against those reviewed above. The novel contributions of this work are boldly contrasted to the literature at a glance in Table I and are also elaborated below in a point-wise fashion.

- To begin with, hybrid transceiver designs are presented for MSE minimization at the FC. The RF TPCs and RCs are determined using the dominant array vectors of the channel. Subsequently, the BB RC is derived using the LMMSE rule, while the BB TPC corresponding to each IoTNo is derived in closed-form using Cauchy's interlacing theorem [11].

- For a noisy FC, the MSE minimization problem of hybrid transceiver designs is observed to be non-convex. Hence, to render the optimization problem tractable, the non-convex constant-magnitude constraints are eliminated by initially considering the design of a fully-digital transceiver. The non-convexity of the optimization objective is handled by exploiting the block-coordinate descent (BCD)-based iterative design procedure, which ultimately yields a closed-form expression for the optimal fully-digital transceiver in each iteration.

- Upon convergence, the simultaneous orthogonal matching pursuit (SOMP) technique is harnessed for decomposing it into its RF and BB components.

- To account for CSI uncertainty in real-world systems, robust transceiver designs are developed for a noisy FC scenario that mitigate the performance degradation imposed by imperfect CSI.

- The centralised MMSE bound has been derived that benchmarks the performance of the proposed LDE schemes. Our numerical results demonstrate the effectiveness of the proposed hybrid transceiver designs, and also validate the improved performance of our robust transceiver design over its non-robust counterpart, which is oblivious of the CSI imperfection.

B. Outline of the paper

The rest of the manuscript is organized as follows. Section-II presents our LDE system model in a mmWave MIMO IoTNe. Section-III describes our hybrid transceiver relying on MSE minimization for a noiseless FC, while Section-IV describes its counterpart designed for a general system relying on a noisy FC. Section-V describes our robust hybrid transceiver designs incorporating CSI uncertainty. Finally, Section-VI and Section-VII present our numerical results and our conclusions, respectively.

Notation: The following notation is used throughout this paper. Lowercase ($\mathbf{c}$) and uppercase ($\mathbf{C}$) letters denote vectors and matrices, respectively. The operators $(.)^T$, $(.)^H$ and $|.|$ denote the transpose, Hermitian and magnitude. Tr$(.)$ and $\mathbb{E}[.]$ represent the trace and statistical expectation operators, respectively. For a matrix $\mathbf{C}$, its (i,j)th element is denoted as $\mathbf{C}(i,j)$, while its column and row spaces are denoted by $\mathcal{C}(\mathbf{C})$ and $\mathcal{R}(\mathbf{C})$, respectively. The Frobenius and l_0 norms are represented as $||.||_F$ and $||.||_0$. Finally, $\mathbf{c} = \text{diag}(\mathbf{C})$ denotes a vector comprising of the principal diagonal elements of the matrix $\mathbf{C}$.

II. mmWave WSN System and Channel Model

A mmWave-based MIMO IoTNe supporting N IoTNos and a FC is considered as shown in Fig. 1. Each IoTNo employs N_t transmit antennas (TAs) and N_{RF}^n RF chains, while the FC has N_r receive antennas (RAs) and N_{RF} RF chains, where $N_{\text{RF}}^n, N_{\text{RF}} << \min(N_t, N_r)$. The IoTNos are sensing/monitoring a common parameter vector $\boldsymbol{\theta} \in \mathbb{C}^{q \times 1}$ comprised of q different physical quantities that have to be estimated. The observation vector of IoTNo n, represented by $\mathbf{x}_n \in \mathbb{C}^{l \times 1}$ can be modeled as

$$\mathbf{x}_n = \mathbf{C}_n \boldsymbol{\theta} + \mathbf{v}_n, \tag{1}$$

where $\mathbf{C}_n \in \mathbb{C}^{l \times q}$ and $\mathbf{v}_n \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_n) \in \mathbb{C}^{l \times 1}$ represent the observation matrix and the observation noise corresponding to the nth IoTNo, respectively. Each node employs the hybrid TPC comprised of the cascaded BB and RF TPC matrices $\mathbf{P}_{\text{BB},n} \in \mathbb{C}^{N_{\text{RF}}^n \times l}$ and $\mathbf{P}_{\text{RF},n} \in \mathbb{C}^{N_t \times N_{\text{RF}}^n}$, respectively. Subsequently, each IoTNo transmits its precoded observations to the FC over a coherent multiple access channel (MAC). Therefore, the signal $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ received at the FC can be modeled as

$$\begin{aligned}
\mathbf{y} &= \sum_{n=1}^{N} \mathbf{H}_n \mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{x}_n + \mathbf{w} \\
&= \sum_{n=1}^{N} \mathbf{H}_n \mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{C}_n \boldsymbol{\theta} + \sum_{n=1}^{N} \mathbf{H}_n \mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{v}_n + \mathbf{w} \\
&= \mathbf{H} \mathbf{P}_{\text{RF}} \mathbf{P}_{\text{BB}} \mathbf{C} \boldsymbol{\theta} + \mathbf{H} \mathbf{P}_{\text{RF}} \mathbf{P}_{\text{BB}} \mathbf{v} + \mathbf{w},
\end{aligned} \tag{2}$$

where the matrix $\mathbf{H}_n \in \mathbb{C}^{N_r \times N_t}$ represents the MIMO channel matrix between each IoTNo n and the FC, whereas the vector $\mathbf{w} \in \mathbb{C}^{N_r \times 1}$ represents the FC noise and follows the distribution $\mathcal{CN}(\mathbf{0}, \mathbf{R}_w)$. The stacked observation matrix $\mathbf{C} \in \mathbb{C}^{Nl \times q}$, concatenated channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times NN_t}$, the overall block-diagonal BB TPC $\mathbf{P}_{\text{BB}} \in \mathbb{C}^{NN_{\text{RF}}^n \times Nl}$ and RF TPC $\mathbf{P}_{\text{RF}} \in \mathbb{C}^{NN_t \times NN_{\text{RF}}^n}$ are defined as

$$\mathbf{C} = [\mathbf{C}_1^T, \mathbf{C}_2^T, \dots, \mathbf{C}_N^T]^T, \tag{3}$$
$$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_N], \tag{4}$$
$$\mathbf{P}_{\text{BB}} = \text{diag}(\mathbf{P}_{\text{BB},1}, \dots, \mathbf{P}_{\text{BB},N}), \tag{5}$$
$$\mathbf{P}_{\text{RF}} = \text{diag}(\mathbf{P}_{\text{RF},1}, \dots, \mathbf{P}_{\text{RF},N}). \tag{6}$$

At the FC, the hybrid RC operation is performed using the cascaded RF and BB combining matrices $\mathbf{A}_{\text{RF}} \in \mathbb{C}^{N_r \times N_{\text{RF}}}$ and $\mathbf{A}_{\text{BB}} \in \mathbb{C}^{N_{\text{RF}} \times q}$, respectively, to yield the estimate $\widehat{\boldsymbol{\theta}}$ as

$$\widehat{\boldsymbol{\theta}} = \mathbf{A}_{\text{RF}}^H \mathbf{A}_{\text{BB}}^H \mathbf{y}. \tag{7}$$

Since, the hybrid MIMO architecture employs the constant magnitude RF TPC and RC matrices, their elements can be constrained as follows

$$|\mathbf{P}_{\text{RF},n}(i,j)| = \frac{1}{\sqrt{N_t}}, \ 1 \le n \le N, \ \forall \, i, j,$$
$$|\mathbf{A}_{\text{RF}}(i,j)| = \frac{1}{\sqrt{N_r}}, \ \forall \, i, j. \tag{8}$$

Furthermore, the MIMO channel matrix $\mathbf{H}_n$ is defined as [6], [41]

$$\mathbf{H}_n = \sqrt{\frac{N_r N_t}{K}} \sum_{k=1}^{K} \alpha_{k,n} \mathbf{a}_{\text{FC}}(\phi_k) \mathbf{a}_{\text{s}}^H(\theta_{k,n}), \tag{9}$$

where the 3-tuple $(\alpha_{k,n}, \phi_k, \theta_{k,n})$ represents the complex gain $\alpha_{k,n}$, angle of arrival (AoA) ϕ_k at the FC, and angle of departure (AoD) $\theta_{k,n}$ at the nth IoTNe associated with the kth cluster and K denotes the total number of clusters. The vectors $\mathbf{a}_{\text{FC}}(\phi_k) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_{\text{s}}(\theta_{k,n}) \in \mathbb{C}^{N_t \times 1}$ represent the array response vectors at the FC and the nth IoTNo, respectively, corresponding to the kth cluster, which are defined as

$$\mathbf{a}_{\text{FC}}(\phi_k) = \frac{1}{\sqrt{N_R}} \left[1, e^{-j\tilde{\phi}_k}, \cdots, e^{-j(r-1)\tilde{\phi}_k}\right]^T, \tag{10}$$
$$\mathbf{a}_{\text{s}}(\theta_{k,i}) = \frac{1}{\sqrt{N_T}} \left[1, e^{-j\tilde{\theta}_{k,n}}, \cdots, e^{-j(t-1)\tilde{\theta}_{k,n}}\right]^T, \tag{11}$$

where $\tilde{\phi}_k = \frac{2\pi}{\lambda} d_R \cos \phi_k$ and $\tilde{\theta}_{k,n} = \frac{2\pi}{\lambda} d_T \cos \theta_{k,n}$. The quantities λ, d_R, and d_T denote the carrier's wavelength and inter-antenna spacings at the FC and each IoTNo, respectively. The mmWave MIMO channel $\mathbf{H}_n$ can be equivalently expressed in the compact form of

$$\mathbf{H}_n = \mathbf{A}_{\text{FC}} \mathbf{D}_n \mathbf{A}_{\text{s},n}^H, \tag{12}$$

where $\mathbf{A}_{\text{FC}} = [\mathbf{a}_{\text{FC}}(\phi_1), \cdots, \mathbf{a}_{\text{FC}}(\phi_K)] \in \mathbb{C}^{N_r \times K}$, $\mathbf{A}_{\text{s},n} = [\mathbf{a}_{\text{s}}(\theta_{1,n}), \cdots, \mathbf{a}_{\text{s}}(\theta_{K,n})] \in \mathbb{C}^{N_t \times K}$ are termed the array response matrices corresponding to the FC and the nth IoTNo, respectively, and $\mathbf{D}_n = \sqrt{\frac{N_r N_t}{K}} \text{diag}(\alpha_{1,n}, \cdots, \alpha_{K,n})$ comprises of the path-gains. We assume that the FC has perfect knowledge of the array response matrices $\mathbf{A}_{\text{FC}}$, $\mathbf{A}_{\text{s},n}$ and the gain matrix $\mathbf{D}_n$. The next section describes our novel hybrid transceiver designed for the MSE minimization of a noiseless FC.

III. Hybrid Transceiver Design for a Noiseless FC

The purpose of this study is to determine the minimum number of RF chains required at each sensor to achieve the minimum MSE. We would like to emphasize that, in this scenario, the observation noise dominates the thermal noise at the receiver, which renders the latter negligible. Thus, the measurements of each sensor are corrupted by observation noise, which necessitates the design of the hybrid TPCs. On account of a noiseless FC, the received vector $\mathbf{y}$ can be obtained by substituting $\mathbf{w} = \mathbf{0}$ in (2). The estimate $\widehat{\boldsymbol{\theta}} \in \mathbb{C}^{m \times 1}$ at the output of the LMMSE RC for a noiseless coherent MAC is given by

$$\begin{aligned}
\widehat{\boldsymbol{\theta}} &= \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \mathbf{y} \\
&= \mathbf{A}_{\text{BB}}^H \underbrace{\mathbf{A}_{\text{RF}}^H \mathbf{H} \mathbf{P}_{\text{RF}}}_{\widetilde{\mathbf{H}}} \mathbf{P}_{\text{BB}} \mathbf{C} \boldsymbol{\theta} + \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \mathbf{H} \mathbf{P}_{\text{RF}} \mathbf{P}_{\text{BB}} \mathbf{v}. \tag{13}
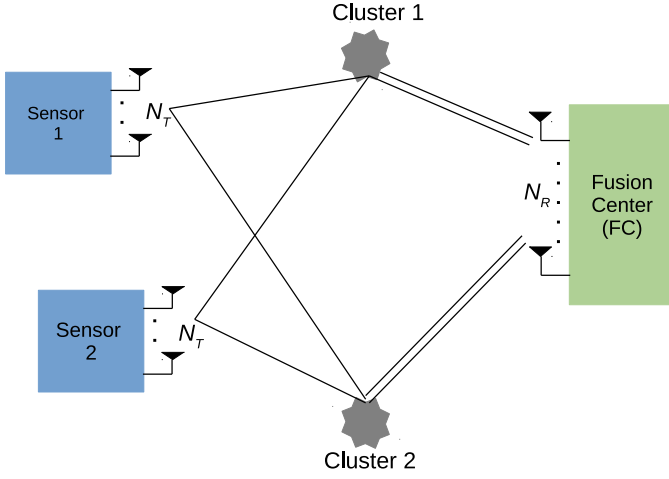\end{aligned}$$

