Joint Beamforming and Jamming Design for mmWave Information Surveillance Systems

Yunlong Cai, *Senior Member, IEEE*, Cunzhuo Zhao, Qingjiang Shi, Geoffrey Ye Li, *Fellow, IEEE*, and Benoit Champagne, *Senior Member, IEEE*

***Abstract*— This paper addresses the design of joint beamforming and jamming for a millimeter wave (mmWave) information surveillance system where a suspicious transmitter in the network sends messages to a suspicious receiver under the supervision of a surveillant controller (SC), which not only carries out the duty of a base station or other access point, but also legitimately monitor the suspicious link. Specifically, we seek to maximize the effective monitoring rate for information surveillance by jointly optimizing the analog transmit and receive beamforming vectors of the suspicious link, the analog jamming and monitoring beamforming vectors at the SC and the jamming signal's power level under transmit power, successful monitoring, and self-interference power constraints at the SC, along with unit modulus constraint on the elements of the radio frequency analog beamforming vectors. The resulting optimization problem is quite challenging due to the tight coupling of the design variables in the objective function and constraints. To solve it, we develop a novel algorithm based on the penalty dual decomposition (PDD) technique, where the exacting constraints are penalized and dualized into the objective function as augmented Lagrangian components. The proposed PDD-based algorithm performs double-loop iterations, i.e., the inner loop resorts to the concave–convex procedure to update the optimization variables; while the outer loop adjusts the Lagrange multipliers and penalty parameter of the augmented Lagrangian cost function. We show that the proposed PDD-based joint beamforming and jamming algorithm converges to a stationary solution of the original problem. Based on our simulation results, the proposed algorithm achieves significantly better performance than the conventional beamforming and jamming algorithms.**

***Index Terms*— Analog beamforming, multiple antennas, mmWave system, wireless information surveillance, penalty dual decomposition.**

Manuscript received September 15, 2017; revised January 30, 2018; accepted February 16, 2018. Date of publication April 11, 2018; date of current version October 18, 2018. The work of Y. Cai was supported in part by the National Natural Science Foundation of China under Grant 61471319 and in part by the Fundamental Research Funds for the Central Universities. The work of Q. Shi was supported in part by the National Nature Science Foundation of China under Grant 61671411 and in part by the Fundamental Research Funds for the Central Universities. *(Corresponding author: Qingjiang Shi.)*

Y. Cai and C. Zhao are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: ylcai@zju.edu.cn; zhaocz@zju.edu.cn).

Q. Shi is with the School of Software Engineering, Tongji University, Shanghai 201804, China (e-mail: qing.j.shi@gmail.com).

G. Y. Li is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: liye@ece.gatech.edu).

B. Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: benoit.champagne@mcgill.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2018.2825718

I. Introduction

MILLIMETER wave (mmWave) communications with operating frequency in the range from 30 to 300GHz have become a key enabling technology for 5G networks and beyond, as they hold the potential to mitigate spectrum shortage and significantly increase data transmission rates [1]–[4]. Moreover, the very small wavelength of mmWave signals enables the packing of a large number of miniaturized antennas in physical devices with small dimensions. Hence, the application of multiple antenna techniques in mmWave systems is currently the focus of great attention within the research community [5]–[8]. Due to the large fabrication costs of radio frequency (RF) chains and analog-to-digital (A/D) converters at such high frequencies, the traditional fully digital transceiver architecture cannot be utilized in mmWave-based multiple antenna systems. As an alternative, a hybrid transceiver structure, consisting of a baseband multi-channel processor followed by RF analog beamformers, has been widely adopted since it allows to employ a much smaller number of RF chains than that of antenna elements.

A common method for hybrid transceiver implementation (in particular for analog beamformers) is to use analog phase shifters [9]–[12]. In this case, a unit modulus constraint will be imposed on the elements of the analog beamforming matrices. Some distinctive algorithms have been proposed in [11]–[15] to design RF analog beamformers. The studies in [11], [12] developed an orthogonal matching pursuit (OMP) algorithm, where the columns of the analog beamforming matrix can be selected from the array response vectors of the channel matrices. As an extension of the OMP, an analog beamforming design algorithm based on channel matching has been proposed in [13]. In addition, to optimize the analog beamformers under the unit modulus constraint, an alternating algorithm based on manifold optimization has been proposed in [14] and [15].