Fig. 2. Spatial mmWave MIMO channel model

Furthermore, the resultant MSE at the FC, which is defined as $\eta\left(\mathbf{P}_{\mathrm{RF}}, \mathbf{P}_{\mathrm{BB}}, \mathbf{A}_{\mathrm{RF}}\right) = \mathbb{E}\left[\left\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right\|^2\right]$, can be expressed as

$$\eta\left(\mathbf{P}_{\mathrm{RF}}, \mathbf{P}_{\mathrm{BB}}, \mathbf{A}_{\mathrm{RF}}\right) = \mathrm{Tr}\left[\sigma_{\theta}^2 \mathbf{I}_m + \mathbf{C}^H \mathbf{P}_{\mathrm{BB}}^H \mathbf{P}_{\mathrm{RF}}^H \bar{\mathbf{H}}^H \left(\mathbf{A}_{\mathrm{RF}} \bar{\mathbf{H}}\right.\right.$$

$$\left.\left. \mathbf{P}_{\mathrm{RF}} \mathbf{P}_{\mathrm{BB}} \mathbf{R}_v \mathbf{P}_{\mathrm{BB}}^H \mathbf{P}_{\mathrm{RF}}^H \bar{\mathbf{H}}^H \mathbf{A}_{\mathrm{RF}}\right)^{-1} \mathbf{A}_{\mathrm{RF}}^H \bar{\mathbf{H}} \mathbf{P}_{\mathrm{RF}} \mathbf{P}_{\mathrm{BB}} \mathbf{C}\right]^{-1}. \quad (14)$$

Hence, we can formulate the optimization problem of our hybrid transceiver with an objective to minimize the MSE in (14) as follows

$$\underset{\mathbf{P}_{\mathrm{RF}}, \mathbf{P}_{\mathrm{BB}}, \mathbf{A}_{\mathrm{RF}}}{\text{minimize}} \quad \eta\left(\mathbf{P}_{\mathrm{RF}}, \mathbf{P}_{\mathrm{BB}}, \mathbf{A}_{\mathrm{RF}}\right)$$

$$\text{subject to } |\mathbf{P}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{N_t}}, 1 \le n \le N, \forall\, i,j,$$

$$|\mathbf{A}_{\mathrm{RF}}(i,j)| = \frac{1}{\sqrt{N_r}}, \forall\, i,j. \quad (15)$$

The constant magnitude of the elements in the RF TPC $\mathbf{P}_{\mathrm{RF},n}$ and RC $\mathbf{A}_{\mathrm{RF}}$ results in the non-convexity of the above optimization problem, and makes it intractable. In this context, suitable RF TPCs $\mathbf{P}_{\mathrm{RF},n}$ and RC $\mathbf{A}_{\mathrm{RF}}$ can be obtained by using the procedure outlined in the next subsection, which is followed by the design of the optimal BB TPC $\mathbf{P}_{\mathrm{BB}}$ for each IoTNo n. This section considers the variance of each parameter to be $\sigma_{\theta}^2 = 1$ and the observation noise covariance to be $\mathbf{R}_n = \mathbf{I}_l$.

A. Design of RF TPCs $\mathbf{P}_{RF,n}$ and RC $\mathbf{A}_{RF}$

Since the columns of the matrices $\mathbf{A}_{\mathrm{FC}}$ and $\mathbf{A}_{\mathrm{s},n}$ are the array response vectors corresponding to the AoAs at the FC and AoDs at the nth IoTNo, the N_{RF} columns of the RF RC $\mathbf{A}_{\mathrm{RF}}$ and TPC $\mathbf{P}_{\mathrm{RF},n}$ can be set using the N_{RF} dominant array response vectors, respectively. In order to design the RF TPC for the nth IoTNo, we organise the complex gains $\alpha_{k,n}$ in decreasing order of magnitude so that $|\alpha_{k_1,n}| \ge |\alpha_{k_2,n}| \ge \cdots \ge |\alpha_{k_K,n}|$. The RF TPC $\mathbf{P}_{\mathrm{RF},n}$ for the nth IoTNo can

be designed by selecting the N_{RF} columns denoted by the set $\mathcal{K}_n = \{k_1, k_2, \ldots, k_{N_{\mathrm{RF}}^n}\}$, which contains the indices of the N_{RF} dominant transmit array response vectors. This is mathematically represented as

$$\mathbf{P}_{\mathrm{RF},n} = \mathbf{A}_{\mathrm{s},n}\left(\,:\,,\,\mathcal{K}_n\,\right). \quad (16)$$

The design of the RF RC $\mathbf{A}_{\mathrm{RF}}$ is described next. Let the quantity $\alpha_k = \sum_{n=1}^{N} |\alpha_{k,n}|$ denote the sum of the magnitudes of the complex path-gains of each IoTNo corresponding to the kth cluster. One can now arrange the quantities α_l in decreasing order of their magnitude, so that $|\alpha_{k_1}| \ge |\alpha_{k_2}| \ge \cdots \ge |\alpha_{k_N}|$. The RF RC $\mathbf{A}_{\mathrm{RF}}$ is obtained as

$$\mathbf{A}_{\mathrm{RF}} = \mathbf{A}_{\mathrm{FC}}\left(\,:\,,\,\mathcal{K}\,\right), \quad (17)$$

where we have $\mathcal{K} = \{k_1, k_2, \ldots, k_{N_{\mathrm{RF}}}\}$. Employing these settings of the RF TPC $\mathbf{P}_{\mathrm{RF},n}$ and RC $\mathbf{A}_{\mathrm{RF}}$, the design of the optimal BB TPC $\mathbf{P}_{\mathrm{BB},n}$ for each IoTNo n is described next.

B. Baseband TPC $\mathbf{P}_{BB,n}$ design

Upon employing the LMMSE BB RC $\mathbf{A}_{\mathrm{BB}}$ at the FC as shown in (13), the resultant error covariance matrix $\mathbf{E}$ is given by

$$\mathbf{E} = \left[\mathbf{I}_q + \mathbf{C}^H \widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1} \widetilde{\mathbf{H}} \mathbf{C}\right]^{-1}, \quad (18)$$

where we have $\widetilde{\mathbf{H}} = \bar{\mathbf{H}} \mathbf{P}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}}^n \times Nl}$. Furthermore, assuming that the effective channel $\bar{\mathbf{H}}_n = \mathbf{A}_{\mathrm{RF}}^H \mathbf{H}_n \mathbf{P}_{\mathrm{RF},n} \in \mathbb{C}^{N_{\mathrm{RF}}^n \times l}$ between each IoTNo n and the FC is invertible, solving the optimization problem for $\mathbf{P}_{\mathrm{BB},n}$ is equivalent to solving it for $\widetilde{\mathbf{H}}_n = \bar{\mathbf{H}}_n \mathbf{P}_{\mathrm{BB},n}$. As a result, the optimization problem in (15) can be restated as

$$\underset{\{\widetilde{\mathbf{H}}_n\}_{n=1}^{N}}{\text{minimize}} \; \mathrm{Tr}\left(\mathbf{E}\right)$$

$$\text{subject to } \mathbf{E}^{-1} = \left[\mathbf{I}_q + \mathbf{C}^H \widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1} \widetilde{\mathbf{H}} \mathbf{C}\right]. \quad (19)$$

The above problem can be solved using Cauchy's interlacing theorem which is described next in detail. Let us assume that the singular value decomposition (SVD) of $\widetilde{\mathbf{H}}$ is given by

$$\widetilde{\mathbf{H}} = \mathbf{U}_{\tilde{H}} \boldsymbol{\Sigma}_{\tilde{H}} \mathbf{V}_{\tilde{H}}^H, \quad (20)$$

where $\mathbf{U}_{\tilde{H}} \in \mathbb{C}^{N_{\mathrm{RF}}^n \times N_{\mathrm{RF}}^n}$, $\mathbf{V}_{\tilde{H}} \in \mathbb{C}^{Nl \times Nl}$ are unitary matrices and $\boldsymbol{\Sigma}_{\tilde{H}} \in \mathbb{C}^{N_{\mathrm{RF}}^n \times Nl}$ is the matrix containing singular values on its principal diagonal. Now, substituting the SVD of $\widetilde{\mathbf{H}}$ into $\widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1} \widetilde{\mathbf{H}}$, one obtains

$$\widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H\right)^{-1} \widetilde{\mathbf{H}} = \mathbf{V}_{\tilde{H}} \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}_{\tilde{H}}^H, \quad (21)$$

where we have $r = \text{rank}\left(\widetilde{\mathbf{H}}\right)$. Upon substituting the SVD of $\mathbf{C}$ as $\mathbf{C} = \mathbf{U}_C \mathbf{\Sigma}_C \mathbf{V}_C^H$ into $\mathbf{V}_C^H \mathbf{E}^{-1} \mathbf{V}_C$, we arrive at:

$$
\begin{aligned}
\mathbf{V}_C^H \mathbf{E}^{-1} \mathbf{V}_C &= \mathbf{V}_C^H \left(\mathbf{I}_q + \mathbf{C}^H \widetilde{\mathbf{H}}^H \left(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H \right)^{-1} \widetilde{\mathbf{H}} \mathbf{C} \right) \mathbf{V}_C \\
&= \mathbf{I}_q + \mathbf{\Sigma}_C^H \mathbf{U}_C^H \mathbf{V}_{\tilde{H}} \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}_{\tilde{H}}^H \mathbf{U}_C \mathbf{\Sigma}_C \\
&= \mathbf{I}_q + \mathbf{\Sigma}_C^H \mathbf{Q}^H \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q} \mathbf{\Sigma}_C, \quad (22)
\end{aligned}
$$

where the unitary matrix obeys $\mathbf{Q} = \mathbf{V}_{\tilde{H}}^H \mathbf{U}_C \in \mathbb{C}^{Nl \times Nl}$. Let $\lambda_k\left(\mathbf{X}\right)$ denote the kth smallest eigenvalue of any symmetric matrix $\mathbf{X}$. Then it follows from (22) that

$$
\begin{aligned}
\lambda_k\left(\mathbf{E}^{-1}\right) &= \lambda_k\left(\mathbf{V}_C^H \mathbf{E}^{-1} \mathbf{V}_C\right) \\
&= \lambda_k\left(\mathbf{I}_q + \mathbf{\Sigma}_C^H \mathbf{Q}^H \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q} \mathbf{\Sigma}_C\right) \\
&= \lambda_k\left(\mathbf{I}_{Nl} + \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right) \\
&= \lambda_k\left(\mathbf{I}_r + \left(\mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H\right)_{(1:r,1:r)}\right), \quad (23)
\end{aligned}
$$

where $\left(\mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H\right)_{(1:r,1:r)}$ represents the upper $r \times r$ block of the matrix $\mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H$. By exploiting Cauchy's interlacing theorem for the eigenvalues of a subblock matrix, we have

$$
\lambda_k\left(\left(\mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H\right)_{(1:r,1:r)}\right) \leq \lambda_k\left(\mathbf{\Sigma}_C \mathbf{\Sigma}_C^H\right) = \lambda_k\left(\mathbf{CC}^H\right),
$$
$$(24)$$

for $1 \leq k \leq r$. Therefore, upon using (23) and (24), we may conclude that

$$
\begin{aligned}
\lambda_k\left(\mathbf{E}^{-1}\right) &= \lambda_k\left(\mathbf{I}_r + \left(\mathbf{Q} \mathbf{\Sigma}_C \mathbf{\Sigma}_C^H \mathbf{Q}^H\right)_{(1:r,1:r)}\right) \\
&\leq 1 + \lambda_k\left(\mathbf{CC}^H\right), \text{ for } 1 \leq k \leq r \quad (25)
\end{aligned}
$$

and $\lambda_k\left(\mathbf{E}^{-1}\right) = 1$, when $r < k \leq m$. Hence, the MSE can be lower-bounded as

$$
\begin{aligned}
\text{MSE} = \text{Tr}\left(\mathbf{E}\right) &= \sum_{k=1}^{q} \lambda_k\left(\mathbf{E}\right) \\
&\geq \sum_{k=1}^{r} \frac{1}{1 + \lambda_k\left(\mathbf{CC}^H\right)} + (q - r). \quad (26)
\end{aligned}
$$

It follows from the above relationship that the minimum MSE is obtained when $q = r$. Furthermore, observe that $r = \text{rank}\left(\widetilde{\mathbf{H}}\right) \leq N_{\text{RF}}^n$, hence the lower bound is achieved when $N_{\text{RF}}^n \geq q$, i.e., the minimum number of RF chains required is equal to the parameter dimension. This implies that the singular matrix $\mathbf{\Sigma}_{\tilde{H}}$ has the following structure

$$
\mathbf{\Sigma}_{\tilde{H}} = \begin{bmatrix} \mathbf{\Sigma}_{N_{\text{RF}}^n \times N_{\text{RF}}^n} & \mathbf{0}_{N_{\text{RF}}^n \times \left(Nl - N_{\text{RF}}^n\right)} \end{bmatrix}. \quad (27)
$$

Furthermore, in order to achieve the lower bound in (26), the inequality in (24) should be satisfied with equality. This implies that the matrix $\mathbf{Q}$ should be set as an identity matrix of dimension $N_{\text{RF}}^n \times N_{\text{RF}}^n$, which results in the relationship

$\mathbf{V}_{\tilde{H}} = \mathbf{U}_C$. Hence, the optimal matrix $\widetilde{\mathbf{H}}$, denoted by $\widetilde{\mathbf{H}}_{\text{opt}}$, can be set as

$$
\widetilde{\mathbf{H}}_{\text{opt}} = \mathbf{U}_{\tilde{H}} \mathbf{\Sigma}_{\tilde{H}} \mathbf{U}_C^H, \quad (28)
$$

where $\mathbf{U}_{\tilde{H}}$ is any unitary matrix, since it does affect the MSE. Subsequently, the baseband precoder matrix for the individual sensors can be extracted using

$$
\mathbf{P}_{\text{BB},n} = \bar{\mathbf{H}}_n^{-1} \widetilde{\mathbf{H}}_{n,\text{opt}}, \quad (29)
$$

where $\widetilde{\mathbf{H}}_{n,\text{opt}} = \widetilde{\mathbf{H}}_{\text{opt}}(\,:\,,(n-1)q+1:nq)$. Finally, the relationship between the MSE performance achieved and the number of RF chains used can be summarized as follows:

$$
\text{MSE} = \begin{cases} (q - N_{\text{RF}}^n) + \sum_{k=1}^{N_{\text{RF}}^n} \frac{1}{1 + \lambda_k(\mathbf{CC}^H)}, & N_{\text{RF}}^n < q, \\ \sum_{k=1}^{q} \frac{1}{1 + \lambda_k(\mathbf{CC}^H)}, & N_{\text{RF}}^n \geq q. \end{cases}
$$

The next section considers the most general scenario, where both realistic observation and FC noise are present and develops the hybrid minimum-MSE transceiver design, while also taking the available transmit power at each IoTNo into account.

IV. Hybrid Transceiver Design for a Noisy FC

Again, this section develops a novel hybrid transceiver design for the more general scenario, where realistic channel-noise is also present. Additionally, another challenging aspect is also considered, where one does not have perfect knowledge of the dominant array response matrices. Hence, in this scenario, initially an iterative algorithm is developed for designing the minimum-MSE optimal fully-digital TPCs and RC matrices. Once convergence is achieved, the fully-digital matrices are decomposed into their RF and baseband counterparts using the popular SOMP algorithm. The detailed procedure is described next. The estimate $\widehat{\boldsymbol{\theta}}$ for the received vector $\mathbf{y}$ in (2) can be obtained as

$$
\begin{aligned}
\widehat{\boldsymbol{\theta}} &= \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \mathbf{y} \\
&= \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \sum_{n=1}^{N} \mathbf{H}_n \mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{C}_n \boldsymbol{\theta} + \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \times \\
&\quad \sum_{n=1}^{N} \mathbf{H}_n \mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{v}_n + \mathbf{A}_{\text{BB}}^H \mathbf{A}_{\text{RF}}^H \mathbf{w}. \quad (31)
\end{aligned}
$$

The expression of the resultant MSE, which denoted by $\zeta\left(\mathbf{A}_{\text{RF}}, \mathbf{A}_{\text{BB}}, \mathbf{P}_{\text{RF},n}, \mathbf{P}_{\text{BB},n}\right)$, is given in (30). The average transmit power corresponding to the nth IoTNo can be evaluated as

$$
\begin{aligned}
&\mathbb{E}\left[||\mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \mathbf{x}_n||^2\right] \\
&= \text{Tr}\left[\mathbf{P}_{\text{RF},n} \mathbf{P}_{\text{BB},n} \left(\mathbf{C}_n \mathbf{R}_{\boldsymbol{\theta}} \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{P}_{\text{BB},n}^H \mathbf{P}_{\text{RF},n}^H\right] \leq \rho_n, \\
&\hspace{8cm}(32)
\end{aligned}
$$

where ρ_n represents the maximum available transmit power at the nth IoTNo. Hence, the MSE-minimization optimization

$$\zeta\left(\mathbf{A}_{\mathrm{RF}}, \mathbf{A}_{\mathrm{BB}}, \mathbf{P}_{\mathrm{RF},n}, \mathbf{P}_{\mathrm{BB},n}\right) = \mathrm{Tr}\left[\mathbf{A}\left(\sum_{n=1}^{N}\mathbf{H}_n\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_n^H\mathbf{H}_n^H\right)\mathbf{A}^H + \mathbf{A}\left(\sum_{n=1}^{N}\sum_{j=1,j\neq n}^{N}\mathbf{H}_n\mathbf{P}_n\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_j^H\mathbf{P}_j^H\mathbf{H}_j^H\right)\right.$$
$$\left.\mathbf{A}^H - \mathbf{A}\left(\sum_{n=1}^{N}\mathbf{H}_n\mathbf{P}_n\mathbf{C}_n\right)\mathbf{R}_\theta - \mathbf{R}_\theta\left(\sum_{n=1}^{N}\mathbf{C}_n^H\mathbf{P}_n^H\mathbf{H}_n^H\right)\mathbf{A}^H + \mathbf{A}\mathbf{R}_w\mathbf{A}^H + \mathbf{R}_\theta\right]. \tag{30}$$

---

problem for designing the hybrid transceiver can be formulated as

$$\begin{aligned}
\underset{}{\text{minimize}} \quad & \zeta\left(\mathbf{A}_{\mathrm{RF}}, \mathbf{A}_{\mathrm{BB}}, \mathbf{P}_{\mathrm{RF},n}, \mathbf{P}_{\mathrm{BB},n}\right) \\
\text{subject to} \quad & |\mathbf{P}_{\mathrm{RF}}(i,j)| = \frac{1}{\sqrt{N_t}}, |\mathbf{A}_{\mathrm{RF}}(i,j)| = \frac{1}{\sqrt{N_r}}, \, \forall\, i,j \\
& \mathrm{Tr}\left[\mathbf{P}_{\mathrm{RF},n}\mathbf{P}_{\mathrm{BB},n}\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_{\mathrm{BB},n}^H\mathbf{P}_{\mathrm{RF},n}^H\right] \leq \rho_n.
\end{aligned} \tag{33}$$

The above optimization problem is once again non-convex due to the constant-magnitude constraint as well as the non-convex objective function. However, for designing the fully-digital transceiver, one can ignore the non-convex constant magnitude constraint. Furthermore, the block-convex nature of the objective function in terms of the TPCs at the IoTNos and RC at the FC can be exploited using the BCD-based iterative framework, which designs the minimum-MSE TPC and combiner in an iterative fashion.

Employing the BCD algorithm, the optimization framework of the fully digital RC matrix assuming that all the TPC matrices are known reduces to the unconstrained problem of

$$\underset{\mathbf{A}}{\text{minimize}} \quad \zeta\left(\mathbf{A}\right). \tag{34}$$

Since the MSE objective in (30) is a quadratic convex function in terms of the RC matrix $\mathbf{A}$, the optimal RC matrix $\mathbf{A}^*$ can be obtained by differentiating it with respect to $\mathbf{A}$ and equating it to zero, which is given as

$$\mathbf{A}^* = \mathbf{R}_\theta\left(\mathbf{C}_n^H\mathbf{P}_n^H\mathbf{H}_n^H\right)\left[\sum_{n=1}^{N}\mathbf{H}_n\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n + \mathbf{R}_n\right)\right.$$
$$\left.\mathbf{P}_n^H\mathbf{H}_n^H + \sum_{n=1}^{N}\sum_{j=1,j\neq n}^{N}\mathbf{H}_n\mathbf{P}_n\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_j^H\mathbf{P}_j^H\mathbf{H}_j^H + \mathbf{R}_w\right]^{-1}. \tag{35}$$

Given the minimum-MSE fully digital RC, the TPC matrices $\mathbf{P}_n$ corresponding to each IoTNo n can be designed by formulating the optimization problem of

$$\begin{aligned}
\underset{(\mathbf{P}_n)_{n=1}^{N}}{\text{minimize}} \quad & \zeta\left(\left(\mathbf{P}_n\right)_{n=1}^{N}\right) \\
\text{subject to} \quad & \mathrm{Tr}\left[\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_n^H\right] \leq \rho_n, \, 1 \leq n \leq N.
\end{aligned} \tag{36}$$

The above optimization problem is convex in nature for a particular TPC matrix $\mathbf{P}_n$ corresponding to the nth IoTNo assuming that all the TPC matrices corresponding to the other IoTNos are known in addition to the RC $\mathbf{A}^*$. Upon invoking

the Karush-Kuhn-Tucker (KKT) framework [42] for the above problem, the Lagrangian function is given by

$$\mathcal{L}\left(\mathbf{P}_n, \lambda_n\right) = \mathrm{Tr}\left[\mathbf{A}\left(\sum_{n=1}^{N}\mathbf{H}_n\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_n^H\mathbf{H}_n^H\right)\right.$$
$$\mathbf{A}^H + \mathbf{A}\left(\sum_{n=1}^{N}\sum_{j=1,j\neq n}^{N}\mathbf{H}_n\mathbf{P}_n\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_j^H\mathbf{P}_j^H\mathbf{H}_j^H\right)\mathbf{A}^H + \mathbf{A}$$
$$\mathbf{R}_w\mathbf{A}^H - \mathbf{A}\left(\sum_{n=1}^{N}\mathbf{H}_n\mathbf{P}_n\mathbf{C}_n\right)\mathbf{R}_\theta - \mathbf{R}_\theta\left(\sum_{n=1}^{N}\mathbf{C}_n^H\mathbf{P}_n^H\mathbf{H}_n^H\right)$$
$$\left.\mathbf{A}^H + \mathbf{R}_\theta\right] - \lambda_n\left[\mathrm{Tr}\left[\mathbf{F}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{F}_n^H\right] - \rho_n\right], \tag{37}$$

where λ_n is the dual variable corresponding to the nth sensor's power constraint. Upon employing the first order optimality KKT condition, the minimum-MSE TPC matrix corresponding to the nth IoTNo, which is denoted by $\mathbf{P}_n^*$, can be obtained as

$$\mathbf{P}_n^* = \left[\mathbf{H}_n^H\mathbf{A}^H\mathbf{A}\mathbf{H}_n + \lambda_n\mathbf{I}\right]^{-1}\left[\mathbf{H}_n^H\mathbf{A}^H\mathbf{R}_\theta\mathbf{C}_n^H - \right.$$
$$\left.\sum_{\substack{j=1\\j\neq n}}^{N}\mathbf{H}_n^H\mathbf{A}^H\mathbf{A}\mathbf{H}_j\mathbf{P}_j\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_j^H\right)^H\right]\left[\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right]^{-1}. \tag{38}$$

The optimal dual variable corresponding to each IoTNo n can be obtained by solving the complementary slackness condition:

$$\lambda_n\left[\mathrm{Tr}\left[\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_n^H\right] - \rho_n\right] = 0. \tag{39}$$

The calculation of dual variables is shown in Appendix A. The above iterative algorithm terminates, when either a desired MSE level is achieved or a certain pre-set number of iterations have been performed. Once, the minimum-MSE fully-digital TPCs and RC matrices are derived, the next two subsections describe the SOMP algorithm based procedure of decomposing them into their RF and BB counterparts.