The explosive growth in the number and types of mobile devices in current and future wireless networks requires innovative physical layer security (PLS) techniques to enhance the secrecy performance in the presence of malicious eavesdroppers [16]–[18]. Besides, the wide availability of extensive wireless services at low cost facilitates the cooperation between criminals or terrorists, therefore posing significant threats on national security. Thus, in order to resist crimes or terrorist attacks, there is a growing need for government agencies worldwide to legitimately monitor (i.e. proactive

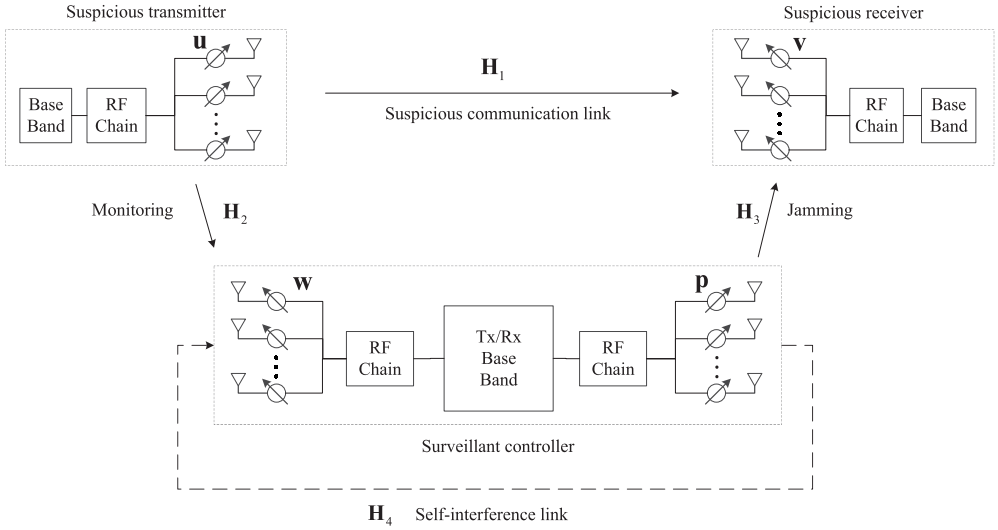

Fig. 1. Millimeter wave surveillance system.

eavesdropping) any suspicious communication links and detect abnormal behaviors within commercial wireless networks. As mentioned in [19], a terrorist surveillance program has been conducted by the National Security Agency (NSA) of the USA to legitimately monitor wireless devices, and the US military has adopted jammers to protect travelling convoys from mobile phone triggered roadside bombs. This recent trend has motivated a paradigm shift from traditional PLS to so-called wireless surveillance systems. In the latter, a legitimate eavesdropper (also referred to as a surveillant controller or monitor in the sequel) is employed to control and supervise the transmission between the suspicious transmitter and receiver for successful monitoring. A direct approach for surveillance is passive eavesdropping, where the legitimate monitor remains silent to listen to the suspicious links, and does not transmit any jamming signals. This method is suitable when the legitimate monitor is close to the suspicious transmitter, so that the transmitted information can be decoded successfully, but does not apply when the transmitter is far-away. To overcome this limitation, several alternative schemes have been proposed for wireless information surveillance [20]–[26]. The design of a full-duplex legitimate monitor is addressed in [20], where the aim is to optimally allocate the jamming signal power in order to maximize either the monitoring non-outage probability or the monitoring rate under perfect self-interference cancelation and global channel state information (CSI) for all relevant links. In [21], the optimal jamming power at the legitimate monitor is selected to maximize the average monitoring rate over Rayleigh fading channels, given only the channel distribution information (CDI). This work is further extended in [22] by considering the use of multiple antennas at the surveillant controller (SC). Subsequently, [23] considers the design and performance analysis of a legitimate multi-antenna surveillance system, where the jamming power and beamforming vectors at a multi-antenna SC are adjusted to maximize the monitoring non-outage probability. Three possible spoofing relay strategies are proposed in [24] to maximize the achievable monitoring rate. Furthermore, information surveillance in a relaying sub-network is investigated in [25] and [26] for the half-duplex mode of operation.

Considering that mmWave is a promising technology for next generation wireless communications, which offers the potential communication channels with larger bandwidth and cellular networks with higher capacity, there is a great need to investigate information surveillance within the framework of mmWave systems. However, due to hardware constraints of mmWave transmitters and receivers, the design of mmWave information surveillance systems is far more challenging than for traditional wireless systems operating at lower frequencies, and to the best of our knowledge, it has not yet been addressed in the literature.

In this paper, we investigate information surveillance in a device-to-device mmWave system where a suspicious transmitter in the network sends messages to a suspicious receiver under the supervision of a surveillant controller (SC), as shown in Fig. 1. We focus on a single data stream transmission scenario and assume that the suspicious transmitter, suspicious receiver and legitimate SC are all equipped with multiple antennas. The transmit and receive antennas at the SC are assumed to be separately located, so that the self-interference can be adequately limited by means of state-of-the-art self-interference cancellation techniques [27]–[30]. In addition, due to practical constraints on mmWave hardware mentioned earlier, the transmit and receive antenna arrays in the system are assumed to each comprise a single RF chain. The SC not only carries out the duty of a base station but also legitimately monitors the suspicious transmission link. We consider the joint optimization problem for the suspicious transmitter and receiver's analog beamforming vectors, the SC's analog jamming and monitoring beamforming vectors, and the jamming signal's power level, where the aim is to maximize the effective monitoring rate under transmit power, successful monitoring and self-interference power constraints at the SC, as well as

a unit modulus constraint on the entries of all the analog beamforming vectors.

It is difficult (if not impossible), to globally solve the resulting optimization problem since the optimized variables are highly coupled in both the objective function and constraints. By introducing a set of auxiliary variables and equality constraints, we first recast this problem into an equivalent but more tractable form. We then propose a new algorithm based on the innovative penalty dual decomposition (PDD) technique [31], [32] to handle the problematic coupled constraints and jointly optimize the analog beamforming vectors and the jamming signal's power level. With the aid of the PDD optimization method, we penalize and dualize the newly introduced equality constraints into the objective function as augmented Lagrangian components [33], [34]. The resulting augmented Lagrangian problem is then optimized via a novel iterative algorithm with a double-loop structure. In the inner loop, we resort to the concave-convex procedure (CCCP) [35]–[37] to update the optimization variables in a block coordinate descent (BCD) fashion; while in the outer loop, we adjust the Lagrange multipliers and penalty parameter of the augmented Lagrangian cost function. The proposed PDD-based algorithm is shown to converge to a stationary solution of the original optimization problem. We also analyze the computational complexity of the proposed algorithm.

This paper is organized as follows. Section II briefly describes the system model and formulates the optimization problem of interest. Section III presents the new PDD-based joint beamforming and jamming design algorithm and discusses its complexity and convergence behavior. Supporting simulation results are presented in Section IV and finally, conclusions are drawn in Section V.

Notations: Scalars, vectors, and matrices are respectively denoted by lower case, boldface lower case, and boldface upper case letters, respectively. $\mathbf{I}_m$ represents an identity matrix with dimension $m \times m$, and $\mathbf{0}_{m\times n}$ denotes an all-zero matrix with dimension $m \times n$. For a matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$ and $\|\mathbf{A}\|$ denote its transpose, conjugate, conjugate transpose and Frobenius norm, respectively. For a square matrix $\mathbf{A}$, $\mathrm{Tr}(\mathbf{A})$ and $\mathbf{A}(i,i)$ denote its trace and the ith diagonal element, respectively. For a vector $\mathbf{a}$, $\|\mathbf{a}\|$ represents its Euclidean norm and $\mathbf{a}(i)$ denotes the ith element. $E[.]$ denotes the statistical expectation. $\Re\{.\}$ ($\Im\{.\}$) denotes the real (imaginary) part of a variable. $\mathrm{diag}\{a_1,\ldots,a_n\}$ denotes a square diagonal matrix with $a_1,\ldots,a_n$ on the diagonal. The operator $\mathrm{vec}(\cdot)$ stacks the elements of a matrix in one long column vector. $|\cdot|$ denotes the absolute value of a complex scalar. The symbol $\otimes$ denotes the Kronecker product. $\mathbb{C}^{m\times n}$ ($\mathbb{R}^{m\times n}$) denotes the space of $m \times n$ complex (real) matrices. $\mathcal{CN}(\delta,\sigma^2)$ represents a complex Gaussian distribution with mean δ and variance σ^2.

II. System Model and Problem Formulation

In this section, we first introduce the mmWave surveillance system under study and then formulate the corresponding design problem.

A. System Model

As illustrated in Fig. 1, consider a device-to-device information surveillance system, where a suspicious transmitter communicates with a suspicious receiver under the supervision of an SC at mmWave frequency. All three entities are part of a larger wireless network, with the SC (which can be a base station or some other type of access point) serving as a legitimate monitor. We focus on a single data stream transmission scenario and assume that the suspicious transmitter and receiver are equipped with N_t and N_r antennas, respectively. The SC is equipped with M_r receive antennas for monitoring and M_t transmit antennas for sending jamming signals[1] and operates in a full-duplex mode, allowing to jam and monitor at the same time. Considering the large fabrication cost of RF chains in mmWave systems, we assume that the suspicious transmitter and receiver and the SC are all equipped with a single RF chain.