A. SOMP based hybrid transceiver design

1) Precoder design: This subsection describes the detailed procedure of decomposing the minimum-MSE fully-digital TPC matrix $\mathbf{P}_n$ corresponding to each IoTNo n into its RF and BB counterparts $\mathbf{P}_{\mathrm{RF},n}$ and $\mathbf{P}_{\mathrm{BB},n}$, respectively. The

corresponding optimization problem can be readily formulated as

$$\underset{\mathbf{P}_{\mathrm{RF},n},\mathbf{P}_{\mathrm{BB},n}}{\text{minimize}} \quad ||\mathbf{P}_n - \mathbf{P}_{\mathrm{RF},n}\mathbf{P}_{\mathrm{BB},n}||_F^2$$
$$\text{subject to} \quad |\mathbf{P}_{\mathrm{RF},n}(i,j)| = \frac{1}{\sqrt{N_t}} \; \forall \, i,j. \quad (40)$$

The above problem is non-convex and intractable in nature, since the elements of the RF TPC must have constant magnitude. However, one can exploit an interesting relationship between the optimal TPC $\mathbf{P}_n$ and the transmit array response matrix $\mathbf{A}_{s,n}$, which is shown next in detail. It allows us to design $\mathbf{P}_{\mathrm{RF},n}$ using the columns of the $\mathbf{A}_{s,n}$. To this end, using the expression for the minimum-MSE TPC $\mathbf{P}_n^*$ in (38) and subsequently substituting the SVD of the matrix $\mathbf{H}_n = \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{V}_n^H$, one obtains

$$\mathbf{P}_n^* = \left[\mathbf{V}_n \mathbf{\Sigma}_n^H \mathbf{U}_n^H \mathbf{A}^H \mathbf{A} \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{V}_n^H + \lambda_n \mathbf{I} \right]^{-1} \Bigg[\mathbf{V}_n \mathbf{\Sigma}_n^H \mathbf{U}_n^H$$

$$\mathbf{A}^H \mathbf{R}_\theta \mathbf{C}_n^H - \sum_{j=1,j\neq n}^{N} \mathbf{V}_n \mathbf{\Sigma}_n^H \mathbf{U}_n^H \mathbf{A}^H \mathbf{A} \mathbf{U}_j \mathbf{\Sigma}_j \mathbf{V}_j^H \mathbf{P}_j \big(\mathbf{C}_n \mathbf{R}_\theta$$

$$\mathbf{C}_j \big)^H \Bigg] \left[\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right]^{-1}$$

$$= \mathbf{V}_n \left[\mathbf{\Sigma}_n^H \mathbf{U}_n^H \mathbf{A}^H \mathbf{A} \mathbf{U}_n \mathbf{\Sigma}_n + \lambda_n \mathbf{I} \right]^{-1} \mathbf{V}_n^H \Bigg[\mathbf{V}_n \mathbf{\Sigma}_n^H \mathbf{U}_n^H \mathbf{A}^H$$

$$\mathbf{R}_\theta \mathbf{C}_n^H - \sum_{j=1,j\neq n}^{N} \mathbf{V}_n \mathbf{\Sigma}_n^H \mathbf{U}_n^H \mathbf{A}^H \mathbf{A} \mathbf{U}_j \mathbf{\Sigma}_j \mathbf{V}_j^H \mathbf{P}_j \big(\mathbf{C}_n \mathbf{R}_\theta$$

$$\mathbf{C}_j^H \big)^H \Bigg] \left[\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right]^{-1}. \quad (41)$$

This implies that the column space of the TPC matrix $\mathbf{P}_n^*$ corresponding to the nth IoTNo lies in the column space of the unitary matrix $\mathbf{V}_n$. Moreover, the columns of $\mathbf{V}_n$ span the row space of the channel matrix $\mathbf{H}_n$. Thus, it can be concluded that

$$\mathcal{C}\left(\mathbf{P}_n\right) \subseteq \mathcal{R}\left(\mathbf{H}_n\right). \quad (42)$$

It can be observed from (12) that the row-space of the mmWave MIMO channel matrix $\mathbf{H}_n$ is a subset of the column-space of the transmit array response matrix $\mathbf{A}_{s,n}$. Thus, we can write

$$\mathcal{C}\left(\mathbf{P}_n\right) \subseteq \mathcal{R}\left(\mathbf{H}_n\right) \subseteq \mathcal{C}\left(\mathbf{A}_{s,n}\right). \quad (43)$$

Hence, we can design the RF TPC matrix $\mathbf{P}_{\mathrm{RF},n}$ by extracting the N_{RF}^n columns of $\mathbf{A}_{s,n}$, which by default satisfies the constant-magnitude condition. Thus, the optimization problem of designing the BB TPC $\mathbf{P}_{\mathrm{BB},n}$ can be formulated as

$$\underset{\bar{\mathbf{P}}_{\mathrm{BB},n}}{\text{minimize}} \quad ||\mathbf{P}_n - \mathbf{A}_{s,n}\bar{\mathbf{P}}_{\mathrm{BB},n}||_F^2$$
$$\text{subject to} \quad \left|\left|\mathrm{diag}\left(\bar{\mathbf{P}}_{\mathrm{BB},n}\bar{\mathbf{P}}_{\mathrm{BB},n}^H\right)\right|\right|_0 = N_{\mathrm{RF}}^n, \quad (44)$$

where the first constraint implies that the matrix $\bar{\mathbf{P}}_{\mathrm{BB},n}$ is block-sparse in nature, since only N_{RF} rows out of N rows are non-zero. Hence, similar to Ayach et al. [8], the simultaneous

---

Algorithm 1 SOMP- based hybrid precoder design

Require: $\{\mathbf{P}_n^{\mathrm{opt}}\}$ and $N_{\mathrm{RF}}^n \quad \forall \, n$
1: **for** $1 \leq n \leq N$ **do**
2: $\quad \mathbf{P}_{\mathrm{RF},n} = [\,]$
3: $\quad \mathbf{P}_{\mathrm{res}} = \mathbf{P}_n$
4: $\quad$ **for** $k \leq N_{\mathrm{RF}}^n$ **do**
5: $\quad\quad \mathbf{\Psi} = \mathbf{A}_{s,n}^H \mathbf{P}_{\mathrm{res}}$
6: $\quad\quad m = \arg\max_{m=1,\dots,N_{cl}} (\mathbf{\Psi}\mathbf{\Psi}^H)_{l,l}$
7: $\quad\quad \mathbf{P}_{\mathrm{RF},n} = [\mathbf{P}_{\mathrm{RF},n} \mid \mathbf{A}_{s,n}^{(n)}]$
8: $\quad\quad \mathbf{P}_{\mathrm{BB},n} = (\mathbf{P}_{\mathrm{RF},n}^H \mathbf{P}_{\mathrm{RF},n})^{-1} \mathbf{P}_{\mathrm{RF},n}^H \mathbf{P}_n$
9: $\quad\quad \mathbf{P}_{\mathrm{res}} = \frac{\mathbf{P}_n - \mathbf{P}_{\mathrm{RF},n}\mathbf{P}_{\mathrm{BB},n}}{||\mathbf{P}_n - \mathbf{P}_{\mathrm{RF},n}\mathbf{P}_{\mathrm{BB},n}||_F}$
10: $\quad$ **end for**
11: $\quad \mathbf{P}_{\mathrm{BB},n} = \frac{\mathbf{P}_{\mathrm{BB},n}||\mathbf{P}_n||_F}{||\mathbf{P}_{\mathrm{RF},n}\mathbf{P}_{\mathrm{BB},n}||_F}$
12: $\quad$ **return** $\mathbf{P}_{\mathrm{RF},n}$, $\mathbf{P}_{\mathrm{BB},n}$
13: $\quad \mathbf{P}_{\mathrm{RF}} = \mathrm{blkdiag}\left(\mathbf{P}_{\mathrm{RF}}, \mathbf{P}_{\mathrm{RF},n}\right)$
14: $\quad \mathbf{P}_{\mathrm{BB}} = \mathrm{blkdiag}\left(\mathbf{P}_{\mathrm{BB}}, \mathbf{P}_{\mathrm{BB},n}\right)$
15: **end for**

---

orthogonal matching pursuit (SOMP) algorithm is employed for design of the RF and BB components of the TPC, since the latter is a simultaneously sparse matrix. Algorithm-1 describes the proposed SOMP-based hybrid TPC design technique in detail. The next subsection describes the detailed procedure to decompose the minimum-MSE RC matrix $\mathbf{A}$ into its RF and BB counterparts.

2) Combiner design: Once the optimal RF and baseband TPC matrices corresponding to each IoTNo n are derived, the next objective is to decompose the minimum-MSE RC matrix into its RF and baseband counterparts. The corresponding optimization is formulated as

$$\underset{\mathbf{A}_{\mathrm{RF}},\mathbf{A}_{\mathrm{BB}}}{\text{minimize}} \quad ||\boldsymbol{\theta} - \mathbf{A}_{\mathrm{BB}}^H, \mathbf{A}_{\mathrm{RF}}^H \mathbf{y}||_F^2$$
$$\text{subject to} \quad |\mathbf{A}_{\mathrm{RF}}(i,j)| = \frac{1}{\sqrt{N_r}} \; \forall \, i,j. \quad (45)$$

Upon, ignoring the constraint in the above optimization problem, the optimal solution of the unconstrained optimization problem is the RC matrix given in (35). As shown in [8], the above optimization problem may also be modified as

$$\underset{\mathbf{A}_{\mathrm{RF}},\mathbf{A}_{\mathrm{BB}}}{\text{minimize}} \quad ||\mathbf{R}_{yy}^{\frac{1}{2}}\left(\mathbf{A} - \mathbf{A}_{\mathrm{RF}}\mathbf{A}_{\mathrm{BB}}\right)||_F^2$$
$$\text{subject to} \quad |\mathbf{A}_{\mathrm{RF}}(i,j)| = \frac{1}{\sqrt{N_r}} \; \forall \, i,j. \quad (46)$$

The above problem is non-convex due to the constant-magnitude constraint. However, similar to the previous subsection, it can be shown that the column space of the RC matrix $\mathbf{A}$ lies in the column space of the concatenated mmWave MIMO channel $\mathbf{H}$. The detailed justification for this follows next. The optimal fully digital RC expression given in (35) can also be written equivalently as

$$\mathbf{A}^H = \left[\mathbf{H}\mathbf{P}\left(\mathbf{C}\mathbf{R}_\theta \mathbf{C}^H + \mathbf{R}_n\right)\mathbf{P}^H \mathbf{H}^H + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{H}\mathbf{P}\mathbf{C}\mathbf{R}_\theta, \quad (47)$$

where the block-diagonal TPC matrix $\mathbf{P} \in \mathbb{C}^{NN_t \times Nl}$ is defined as

$$\mathbf{P} = \mathrm{blkdiag}\left[\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N\right],$$

and $\mathbf{P}_n = \mathbf{P}_{\text{RF},n}\mathbf{P}_{\text{BB},n}$. Substituting $\mathbf{R}_n = \sigma_n^2 \mathbf{I}_l$, and using the property $(\mathbf{I} + \mathbf{GQ})^{-1}\mathbf{G} = \mathbf{G}(\mathbf{I} + \mathbf{QG})^{-1}$, for any suitable matrices $\mathbf{G}$ and $\mathbf{Q}$, the optimal fully-digital RC expression above can further be recast as

$$\mathbf{A}^H = \mathbf{HP}\left[\left(\mathbf{CR}_\theta\mathbf{C}^H + \mathbf{R}_n\right)\mathbf{P}^H\mathbf{H}^H\mathbf{HP} + \sigma_n^2\mathbf{I}\right]^{-1}\mathbf{CR}_\theta, \tag{48}$$

where $\mathbf{G} = \mathbf{HP}$ and $\mathbf{Q} = \left(\mathbf{CR}_\theta\mathbf{C}^H + \mathbf{R}_n\right)\mathbf{P}^H\mathbf{H}^H$. It can now be observed that any column of $\mathbf{A}^H$ can be written as a linear combination of the columns of $\mathbf{H}$. Furthermore, the columns of the matrix $\mathbf{H}$ are linear combinations of the columns of the receive array response matrix $\mathbf{A}_{\text{FC}}$. Thus, it can be concluded that

$$\mathcal{C}\left(\mathbf{A}^H\right) \subseteq \mathcal{C}\left(\mathbf{H}\right) \subseteq \mathcal{C}\left(\mathbf{A}_{\text{FC}}\right). \tag{49}$$

Hence, we can design $\mathbf{A}_{\text{RF}}$ by extracting N_{RF} columns from $\mathbf{A}_{\text{FC}}$, which also satisfy the constant gain condition on the elements of the RF RC matrix. Subsequently, one can develop the BB RC design problem as

$$\begin{aligned}\underset{\mathbf{A}_{\text{BB}}}{\text{minimize}} \quad & ||\mathbf{R}_{yy}^{\frac{1}{2}}\left(\mathbf{A} - \mathbf{A}_{\text{FC}}\bar{\mathbf{A}}_{\text{BB}}\right)||_F^2 \\ \text{subject to} \quad & \left\|\text{diag}\left(\bar{\mathbf{A}}_{\text{BB}}\bar{\mathbf{A}}_{\text{BB}}^H\right)\right\|_0 = N_{\text{RF}}.\end{aligned} \tag{50}$$

The above optimization problem is once again a sparse signal recovery problem and can be solved similar to (44) upon employing the SOMP algorithm. The complete iterative procedure of designing the hybrid transceiver is summarized in Algorithm-2. The next subsection presents a proof of the convergence of the proposed BCD-based iterative algorithm.