We define $\mathbf{u} \in \mathbb{C}^{N_t\times 1}$ as the analog transmit beamforming vector at the suspicious transmitter and $\mathbf{v} \in \mathbb{C}^{N_r\times 1}$ as the analog receive beamforming vector at the suspicious receiver, and define $\mathbf{w} \in \mathbb{C}^{M_r\times 1}$ and $\mathbf{p} \in \mathbb{C}^{M_t\times 1}$ as the analog monitoring and jamming beamforming vectors, respectively, at the SC. Since the transmit and receive beamformers are implemented by using low-cost phase shifters in the analog domain, every element of all the analog beamforming vectors follows the unit modulus constraint [9]–[12]. At a given symbol time instance, the received signal vector at the suspicious receiver, $\mathbf{y} \in \mathbb{C}^{N_r\times 1}$, can be expressed as

$$\mathbf{y} = \mathbf{H}_1\mathbf{u}b + \mathbf{H}_3\mathbf{p}f + \mathbf{n}_1 \tag{1}$$

where b denotes the symbol emitted by the suspicious transmitter, which is modeled as a zero-mean circular complex Gaussian random variable with variance $\sigma_b^2 = E[|b|^2]$, f denotes the jamming symbol sent by the SC, which is modeled as a circular complex Gaussian random variable with zero mean and variance $\sigma_f^2 = E[|f|^2]$, $\mathbf{H}_1 \in \mathbb{C}^{N_r\times N_t}$ and $\mathbf{H}_3 \in \mathbb{C}^{N_r\times M_t}$ denote the mmWave channel matrices[2] of the suspicious communication link and of the jamming link, respectively, and $\mathbf{n}_1 \in \mathbb{C}^{N_r\times 1}$ is the additive zero-mean circular complex Gaussian noise vector at the suspicious receiver with covariance matrix $E[\mathbf{n}_1\mathbf{n}_1^H] = \sigma_1^2\mathbf{I}_{N_r}$, where σ_1^2 denotes the noise variance.

We assume that the transmit and receive antennas at the SC are separately located. By properly using self-interference cancellation techniques [27]–[30], the residual self-interference to the SC's receive antennas, $\mathbf{n}_s$, can be modeled as a zero-mean Gaussian additive noise with covariance matrix $E[\mathbf{n}_s\mathbf{n}_s^H] = \sigma_s^2\mathbf{I}_{M_r}$ provided [30] the following constraint is met:

$$E[\|\mathbf{H}_4\mathbf{p}f\|^2] = \sigma_f^2\|\mathbf{H}_4\mathbf{p}\|^2 \le \delta, \tag{2}$$

[1]The SC purposely sends jamming signals to interfere with the suspicious transmission link, so as to decrease the achievable data rate at the suspicious receiver, while improving the monitoring ability of the SC, as further explain below.

[2]In this work, we consider a narrowband propagation where all the channels are assumed to be exhibit flat fading.

where $\mathbf{H}_4 \in \mathbb{C}^{M_r \times M_t}$ denotes the mmWave channel matrix of the residual self-interference link[3] between the transmit and the receive antennas at the SC and δ denotes the maximum tolerable self-interference power. Therefore, the received signal vector at the SC is given by

$$\mathbf{y}^{(M)} = \mathbf{H}_2 \mathbf{u} b + \mathbf{n}_2 + \mathbf{n}_s, \tag{3}$$

where $\mathbf{H}_2 \in \mathbb{C}^{M_r \times N_t}$ denotes the mmWave channel matrix of the monitoring link between the suspicious transmitter and the SC receiver, while $\mathbf{n}_2 \in \mathbb{C}^{M_r \times 1}$ denotes the additive zero-mean circular complex Gaussian noise vector at the SC with covariance matrix $E[\mathbf{n}_2 \mathbf{n}_2^H] = \sigma_2^2 \mathbf{I}_{M_r}$, where σ_2^2 denotes the noise variance.

The SC is intended to carry out the following duties:

1) In the role of *base station*, it provides the required signalling and algorithm implementation services for the suspicious device-to-device transmission link. Due to hardware limitation of mobile devices, the coefficients of the analog transmit and receive beamforming vectors, $\mathbf{u}$ and $\mathbf{v}$, are optimized at the SC and then fed to the suspicious transmitter and receiver through signalling channels for communications.
2) In the role of *legitimate monitor*, it optimizes the analog monitoring and jamming beamforming vectors at the SC, i.e., $\mathbf{w}$ and $\mathbf{p}$, and adjust the jamming power level σ_f in order to efficiently overhear the suspicious transmission link.

In this scenario, the channel capacity of the suspicious transmission link will be

$$R_D = \log\left(1 + \gamma_0\right), \tag{4}$$

where γ_0 denotes the signal-to-noise ratio (SNR) at the suspicious receiver, which can be expressed as

$$\gamma_0 = \frac{|\mathbf{v}^H \mathbf{H}_1 \mathbf{u}|^2 \sigma_b^2}{\sigma_1^2 N_r + \sigma_f^2 |\mathbf{v}^H \mathbf{H}_3 \mathbf{p}|^2}. \tag{5}$$

In (5), the unit modulus constraint of the analog beamforming vectors has been used. Similarly, the capacity of the monitoring channel will be

$$R_M = \log\left(1 + \gamma_1\right), \tag{6}$$

where γ_1 denotes the SNR at the SC and can be expressed as

$$\gamma_1 = \frac{|\mathbf{w}^H \mathbf{H}_2 \mathbf{u}|^2 \sigma_b^2}{\sigma_2^2 M_r + \sigma_s^2 M_r}. \tag{7}$$

From [22], if $R_M \geq R_D$, the effective monitoring rate will be $\tilde{R} = R_D$; otherwise, i.e. $R_M < R_D$, it is impossible for the SC to decode the information from the suspicious transmitter without errors and the effective monitoring rate will be $\tilde{R} = 0$.

Note that if the SC benefits from a better channel condition than the suspicious receiver, it can overhear the suspicious link without the need to send jamming signals. Otherwise, the SC has to send jamming signals or purposely degrade the beamforming vector $\mathbf{v}$ used by the suspicious receiver to make $R_M \geq R_D$, as needed for efficient monitoring. In this work, we propose a general joint beamforming and jamming algorithm in order to optimize the functionality of the SC in carrying the above tasks.

[3]Here, we assume that the SI power level is first mitigated to an acceptable level based on hardware-aided FD interference cancellation techniques, such as those exposed in [27]–[30]. Then, we jointly optimize the analog beamforming vectors and the jamming signal's power level with the proposed PDD-based algorithm to handle the residual SI.

B. Problem Formulation

The SC aims to jointly optimize the analog beamforming vectors, i.e., $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{p}$, and the jamming power level σ_f^2 so that the monitoring rate is maximized. By taking the successful monitoring condition into account, i.e. $R_M \geq R_D$, the problem can be formulated as

$$\max_{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f} \quad \log\left(1 + \frac{|\mathbf{v}^H \mathbf{H}_1 \mathbf{u}|^2 \sigma_b^2}{\sigma_1^2 N_r + \sigma_f^2 |\mathbf{v}^H \mathbf{H}_3 \mathbf{p}|^2}\right) \tag{8a}$$

$$\text{s.t. } 0 \leq \sigma_f \leq \sqrt{\frac{P_t}{M_t}} \tag{8b}$$

$$\frac{|\mathbf{w}^H \mathbf{H}_2 \mathbf{u}|^2}{\sigma_2^2 \ M_r + \sigma_s^2 M_r} \geq \frac{|\mathbf{v}^H \mathbf{H}_1 \mathbf{u}|^2}{\sigma_1^2 N_r + \sigma_f^2 |\mathbf{v}^H \mathbf{H}_3 \mathbf{p}|^2} \tag{8c}$$

$$\sigma_f^2 \|\mathbf{H}_4 \mathbf{p}\|^2 \leq \delta \tag{8d}$$

$$|\mathbf{u}(i)| = 1, |\mathbf{v}(i)| = 1, |\mathbf{w}(i)| = 1, |\mathbf{p}(i)| = 1, \quad \forall i, \tag{8e}$$

where constraints (8b) and (8c) reflect the bounded transmit power budget of the SC and the successful monitoring condition, respectively. The suppressible self-interference power constraint is shown in (8d). The last four sets of unit modulus constraints in (8e) are due to the fact that all the analog transmit and receive beamformers are implemented using low-cost phase shifters.