B. Convergence analysis

The proposed BCD-based iterative transceiver scheme's theoretical convergence can be proved as follows. Due to the fact that the optimization problems of the fully-digital RC and TPC matrices in (34) and (36) are convex in nature, the following inequalities hold:

$$\begin{aligned}&\text{MSE}\left(\mathbf{A}(i), \{\mathbf{P}_n(i)\}_{n=1}^N\right) \geq \min_{\mathbf{A}} \text{MSE}\left(\mathbf{A}|\{\mathbf{P}_n(i)\}_{n=1}^N\right) \\ &= \text{MSE}\left(\mathbf{A}(i+1), \{\mathbf{P}_n(i)\}_{n=1}^N\right) \\ &\geq \min_{\mathbf{P}_n}\text{MSE}\left(\mathbf{P}_n|\mathbf{A}(i+1), \{\mathbf{P}_k(i+1)\}_{k=1}^{N-1}, \{\mathbf{P}_k(i)\}_{k=n+1}^N\right) \\ &= \text{MSE}\left(\mathbf{P}_n(i+1), \mathbf{A}(i+1)\right).\end{aligned} \tag{51}$$

Hence, $\text{MSE}\left(\{\mathbf{A}\}^{(i)}, \{\mathbf{P}_n\}^{(i)}\right)$ is a monotonically decreasing sequence and it is lower-bounded by zero. This proves that the algorithm converges. On convergence, the SOMP algorithm is employed for obtaining the RF and BB TPCs and RCs from their respective fully-digital counterparts. The next section develops robust transceiver designs for a scenario associated with CSI uncertainty.

V. Robust Hybrid Transceiver Design for a Noisy FC

In practical scenarios, it is impossible to have perfect CSI knowledge between each IoTNo and the FC due to the limited

---

Algorithm 2 BCD-based iterative algorithm for hybrid transceiver design

1: **Input**: Observation vector $\mathbf{y}$, maximum iterations $i_{\max} = 40$ and desired accuracy $\epsilon = 0.0001$.
2: **Initialization**: $i = 1$, initialize precoding matrices $\{\mathbf{P}_n(i-1)\}_{n=1}^N$ randomly.
3: **while** $\|\widehat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta}^{(n)}\|_2 \geq \epsilon$ and $i < i_{max}$ **do**
4:     Evaluate the fully digital combiner matrix $\mathbf{A}(i)$ using (35).
5:     Evaluate the optimal precoding matrices $\{\mathbf{P}_n(i)\}_{n=1}^N$ using (38).
6:     update $i = i + 1$.
7: **end while**
8: Employ the SOMP algorithm detailed in Subsection-IV-A to decompose the fully digital combiner/ precoder into the corresponding hybrid combiner/ precoder.
9: **Output**: $\mathbf{A}_{\text{RF}}, \mathbf{A}_{\text{BB}}, \{\mathbf{P}_{\text{RF},n}\}_{n=1}^N, \{\mathbf{P}_{\text{BB},n}\}_{n=1}^N$.

---

pilot overhead and quantization errors. Hence, in a practical system, one has to take the CSI uncertainty into account. Therefore, motivated by [35], [37], [38], we model the channel between each IoTNo and the FC as

$$\mathbf{H}_n = \widehat{\mathbf{H}}_n + \Delta\mathbf{H}_n, \tag{52}$$

where $\widehat{\mathbf{H}}_n$ denotes the available channel estimate and $\Delta\mathbf{H}_n$ represents the estimation error matrix whose elements obey the distribution $\mathcal{CN}\left(0, \sigma_H^2\right)$. Substituting $\mathbf{H}_n$ from (52) into the expression in (30), the expression for the resultant average MSE defined as $\bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right) = \mathbb{E}_{\Delta\mathbf{H}_n}\zeta\left(\mathbf{A}, \mathbf{P}_n\right)$, is given by

$$\begin{aligned}\bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right) = \text{Tr}&\left[\mathbf{A}\left(\sum_{n=1}^N \widehat{\mathbf{H}}_n\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{F}_n^H\widehat{\mathbf{H}}_n^H\right)\right. \\ &\mathbf{A}^H + \mathbf{A}\left(\sum_{n=1}^N \sigma_H^2\text{Tr}\left[\mathbf{P}_n\left(\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_n^H + \mathbf{R}_n\right)\mathbf{P}_n^H\right]\right)\mathbf{A}^H + \mathbf{R}_\theta \\ &+ \mathbf{AR}_w\mathbf{A}^H + \mathbf{A}\left(\sum_{n=1}^N\sum_{\substack{j=1\\j\neq n}}^N \widehat{\mathbf{H}}_n\mathbf{P}_n\mathbf{C}_n\mathbf{R}_\theta\mathbf{C}_j^H\mathbf{P}_j^H\widehat{\mathbf{H}}_j^H\right)\mathbf{A}^H \\ &\left.- \mathbf{A}\left(\sum_{n=1}^N \widehat{\mathbf{H}}_n\mathbf{P}_n\mathbf{C}_n\right)\mathbf{R}_\theta - \mathbf{R}_\theta\left(\sum_{n=1}^N \mathbf{C}_n^H\mathbf{P}_n^H\widehat{\mathbf{H}}_n^H\right)\mathbf{A}^H\right].\end{aligned} \tag{53}$$

The detailed steps are shown in Appendix B. Hence, the pertinent optimization problem for designing the hybrid transceiver in this framework associated with CSI uncertainty so that the average MSE is minimized at the FC is similar to (33), with the objective function $\zeta\left(\mathbf{A}, \mathbf{P}_n\right)$ replaced by $\bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right)$. The BCD-based iterative framework can once again be employed to design the fully-digital robust transceiver followed by the SOMP algorithm to yield the robust hybrid TPC and RC matrices. One can determine the robust RC matrix at the FC for a given set of robust TPC matrices $\{\mathbf{P}_n\}_{n=1}^N$, by setting
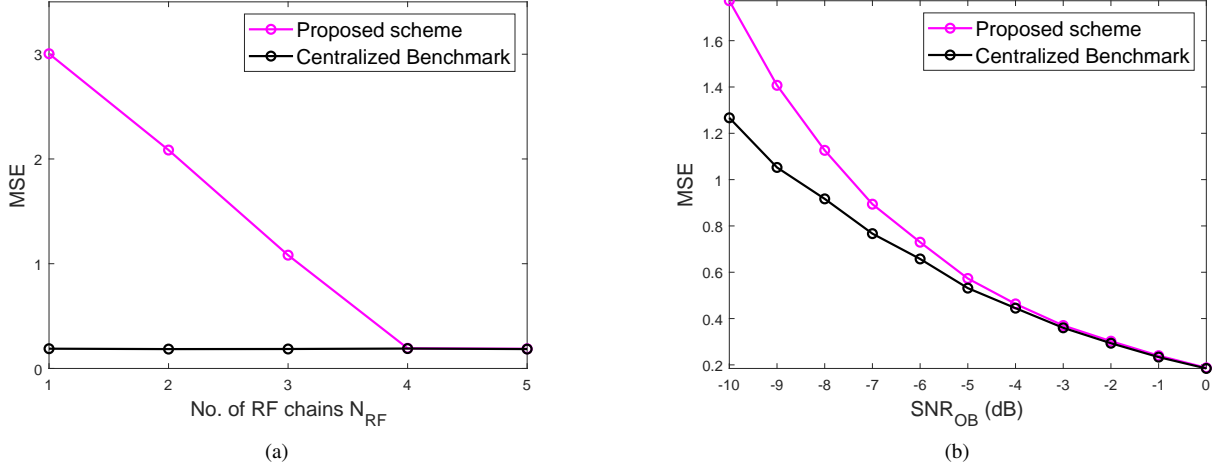
Fig. 3. (a) MSE versus number of RF chains N_{RF}^n at $\text{SNR}_{\text{OB}} = 0$ dB (b) MSE versus SNR_{OB} when $N_{\text{RF}}^n = 4$, for the mmWave MIMO IoTNe with $N = 25, N_t = 5, N_r = 16, q = 4, l = 2, K = 5$.

$\frac{\partial \bar{\zeta}(\mathbf{A})}{\partial \mathbf{A}} = \mathbf{0}$, as

$$\mathbf{A}^* = \mathbf{R}_\theta \left(\mathbf{C}_n^H \mathbf{P}_n^H \widehat{\mathbf{H}}_n^H \right) \left[\sum_{n=1}^{N} \widehat{\mathbf{H}}_n \mathbf{P}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right) \right.$$

$$\mathbf{P}_n^H \widehat{\mathbf{H}}_n^H + \sum_{n=1}^{N} \sigma_H^2 \text{Tr} \left[\mathbf{P}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right) \mathbf{P}_n^H \right]$$

$$+ \left. \sum_{n=1}^{N} \sum_{j=1, j \neq n}^{N} \widehat{\mathbf{H}}_n \mathbf{P}_n \mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_j^H \mathbf{P}_j^H \widehat{\mathbf{H}}_j^H + \mathbf{R}_w \right]^{-1}. \quad (54)$$

The optimization problem of designing the robust TPC matrices, given the fully-digital RC matrix, is equivalent to (36) with $\zeta \left(\{\mathbf{P}_n\}_{n=1}^N \right)$ replaced by $\bar{\zeta} \left(\{\mathbf{P}_n\}_{n=1}^N \right)$. This can be solved along similar lines by invoking the KKT framework of [42] to obtain the robust fully-digital TPC matrices, given as

$$\mathbf{P}_n^* = \left[\widehat{\mathbf{H}}_n^H \mathbf{A}^H \mathbf{A} \widehat{\mathbf{H}}_n + \sigma_H^2 \text{Tr} \left[\mathbf{A}^H \mathbf{A} \right] + \lambda_n \mathbf{I} \right]^{-1} \left[\widehat{\mathbf{H}}_n^H \mathbf{A}^H \mathbf{R}_\theta \right.$$

$$\mathbf{C}_n^H - \sum_{j=1, j \neq n}^{L} \widehat{\mathbf{H}}_n^H \mathbf{A}^H \mathbf{A} \widehat{\mathbf{H}}_j \mathbf{P}_j \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_j^H \right)^H \right] \left[\mathbf{C}_n \right.$$

$$\left. \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right]^{-1}. \quad (55)$$

As described in the previous section, the SOMP-based algorithm can once again be employed for deriving the hybrid TPC and RC matrices for this scenario in the face of CSI uncertainty. The next subsection derives the centralized MMSE benchmark, which acts as a lower bound for the LDE schemes.

A. Centralized MMSE benchmark

The centralized MMSE benchmark represents the best achievable performance where all the IoTNo observations are directly available at the FC. The observation vector $\mathbf{x} \in \mathbb{C}^{Nl \times 1}$ available at the FC can be modeled as

$$\mathbf{x} = \mathbf{C}\boldsymbol{\theta} + \mathbf{v}, \quad (56)$$

where $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_L^T]^T \in \mathbb{C}^{Nl \times 1}$. The MSE is formulated as

$$\text{MSE}_{\text{MMSE}} = \text{Tr} \left[\left(\mathbf{I}_q + \mathbf{C}^H \mathbf{C} \right)^{-1} \right]. \quad (57)$$

The MSE bound above can be used as a benchmark for the evaluation of the performance of the proposed LDE schemes. The next section discusses the simulation results to illustrate the effectiveness of the various designs proposed.

VI. SIMULATION RESULTS

A mmWave MIMO IoTNe is considered where each IoTNo is equipped with $N_t = 5$ TAs, and the FC has $N_r = 10$ RAs. The number of clusters in the mmWave MIMO channel is set to $K = 5$ clusters. The path-gains $\alpha_{k,n}$ are randomly generated from the distribution $\mathcal{CN}(0, 1)$. The dimension of the parameter vector to be estimated is set to $q = 3$ and the number of observations made by each IoTNo is set to $l = 2$. The observation SNR is defined as $\text{SNR}_{\text{OB}} = \frac{1}{\sigma_l^2}$ whereas the $\text{SNR}_{\text{FC}} = \frac{1}{\sigma_n^2}$. The elements of the observation matrix $\mathbf{C}_n$ have been randomly generated from the distribution $\mathcal{CN}(0, 1)$.

Fig. 3(a) illustrates the MSE performance of the proposed hybrid transceiver against the number of RF chains N_{RF} employed at each IoTNo at $\text{SNR}_{\text{OB}} = 0$ dB. Observe that the MSE performance improves upon increasing the number of RF chains, and as claimed in (26) it matches the centralized MMSE benchmark derived in (57) when the number of RF chains N_{RF} becomes equal to the parameter length q. It can also be readily observed that increasing the number of RF chains beyond the parameter dimension does not result in any improvement in the MSE performance at the FC, thus validating our analytical result.