As we can see from (8), the design variables are highly coupled in the objective function and constraints and accordingly, it is in general difficult to find the global optimal solution. In the following section, we propose an efficient joint design algorithm that can find a local stationary solution of the original problem by applying proper convex optimization techniques. In order to implement the proposed algorithm, we assume that the CSI of all links is available at the SC. In practice, the residual self-interference channel matrix $\mathbf{H}_4$ and the monitoring channel matrix $\mathbf{H}_2$ can be estimated at the SC. Assuming that the suspicious user, which is part of the network, respects the transmission protocols in place to avoid being detected, the suspicious transmission channel matrix $\mathbf{H}_1$ and jamming channel matrix $\mathbf{H}_3$ will be estimated at the suspicious receiver using standard techniques and fed back to the SC (which operates as a legitimate base station) through dedicated signaling channels.

III. JOINT BEAMFORMING AND JAMMING DESIGN

In this section, we propose to employ the PDD optimization framework, as summarized in Appendix A, to address problem (8). By applying the PDD technique, we first obtain an alternative formulation to (8) where a number of augmented Lagrangian terms are incorporated into the objective in order to handle highly coupled constraints. We then develop an efficient CCCP algorithm to solve the augmented Lagrangian

problem in the inner loop of the PDD algorithm. In particular, in order to approximate this latter problem as a convex one, a locally tight lower bound for the objective function is derived while linearization is employed to approximate the nonconvex constraint. Finally, we summarize the PDD-based joint beamforming and jamming algorithm and analyze its computational complexity.

A. Augmented Lagrangian Problem

Let $\mathbf{h}_1 \triangleq \text{vec}(\mathbf{H}_1)$ and $\mathbf{h}_3 \triangleq \text{vec}(\mathbf{H}_3)$, so that we have $\mathbf{v}^H\mathbf{H}_1\mathbf{u} = (\mathbf{u}^T \otimes \mathbf{v}^H)\mathbf{h}_1$ and $\mathbf{v}^H\mathbf{H}_3\mathbf{p} = (\mathbf{p}^T \otimes \mathbf{v}^H)\mathbf{h}_3$, and define $\tilde{\mathbf{A}} \triangleq \mathbf{h}_1\mathbf{h}_1^H$, $\tilde{\mathbf{B}} \triangleq \frac{\sigma_s^2}{M_t}\mathbf{I} + \sigma_f^2\mathbf{h}_3\mathbf{h}_3^H$ and $\bar{\sigma} \triangleq \sigma_2^2 M_r + \sigma_s^2 M_r$. By introducing a set of auxiliary variables $\{\tilde{t} \geq 0, t, \tilde{\mathbf{u}}, \tilde{\mathbf{p}}, x, \bar{\mathbf{p}}\}$, problem (8) can be equivalently converted into[4]

$$\max_{\substack{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f,\\ \bar{\mathbf{p}},x,\tilde{\mathbf{u}},\tilde{\mathbf{p}},t,\tilde{t}\geq 0}} \quad t \tag{9a}$$

$$\text{s.t. } 0 \leq \sigma_f \leq \sqrt{\frac{P_t}{M_t}} \tag{9b}$$

$$|x|^2 \geq \bar{\sigma} t^2 \tag{9c}$$

$$\|\mathbf{H}_4\bar{\mathbf{p}}\|^2 \leq \delta \tag{9d}$$

$$|\mathbf{u}(i)| = 1, \quad |\mathbf{v}(i)| = 1, \ |\mathbf{w}(i)| = 1, \ |\mathbf{p}(i)| = 1, \ \forall i \tag{9e}$$

$$\|\tilde{\mathbf{A}}^{\frac{1}{2}}\tilde{\mathbf{u}}\| = t\|\tilde{\mathbf{B}}^{\frac{1}{2}}\tilde{\mathbf{p}}\|, \quad \bar{\mathbf{p}} = \sigma_f\mathbf{p}, \ x = \mathbf{w}^H\mathbf{H}_2\mathbf{u}, \tag{9f}$$

$$\tilde{\mathbf{u}}^H = \mathbf{u}^T \otimes \mathbf{v}^H, \quad \tilde{\mathbf{p}}^H = \mathbf{p}^T \otimes \mathbf{v}^H, \ t = \tilde{t}, \tag{9g}$$

where the equality constraints shown in (9f) and (9g) are introduced to apply the PDD technique.

The equality constraint $\tilde{\mathbf{u}}^H = \mathbf{u}^T \otimes \mathbf{v}^H$ is equivalent to $\tilde{\mathbf{u}}^H\mathbf{A}_k = \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})$, $\forall k \in \{1, \ldots, N_t\}$, while the equality constraint $\tilde{\mathbf{p}}^H = \mathbf{p}^T \otimes \mathbf{v}^H$ is equivalent to $\tilde{\mathbf{p}}^H\mathbf{B}_k = \mathbf{v}^H(\mathbf{b}_k^T\mathbf{p})$, $\forall k \in \{1, \ldots, M_t\}$, where we define the following selection matrices and vectors:

$$\begin{aligned}
\mathbf{A}_k &\triangleq [\mathbf{0}_{N_r\times(k-1)N_r}, \mathbf{I}_{N_r}, \mathbf{0}_{N_r\times(N_t-k)N_r}]^T \\
\mathbf{a}_k &\triangleq [\mathbf{0}_{1\times(k-1)}, 1, \mathbf{0}_{1\times(N_t-k)}]^T \\
\mathbf{B}_k &\triangleq [\mathbf{0}_{N_r\times(k-1)N_r}, \mathbf{I}_{N_r}, \mathbf{0}_{N_r\times(M_t-k)N_r}]^T \\
\mathbf{b}_k &\triangleq [\mathbf{0}_{1\times(k-1)}, 1, \mathbf{0}_{1\times(M_t-k)}]^T.
\end{aligned} \tag{10}$$

Based on the PDD optimization method (see Appendix A), we take all the equality constraints in (9g) into account by augmenting the objective function with Lagrange multipliers λ_1, $\boldsymbol{\lambda}_2$, λ_3, $\{\boldsymbol{\lambda}_{4,k}\}$, $\{\boldsymbol{\lambda}_{5,k}\}$ and λ_6 along with a penalty parameter ρ. Therefore, the augmented Lagrangian problem can be formulated as (11), as shown at the top of the next page. The proposed PDD-based algorithm exhibits a double-loop structure, where the outer loop updates the dual variables and the penalty parameter while the inner loop seeks to optimize the primal variables by solving problem (11). The key to the PDD method is the inner iterations for solving augmented Lagrangian problems.

[4]Note that here we used the fact that $\tilde{\mathbf{p}}^H\tilde{\mathbf{p}} = M_tN_r$, which follows from the unit modulus constraint for each element of the analog beamforming vectors.

B. Proposed CCCP Algorithm for Solving Problem (11)

It is challenging to deal with problem (11) directly due to the difficulties caused by the term $\left(\|\tilde{\mathbf{A}}^{\frac{1}{2}}\tilde{\mathbf{u}}\| - t\|\tilde{\mathbf{B}}^{\frac{1}{2}}\tilde{\mathbf{p}}\| + \rho\lambda_1\right)^2$ in the objective function and by the nonconvex constraint in (9c). In the following, we develop an efficient CCCP algorithm to solve (11) in the inner loop of the PDD optimization framework. In order to obtain a more tractable and convex problem, we first find a locally tight lower bound for the objective function and then linearize the nonconvex constraint based on the CCCP concept [35]–[37].

Let us define the following real vectors and matrices: $\tilde{\mathbf{u}}_{eq} \triangleq [\Re\{\tilde{\mathbf{u}}^T\}, \Im\{\tilde{\mathbf{u}}^T\}]^T$, $\tilde{\mathbf{p}}_{eq} \triangleq [\Re\{\tilde{\mathbf{p}}^T\}, \Im\{\tilde{\mathbf{p}}^T\}]^T$,

$$\mathbf{A}_{eq} \triangleq \begin{bmatrix} \Re\{\tilde{\mathbf{A}}\} & -\Im\{\tilde{\mathbf{A}}\} \\ \Im\{\tilde{\mathbf{A}}\} & \Re\{\tilde{\mathbf{A}}\} \end{bmatrix}, \tag{12}$$

$$\mathbf{B}_{eq} \triangleq \begin{bmatrix} \Re\{\tilde{\mathbf{B}}\} & -\Im\{\tilde{\mathbf{B}}\} \\ \Im\{\tilde{\mathbf{B}}\} & \Re\{\tilde{\mathbf{B}}\} \end{bmatrix}. \tag{13}$$

Then problem (11) can be equivalently formulated as (14), as shown at the top of the next page, where $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{D}_1$ and $\mathbf{D}_2$ are selection matrices defined as $\mathbf{C}_1 \triangleq [\mathbf{I}_{N_tN_r}, \mathbf{0}_{N_tN_r}]$, $\mathbf{C}_2 \triangleq [\mathbf{0}_{N_tN_r}, \mathbf{I}_{N_tN_r}]$, $\mathbf{D}_1 \triangleq [\mathbf{I}_{M_tN_r}, \mathbf{0}_{M_tN_r}]$, $\mathbf{D}_2 \triangleq [\mathbf{0}_{M_tN_r}, \mathbf{I}_{M_tN_r}]$.