Fig. 3(b) illustrates the MSE versus SNR_{OB}, when the number of RF chains N_{RF} is set equal to the length of the parameter q. It can be observed from the figure that the MSE of both the proposed scheme and the centralized benchmark improve upon increasing the observation SNR, which is along expected lines. Moreover, the gap between the proposed scheme and the centralized benchmark decreases
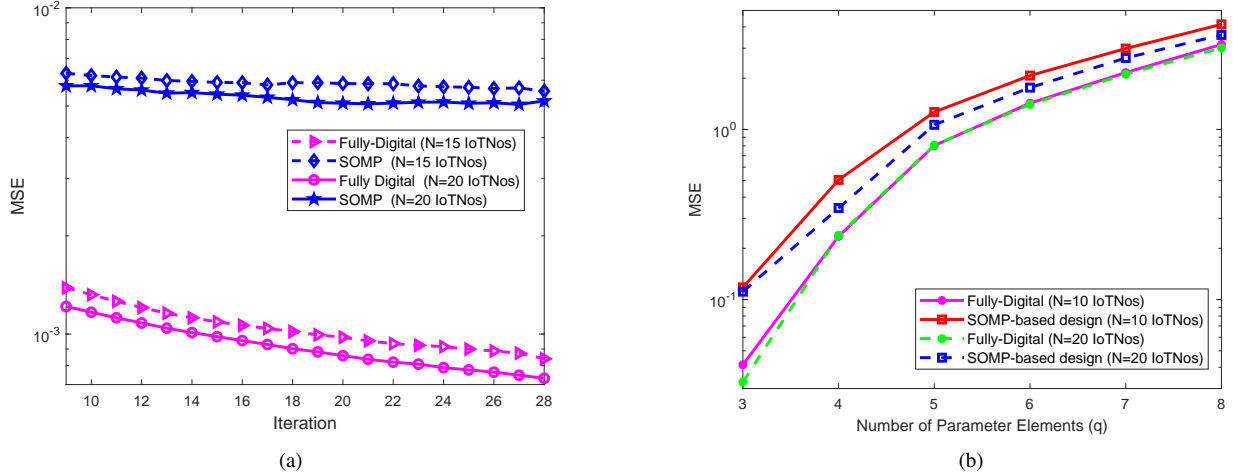
Fig. 4. (a) MSE versus the number of BCD iterations for different number of IoTNos $N \in \{10, 20\}$ (b) MSE versus the number of elements in the parameter vector $\boldsymbol{\theta}$ for different number of IoTNos $N \in \{10, 20\}$ (b).
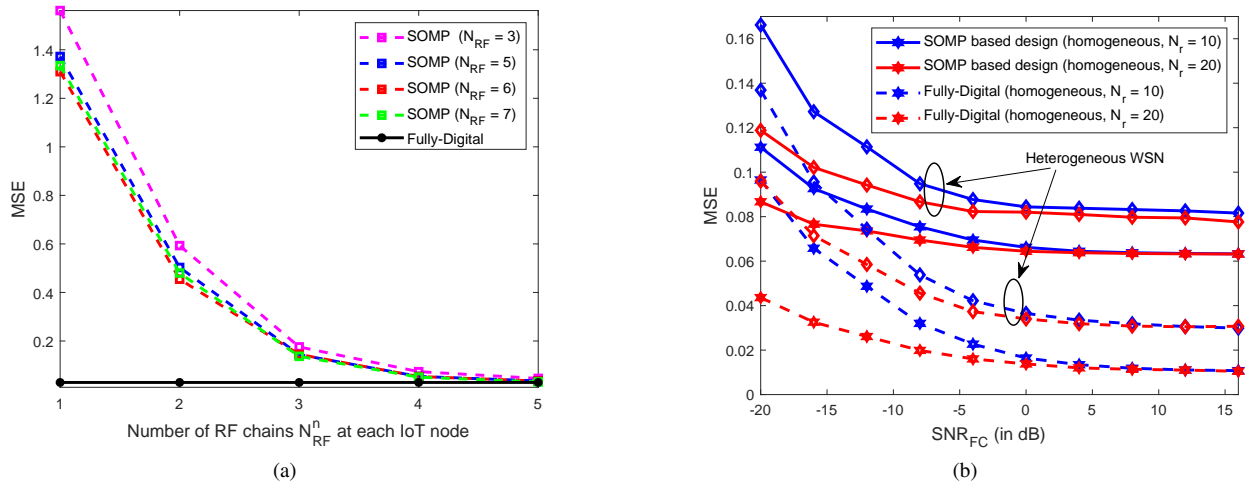

Fig. 5. (a) MSE versus the number of RF chains (N_{RF}^n) for $N = 20$ IoTNos (b) MSE versus SNR$_{FC}$ for different number of RAs $N_r \in \{10, 20\}$.

upon increasing the observation SNR and matches the centralized benchmark at high SNRs, thus validating the effectiveness of the proposed scheme.

Fig. 4(a) depicts the MSE performance as a function of the number of the BCD-iterations for the design proposed in Section-IV for a noisy FC with perfect CSI. The number of IoTNos is $N \in \{15, 20\}$ for this particular plot. To benchmark the SOMP-based hybrid transceiver design's performance, its corresponding fully-digital performance has also been shown. The MSE performance improves as the number of iterations increases. Furthermore, the MSE also decreases as the number of sensors K increases, a trend that is along expected lines due to the availability of a larger number of observations.

Fig. 4(b) illustrates the MSE performance against the dimension of the underlying parameter vector of interest for different number of IoTNos N. Observe that the MSE performance degrades upon increasing the dimensionality of the parameter. When increasing the number of IoTNos, the MSE performance improves as it leads to the availability of more measurements of the parameter, thereby increasing the

estimation accuracy. The performance of the proposed SOMP-based hybrid transceiver design is observed to be very close to that of the fully-digital transceiver design, demonstrating the efficiency of the proposed design.

Fig. 5(a) illustrates the MSE versus the number of RF chains N_{RF}^n at each IoTNo for different number of RF chains N_{RF} at the noisy FC. The number of IoTNos considered in the system is $N = 20$. Again, the MSE performance of the SOMP-based hybrid transceiver improves upon increasing the number of RF chains. It becomes equivalent to the MSE corresponding to the fully-digital design, because N_{RF}^n becomes equal to the total number of TAs $N_t = 5$. It is also observed that upon increasing the number of RF chains at the FC, the MSE performance further improves. However, for $N_{RF} \geq K$ clusters, it is observed that the MSE performance of the hybrid designs becomes similar. This demonstrates that the FC is indeed capable of recovering the active beams using our SOMP-based RF RC when $N_{RF} = \text{rank}(\mathbf{H})$ and there is no further improvement in performance upon increasing N_{RF}.

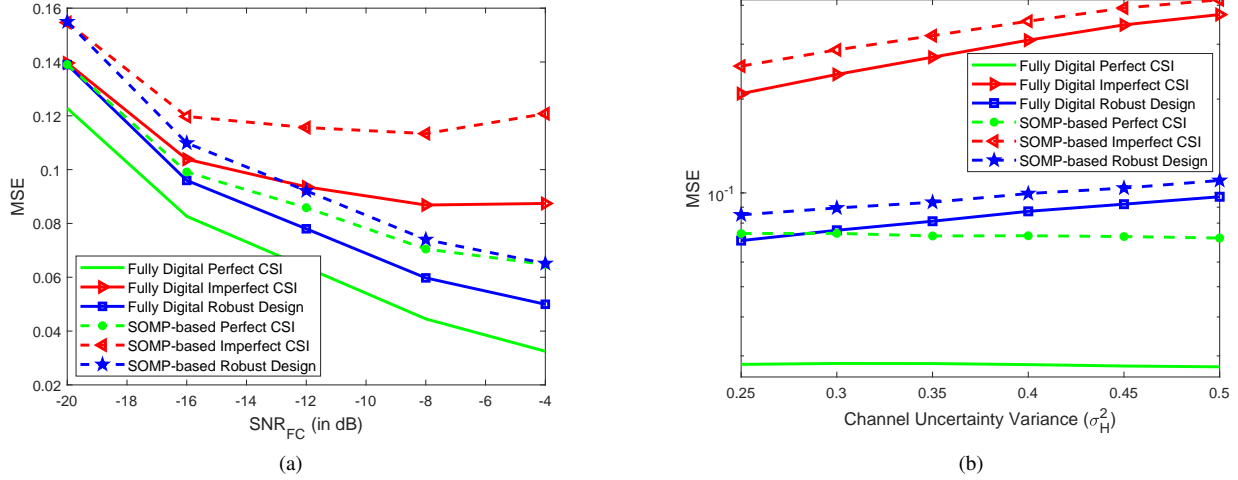Fig. 5(b) illustrates the MSE performance of the hybrid

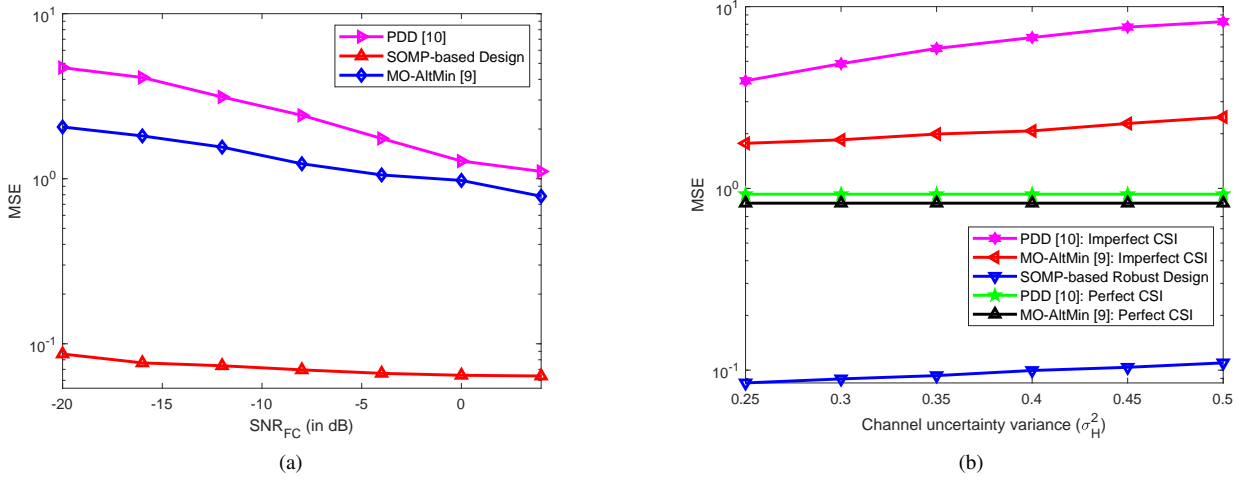Fig. 6. (a) MSE versus SNR_{FC} (b) MSE versus the channel uncertainty variance (σ_H^2) for $N = 20$ IoTNos.

Fig. 7. (a) MSE versus SNR_{FC} (b) MSE versus the channel uncertainty variance (σ_H^2) for $N = 20$ IoTNos.
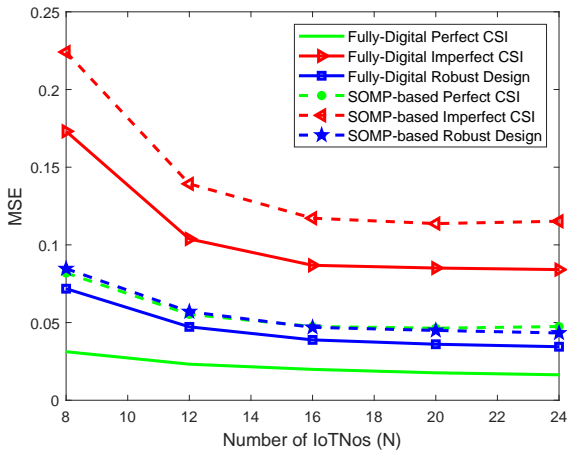

Fig. 8. MSE versus number of IoTNos (N)

transceiver designs versus SNR_{FC}. For comparison, a heterogeneous IoTNe is considered, where the observation noise

power at each sensor takes values from the following set $\{-10, -9, \cdots, 9\}$ dB. This setup is compared to a homogeneous IoTNe, where the observation noise at each sensor is $\sigma_l^2 = -10$dB. The homogeneous mmWave IoTNe is observed to perform better than its counterpart if the SNR value at each IoTNo for the homogeneous mmWave IoTNe is higher than that of heterogeneous mmWave counterpart. Upon increasing the number of RAs N_r from 10 to 20, the MSE performance improves due to the array gain in the noise limited region. As SNR_{FC} increases, the performance gain due to the increasing array gain becomes negligible.

Fig. 6(a) plots the MSE as a function of SNR_{FC} for the hybrid TPC/ RC designs derived in Section-IV and Section-V, respectively. The iterative algorithm is run for 15 iterations for each SNR_{FC} value. As expected, the MSE performance improves as SNR_{FC} increases, and the robust design offers significant performance improvements over the imperfect CSI-based design that ignores the CSI uncertainty. The corresponding performance of the fully-digital transceiver is once again plotted for benchmarking the hybrid transceiver designs.

Fig. 6(b) shows the MSE performance of the hybrid

transceiver designs proposed in Section-IV and Section-V, for scenarios of perfect and imperfect CSI, respectively. To benchmark the performance of the hybrid TPC/ RC designs, the MSE performance corresponding to their fully-digital counterparts is also plotted. The iterative algorithm proposed for our robust transceiver design converges in 15 iterations. It can observed from the figure that the proposed robust design offers an improved MSE performance in comparison to its uncertainty-agnostic counterpart. In addition, its performance is also close to the perfect CSI-based design, which illustrates its value for practical systems.

Fig. 7(a) compares the MSE performance of our hybrid TPC/ RC design of Section IV to that of the corresponding designs in [9] and [10] that proposed the alternating minimization (AltMin) algorithm based on manifold optimization (MO) and penalty dual decomposition (PDD), respectively. The figure clearly demonstrates that the proposed SOMP-based design outperforms both these algorithms. Similarly, Fig. 7(b) compares the MSE performance of the proposed SOMP-based robust hybrid TPC/ RC design of Section V with that of [9] and [10], considering both perfect CSI and CSI-uncertainty agnostic approaches. It can be deduced from the figure that the proposed robust design offers a significant MSE performance improvement over both the agnostic as well as perfect CSI-based hybrid designs of [9] and [10]. This illustrates the efficacy of our proposed robust design.

Fig. 8 shows the MSE performance versus the number of IoTNos N in the network for the schemes proposed in Section-IV and -V, both with perfect and imperfect CSI, respectively. As expected, the MSE decreases upon increasing the number of sensors, which reinforces the trend seen in the previous figures.

VII. CONCLUSION

Hybrid transceiver designs were conceived for the LDE of a parameter vector in a mmWave MIMO IoTNe. Separate design procedures were presented for a noiseless as well as noisy FC. An important observation for the proposed hybrid transceiver design for a noiseless FC is that even in systems having a large number of antennas both at the IoTNos and the FC, increasing the number of RF chains beyond the parameter dimension does not result in any further MSE improvement. For a noisy FC, the fully digital TPC/ RC are initially determined using a BCD-based iterative framework followed by the SOMP technique for decomposing it into its RF and BB components. A robust hybrid design procedure was also formulated for a practical scenario having CSI uncertainty, which provides improved estimation performance over its CSI uncertainty-agnostic counterpart. Our simulation results illustrate the performance of the proposed schemes and corroborate our analytical findings.

APPENDIX A
CALCULATION OF DUAL VARIABLES

To obtain the dual variable λ_n at the nth IoTNo, let the following matrices be defined as

$$\mathbf{X}_n = \mathbf{H}_n^H \mathbf{A}^H \mathbf{A} \mathbf{H}_n, \tag{59}$$

$$\mathbf{Y}_n = \left[\mathbf{H}_n^H \mathbf{A}^H \mathbf{R}_\theta \mathbf{G}_n^H - \sum_{j=1, j\neq n}^{N} \mathbf{H}_n^H \mathbf{A}^H \mathbf{A} \mathbf{H}_j \mathbf{P}_j \right.$$
$$\left. \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_j^H\right)^H \right] \left[\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n \right]^{-1}, \tag{60}$$

where $\mathbf{X}_n \in \mathbb{C}^{N_t \times N_t}$ and $\mathbf{Y}_n \in \mathbb{C}^{N_t \times l}$. The minimum-MSE TPC matrix in (38) can be compactly written as

$$\mathbf{P}_n^* = \left[\mathbf{X}_n + \lambda_n \mathbf{I}\right]^{-1} \mathbf{Y}_n, \tag{61}$$

Since $\mathbf{X}_n$ is a Hermitian symmetric matrix, its eigenvalue decomposition (EVD) is given by $\widetilde{\mathbf{U}}_n \widetilde{\mathbf{\Lambda}}_n \widetilde{\mathbf{U}}_n^H$. Initially, $\mathbf{P}_n^*$ is calculated by setting $\lambda_n = 0$. If the average transmit power in (32) is satisfied, then $\lambda_n = 0$. Otherwise for a non-zero dual variable, the average transmit power is met with equality from the complementary slackness condition

$$\lambda_n \left[\mathrm{Tr}\left[\mathbf{P}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{P}_n^H \right] - \rho_n \right] = 0. \tag{62}$$

The value of the dual variable λ_n for the average transmit power to be satisfied with equality can be calculated as

$$\mathrm{Tr}\left[\mathbf{P}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{P}_n^H \right] = \rho_n \tag{63}$$

$$= \mathrm{Tr}\left[\left[\mathbf{X}_n + \lambda_n \mathbf{I}\right]^{-1} \mathbf{Y}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \right.$$
$$\left. \times \mathbf{Y}_n^H \left[\mathbf{X}_n + \lambda_n \mathbf{I}\right]^{-1} \right]$$

$$= \mathrm{Tr}\left[\left[\widetilde{\mathbf{U}}_n \widetilde{\mathbf{\Lambda}}_n \widetilde{\mathbf{U}}_n^H + \lambda_n \mathbf{I}\right]^{-1} \mathbf{Y}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \right.$$
$$\left. \times \mathbf{Y}_n^H \left[\widetilde{\mathbf{U}}_n \widetilde{\mathbf{\Lambda}}_n \widetilde{\mathbf{U}}_n^H + \lambda_n \mathbf{I}\right]^{-1} \right]$$

$$= \mathrm{Tr}\left[\widetilde{\mathbf{U}}_n \left[\widetilde{\mathbf{\Lambda}}_n + \lambda_n \mathbf{I}\right]^{-1} \widetilde{\mathbf{U}}_n^H \mathbf{Y}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \right.$$
$$\left. \times \mathbf{Y}_n^H \widetilde{\mathbf{U}}_n \left[\mathbf{\Lambda}_n + \lambda_n \mathbf{I}\right]^{-1} \widetilde{\mathbf{U}}_n^H \right]$$

$$= \mathrm{Tr}\left[\left[\widetilde{\mathbf{\Lambda}}_n + \lambda_n \mathbf{I}\right]^{-2} \times \right.$$
$$\left. \underbrace{\widetilde{\mathbf{U}}_n^H \mathbf{Y}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{Y}_n^H \widetilde{\mathbf{U}}_n}_{\mathbf{Z}_n} \right]$$

$$= \sum_{k=1}^{N_t} \frac{z_{kk}}{\left(\tilde{\lambda}_k + \lambda_n\right)^2} = \rho_n,$$

where $\tilde{\lambda}_k$ is the kth eigenvalue of $\mathbf{X}_n$ and z_{kk} is the (k,k)th element of matrix $\mathbf{Z}_n = \widetilde{\mathbf{U}}_n^H \mathbf{Y}_n \left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{Y}_n^H \widetilde{\mathbf{U}}_n$. The dual variable λ_n can now be obtained via the classic bisection method.

$$
\mathbb{E}_{\{\Delta \mathbf{H}_n\}} \zeta\left(\mathbf{A}, \mathbf{P}_n\right) = \bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right) = \operatorname{Tr}\left[\mathbf{A}\left(\sum_{n=1}^{N} \mathbb{E}_{\Delta \mathbf{H}_n}\left\{\mathbf{H}_n \mathbf{P}_n\left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{P}_n^H \mathbf{H}_n^H\right\}\right) \mathbf{A}^H + \mathbf{A} \mathbf{R}_n \mathbf{A}^H + \mathbf{R}_\theta\right.
$$
$$
+ \mathbf{A}\left(\sum_{n=1}^{N} \sum_{j=1, j \neq n}^{N} \mathbb{E}_{\{\Delta \mathbf{H}_n, \Delta \mathbf{H}_j\}} \mathbf{H}_n \mathbf{P}_n \mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_j^H \mathbf{P}_j^H \mathbf{H}_j^H\right) \mathbf{A}^H - \mathbf{A}\left(\sum_{n=1}^{N} \mathbb{E}_{\Delta \mathbf{H}_n} \mathbf{H}_n \mathbf{P}_n \mathbf{C}_n\right) \mathbf{R}_\theta
$$
$$
\left. - \mathbf{R}_\theta\left(\sum_{n=1}^{N} \mathbb{E}_{\Delta \mathbf{H}_n} \mathbf{C}_n^H \mathbf{P}_n^H \mathbf{H}_n^H\right) \mathbf{A}^H\right].
$$
$$(58)$$

$$
\bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right) = \operatorname{Tr}\left[\mathbf{A}\left(\sum_{n=1}^{N} \widehat{\mathbf{H}}_n \mathbf{P}_n\left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right) \mathbf{P}_n^H \widehat{\mathbf{H}}_n^H\right) \mathbf{A}^H + \left(\sum_{n=1}^{N} \sigma_H^2 \operatorname{Tr}\left(\mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_n^H + \mathbf{R}_n\right)\right) \mathbf{A} \mathbf{A}^H + \mathbf{A} \mathbf{R}_n \mathbf{A}^H + \right.
$$
$$
\left. \mathbf{R}_\theta + \mathbf{A}\left(\sum_{n=1}^{N} \sum_{j=1, j \neq n}^{N} \widehat{\mathbf{H}}_n \mathbf{P}_n \mathbf{C}_n \mathbf{R}_\theta \mathbf{C}_j^H \mathbf{P}_j^H \widehat{\mathbf{H}}_j^H\right) \mathbf{A}^H - \mathbf{A}\left(\sum_{n=1}^{N} \widehat{\mathbf{H}}_n \mathbf{P}_n \mathbf{C}_n\right) \mathbf{R}_\theta - \mathbf{R}_\theta\left(\sum_{n=1}^{N} \widehat{\mathbf{C}}_n^H \mathbf{P}_n^H \mathbf{H}_n^H\right) \mathbf{A}^H\right].
$$

APPENDIX B
AVERAGE MSE CALCULATION IN (53)

The channel estimation error matrices $\{\Delta \mathbf{H}_n\}_{n=1}^{N}$ for each IoTNo in (52) are assumed to be independent of each other. The average MSE $\bar{\zeta}\left(\mathbf{A}, \mathbf{P}_n\right)$ is given by (58) shown at the top of this page.

The lemma described below is employed to evaluate the average MSE.

Lemma 1. *Consider a matrix* $\mathbf{H}_n \in \mathbb{C}^{N_R \times N_T}$, *defined as* $\mathbf{H}_n = \widehat{\mathbf{H}}_n + \Delta \mathbf{H}_n$, *where* $\widehat{\mathbf{H}}_n$ *is known and each element of* $\Delta \mathbf{H}_n$ *is assumed to be Gaussian distributed as* $\mathcal{CN}\left(0, \sigma_H^2\right)$. *Then, for a matrix* $\mathbf{K}$, *the following properties hold*

$$
\mathbb{E}_{\Delta \mathbf{H}_n}\left\{\mathbf{H}_n \mathbf{K} \mathbf{K}^H \mathbf{H}_n^H\right\} = \widehat{\mathbf{H}}_n \mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_n^H + \sigma_H^2 \operatorname{Tr}\left(\mathbf{K} \mathbf{K}^H\right) \mathbf{I}_{N_R}
$$
$$
\mathbb{E}_{\Delta \mathbf{H}_n, \Delta \mathbf{H}_j}\left\{\mathbf{H}_n \mathbf{K} \mathbf{K}^H \mathbf{H}_j^H\right\} = \widehat{\mathbf{H}}_n \mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_j^H, n \neq j.
$$

Proof. The first result of Lemma 1 can be proved as follows: The term $\mathbf{H}_n \mathbf{K} \mathbf{K}^H \mathbf{H}_n^H$ can be expanded as

$$
\mathbf{H}_n \mathbf{K} \mathbf{K}^H \mathbf{H}_n^H = \left(\widehat{\mathbf{H}}_n + \Delta \mathbf{H}_n\right) \mathbf{K} \mathbf{K}^H \left(\widehat{\mathbf{H}}_n + \Delta \mathbf{H}_n\right)^H
$$
$$
= \widehat{\mathbf{H}}_n \mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_n^H + \underbrace{\Delta \mathbf{H}_n \mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_n^H}_{\mathbf{T}_1} + \underbrace{\widehat{\mathbf{H}}_n \mathbf{K} \mathbf{K}^H \Delta \mathbf{H}_n^H}_{\mathbf{T}_2}
$$
$$
+ \underbrace{\Delta \mathbf{H}_n \mathbf{K} \mathbf{K}^H \Delta \mathbf{H}_n^H}_{\mathbf{T}_3}. \tag{64}
$$

For compact representation, consider $\mathbf{T}_1, \mathbf{T}_2$ and $\mathbf{T}_3$ as shown above.
Each (a, b) th element of $\mathbf{T}_1$ is the dot product of the ath row of $\Delta \mathbf{H}_n$ and bth column of $\mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_n^H$. This implies that

$$
[\mathbf{T}_1]_{ab} = \sum_{i=1}^{N_T} [\Delta \mathbf{H}_n]_{ai} \left[\mathbf{K} \mathbf{K}^H \widehat{\mathbf{H}}_n^H\right]_{ib}.
$$

Thus, the expected value of any element of $\mathbf{T}_1$ is zero, since $\mathbb{E}\left[\Delta \mathbf{H}_n\right]_{ai} = 0, 1 \leq a \leq N_R, 1 \leq i \leq N_T$. This further implies that

$$
\mathbb{E}_{\Delta \mathbf{H}_n}\left\{\mathbf{T}_1\right\} = \mathbf{0}.
$$

Following the same reasoning as above, it can be shown that

$$
\mathbb{E}_{\Delta \mathbf{H}_n}\left\{\mathbf{T}_2\right\} = \mathbf{0}.
$$

To compute the expected value of $\mathbf{T}_3$, let us denote the EVD of the Hermitian matrix $\mathbf{K} \mathbf{K}^H$ as $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$. Therefore, we have

$$
\mathbb{E}_{\Delta \mathbf{H}_n}\left\{\mathbf{T}_3\right\} = \mathbb{E}_{\Delta \mathbf{H}_n}\left\{\Delta \mathbf{H}_n \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \Delta \mathbf{H}_n^H\right\}
$$
$$
= \mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\Delta \widetilde{\mathbf{H}}_n \mathbf{\Lambda} \Delta \widetilde{\mathbf{H}}_n^H\right\},
$$

where $\Delta \widetilde{\mathbf{H}}_n = \Delta \mathbf{H}_n \mathbf{U}$, whose elements are also distributed as $\mathcal{CN}\left(0, \sigma_H^2\right)$ since $\mathbf{U}$ is a unitary matrix. Thus, the above expression can be simplified as follows:

$$
\mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\Delta \widetilde{\mathbf{H}}_n \mathbf{\Lambda} \Delta \widetilde{\mathbf{H}}_n^H\right\} = \mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\sum_{i=1}^{N_T} \tilde{\lambda}_i \Delta \widetilde{\mathbf{h}}_{n,i} \Delta \widetilde{\mathbf{h}}_{n,i}^H\right\},
$$

where $\Delta \widetilde{\mathbf{h}}_{n,i}$ denotes the ith column of $\Delta \widetilde{\mathbf{H}}_n$. Upon assuming that all the elements of $\Delta \mathbf{H}_n$ are independent and identically distributed, $\mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\Delta \widetilde{\mathbf{h}}_{n,i} \Delta \widetilde{\mathbf{h}}_{n,i}^H\right\} = \sigma_H^2 \mathbf{I}_{N_R}$, we have

$$
\mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\sum_{i=1}^{N_T} \tilde{\lambda}_i \Delta \widetilde{\mathbf{h}}_{n,i} \Delta \widetilde{\mathbf{h}}_{n,i}^H\right\} = \sum_{i=1}^{N_T} \tilde{\lambda}_i \mathbb{E}_{\Delta \widetilde{\mathbf{H}}_n}\left\{\Delta \widetilde{\mathbf{h}}_{n,i} \Delta \widetilde{\mathbf{h}}_{n,i}^H\right\}
$$
$$
= \sum_{i=1}^{N_T} \tilde{\lambda}_i \sigma_H^2 \mathbf{I}_{N_R}
$$
$$
= \sigma_H^2 \operatorname{Tr}\left(\mathbf{K} \mathbf{K}^H\right) \mathbf{I}_{N_R}.
$$

From the above results and (64), we conclude that

$$\mathbb{E}_{\Delta\mathbf{H}_n}\left\{\mathbf{H}_n\mathbf{K}\mathbf{K}^H\mathbf{H}_n^H\right\} = \mathbb{E}_{\Delta\mathbf{H}_n}\left\{\widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_n^H\right\} +$$

$$\mathbb{E}_{\Delta\mathbf{H}_n}\left\{\mathbf{T}_1\right\} + \mathbb{E}_{\Delta\mathbf{H}_n}\left\{\mathbf{T}_2\right\} + \mathbb{E}_{\Delta\mathbf{H}_n}\left\{\mathbf{T}_3\right\}$$

$$= \widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_n^H + \mathbf{0} + \mathbf{0} + \sigma_H^2\mathrm{Tr}\left(\mathbf{K}\mathbf{K}^H\right)\mathbf{I}_{N_R}$$

$$= \widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_n^H + \sigma_H^2\mathrm{Tr}\left(\mathbf{K}\mathbf{K}^H\right)\mathbf{I}_{N_R}.$$

Hence, the first result of Lemma 1 is proved.

The next result of the lemma can be proven by expanding the quantity $\mathbf{H}_n\mathbf{K}\mathbf{K}^H\mathbf{H}_j^H$ as

$$\mathbf{H}_n\mathbf{K}\mathbf{K}^H\mathbf{H}_j^H = \widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_j^H + \Delta\mathbf{H}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_j^H$$
$$+ \widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\Delta\mathbf{H}_j^H + \Delta\mathbf{H}_n\mathbf{K}\mathbf{K}^H\Delta\mathbf{H}_j^H. \tag{65}$$

As explained previously, the expected value of the second and third terms above is equal to zero. The expected value of the fourth term is given by

$$\mathbb{E}_{\Delta\mathbf{H}_n,\Delta\mathbf{H}_j}\left\{\Delta\mathbf{H_n}\mathbf{K}\mathbf{K}^H\Delta\mathbf{H}_j^H\right\}$$
$$\stackrel{a}{=} \mathbb{E}_{\Delta\mathbf{H}_n}\left\{\Delta\mathbf{H_n}\right\}\mathbf{K}\mathbf{K}^H\mathbb{E}_{\Delta\mathbf{H}_j}\left\{\Delta\mathbf{H}_j^H\right\}$$
$$= \mathbf{0},$$

where (a) follows from the fact that $\Delta\mathbf{H}_n$ and $\Delta\mathbf{H}_j$ are independent of each other. This implies that

$$\mathbb{E}_{\Delta\mathbf{H}_n,\Delta\mathbf{H}_j}\left\{\mathbf{H}_n\mathbf{K}\mathbf{K}^H\mathbf{H}_j^H\right\} = \widehat{\mathbf{H}}_n\mathbf{K}\mathbf{K}^H\widehat{\mathbf{H}}_j^H,$$

which proves the second result of Lemma 1. $\qquad\square$

Now, upon employing Lemma 1 in (58), one obtains the desired expression given at the top of this page.

REFERENCES

[1] S. Andreev, C. Dobre, and P. Misra, "Internet of things and sensor networks," *IEEE Communications Magazine*, vol. 58, no. 4, pp. 74–74, 2020.

[2] N. Ahmed, D. De, and I. Hussain, "Internet of things (IoT) for smart precision agriculture and farming in rural areas," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4890–4899, 2018.

[3] H. Harb, A. Mansour, A. Nasser, E. M. Cruz, and I. de la Torre Díez, "A sensor-based data analytics for patient monitoring in connected healthcare applications," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 974–984, 2021.

[4] M. Carminati, O. Kanoun, S. L. Ullo, and S. Marcuccio, "Prospects of distributed wireless sensor networks for urban environmental monitoring," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 6, pp. 44–52, 2019.

[5] O. Cetinkaya, D. Balsamo, and G. V. Merrett, "Internet of MIMO things: UAV-assisted wireless-powered networks for future smart cities," *IEEE Internet of Things Magazine*, vol. 3, no. 1, pp. 8–13, 2020.

[6] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[7] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.

[8] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[9] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, 2016.

[10] Q. Shi and M. Hong, "Spectral efficiency optimization for millimeter wave multiuser MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 3, pp. 455–468, 2018.

[11] 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.

[12] 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.

[13] N. K. D. 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.

[14] K. P. Rajput, Y. Verma, N. K. D. Venkategowda, A. K. Jagannatham, and P. K. Varshney, "Linear MMSE precoder combiner designs for decentralized estimation in wireless sensor networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[15] Q. Zhou, D. Li, S. Kar, L. M. Huie, H. V. Poor, and S. Cui, "Learning-based distributed detection-estimation in sensor networks with unknown sensor defects," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 130–145, 2017.

[16] V. Gupta and S. De, "An energy-efficient edge computing framework for decentralized sensing in WSN-Assisted IoT," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4811–4827, 2021.

[17] A. S. Leong, S. Dey, and J. S. Evans, "Asymptotics and power allocation for state estimation over fading channels," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 1, pp. 611–633, 2011.

[18] A. S. Leong, S. Dey, G. N. Nair, and P. Sharma, "Power allocation for outage minimization in state estimation over fading channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3382–3397, 2011.

[19] J. Akhtar and K. Rajawat, "Distributed sequential estimation in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 86–100, 2017.

[20] X. Cheng, P. Khanduri, B. Chen, and P. K. Varshney, "Joint collaboration and compression design for distributed sequential estimation in a wireless sensor network," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5448–5462, 2021.

[21] N. K. D. Venkategowda, H. Lee, and I. Lee, "Joint transceiver designs for MSE minimization in MIMO wireless powered sensor networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5120–5131, 2018.

[22] N. K. Venkategowda, H. Lee, and I. Lee, "Data precoding and energy transmission for parameter estimation in mimo wireless powered sensor networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–5.

[23] Z. Zhu, Z. Wang, Y. Lin, P. Liu, W. Hao, Z. Wang, and I. Lee, "Robust max-min fair beamforming of secrecy SWIPT IoT systems under a non-linear EH model," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–6.

[24] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. López-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3005–3017, 2015.

[25] A. Shirazinia, S. Dey, D. Ciuonzo, and P. Salvo Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2499–2512, 2016.

[26] D. Ciuonzo, P. S. Rossi, and S. Dey, "Massive MIMO channel-aware decision fusion," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 604–619, 2015.

[27] Y. Liu and C. Li, "Secure distributed estimation over wireless sensor networks under attacks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1815–1831, 2018.

[28] H. Liu, S. Wang, X. Zhao, S. Gong, N. Zhao, and T. Q. Quek, "Analog–digital hybrid transceiver optimization for data aggregation in IoT networks," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11 262–11 275, 2020.

[29] H. Liu, S. Wang, S. Gong, N. Zhao, J. An, and T. Q. S. Quek, "Hybrid LMMSE transceiver optimization for distributed IoT sensing networks with different levels of synchronization," *IEEE Internet of Things Journal*, vol. 8, no. 19, pp. 14 458–14 470, 2021.

[30] S. H. Chae, C. Jeong, and S. H. Lim, "Simultaneous wireless information and power transfer for Internet of Things sensor networks," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2829–2843, 2018.

[31] K. P. Rajput, A. Kumar, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Bayesian learning-based linear decentralized sparse parameter estima-

tion in MIMO wireless sensor networks relying on imperfect CSI," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6236–6250, 2021.

[32] K. P. Rajput, M. F. Ahmed, N. K. D. Venkategowda, A. K. Jagannatham, G. Sharma, and L. Hanzo, "Robust decentralized and distributed estimation of a correlated parameter vector in MIMO-OFDM wireless sensor networks," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6894–6908, 2021.

[33] 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, 2017.

[34] 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.

[35] K. Xu, Y. Cai, M. Zhao, Y. Niu, and L. Hanzo, "MIMO-aided nonlinear hybrid transceiver design for multiuser mmwave systems relying on Tomlinson-Harashima precoding," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6943–6957, 2021.

[36] L. Jiang and H. Jafarkhani, "MmWave amplify-and-forward MIMO relay networks with hybrid precoding/combining design," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1333–1346, 2019.

[37] Y. Cai, Y. Xu, Q. Shi, B. Champagne, and L. Hanzo, "Robust joint hybrid transceiver design for millimeter wave full-duplex MIMO relay systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1199–1215, 2019.

[38] M.-M. Zhao, Y. Cai, M.-J. Zhao, Y. Xu, and L. Hanzo, "Robust joint hybrid analog-digital transceiver design for full-duplex mmwave multicell systems," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4788–4802, 2020.

[39] C. Wang, Z. Li, T.-X. Zheng, H. Chen, and X.-G. Xia, "Robust hybrid precoding design for securing millimeter-wave IoT networks under secrecy outage constraint," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 13 024–13 038, 2021.

[40] K. P. Rajput, S. Srivastava, R. Akarapu, P. Maity, and A. K. Jagannatham, "Hybrid precoder designs for decentralized parameter estimation in millimeter wave (mmWave) sensor networks with perfect and imperfect CSI," *Physical Communication*, vol. 47, p. 101377, 2021.

[41] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

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