First, to find a tight bound on the term $2t\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|$ in the objective function of problem (14), we need the following lemma, which can be directly proved by the Cauchy-Schwartz inequality.

Lemma 1: For real vectors $\mathbf{x}$, $\mathbf{y}$, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$, the following inequality holds true:

$$\|\mathbf{x}\|\|\mathbf{y}\| \geq \frac{1}{\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|}\mathbf{x}^T\tilde{\mathbf{x}}\tilde{\mathbf{y}}^T\mathbf{y}, \quad \forall\tilde{\mathbf{x}} \neq \mathbf{0}, \ \tilde{\mathbf{y}} \neq \mathbf{0}, \ \mathbf{x}, \ \mathbf{y}, \tag{15}$$

with equality satisfied at $\mathbf{x} = \tilde{\mathbf{x}}$ and $\mathbf{y} = \tilde{\mathbf{y}}$.

Let $\tilde{\mathbf{u}}_{eq}^i$ and $\tilde{\mathbf{p}}_{eq}^i$ denote the values of $\tilde{\mathbf{u}}_{eq}$ and $\tilde{\mathbf{p}}_{eq}$ in the ith iteration of the CCCP algorithm. By applying Lemma 1 we obtain the following locally tight lower bound:

$$2t\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\| \geq 2t\tilde{\mathbf{u}}_{eq}^T\boldsymbol{\Omega}\tilde{\mathbf{p}}_{eq} \tag{16}$$

where

$$\boldsymbol{\Omega} \triangleq \frac{\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}^i\tilde{\mathbf{p}}_{eq}^{iT}\mathbf{B}_{eq}}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|}. \tag{17}$$

Let us now turn our attention to the derivation of bounds for the terms $2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|$ and $2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|$ in the objective function. Note that the multiplier λ_1 can be either negative or positive, and each case leads to slight variation in the bounding approach.

In the case of $\lambda_1 \geq 0$, noting that $\|\tilde{\mathbf{p}}_{eq}\| = \sqrt{M_tN_r}$, and using Lemma 1, we obtain the following

$$2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\| = \frac{2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|\|\tilde{\mathbf{p}}_{eq}\|}{\sqrt{M_tN_r}} \geq \frac{2\rho\lambda_1 \ t\tilde{\mathbf{p}}_{eq}^T\boldsymbol{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_tN_r}} \tag{18}$$

where

$$\boldsymbol{\Pi} \triangleq \frac{\mathbf{B}_{eq}\tilde{\mathbf{p}}_{eq}^i\tilde{\mathbf{p}}_{eq}^{iT}}{\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|\|\tilde{\mathbf{p}}_{eq}^i\|}. \tag{19}$$

$$\max_{\substack{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f,\\ \bar{\mathbf{p}},x,\tilde{\mathbf{u}},\tilde{\mathbf{p}},t,\tilde{t}\geq 0}} \; t-\frac{1}{2\rho}\Big(\big(\|\tilde{\mathbf{A}}^{\frac{1}{2}}\tilde{\mathbf{u}}\|-t\|\tilde{\mathbf{B}}^{\frac{1}{2}}\tilde{\mathbf{p}}\|+\rho\lambda_1\big)^2+\|\bar{\mathbf{p}}-\sigma_f\mathbf{p}+\rho\boldsymbol{\lambda}_2\|^2+|x-\mathbf{w}^H\mathbf{H}_2\mathbf{u}+\rho\lambda_3|^2$$
$$+\sum_{k=1}^{N_t}\|\tilde{\mathbf{u}}^H\mathbf{A}_k-\mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})+\rho\boldsymbol{\lambda}_{4,k}\|^2+\sum_{k=1}^{M_t}\|\tilde{\mathbf{p}}^H\mathbf{B}_k-\mathbf{v}^H(\mathbf{b}_k^T\mathbf{p})+\rho\boldsymbol{\lambda}_{5,k}\|^2+|t-\tilde{t}+\rho\lambda_6|^2\Big)$$
$$\text{s.t. } (9b)-(9e). \tag{11}$$

$$\max_{\substack{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f,\\ \bar{\mathbf{p}},x,\tilde{\mathbf{u}}_{eq},\tilde{\mathbf{p}}_{eq},t,\tilde{t}\geq 0}} \; t-\frac{1}{2\rho}\big(\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2+t^2\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2-2t\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|+2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|-2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|$$
$$+\|\bar{\mathbf{p}}-\sigma_f\mathbf{p}+\rho\boldsymbol{\lambda}_2\|^2+|x-\mathbf{w}^H\mathbf{H}_2\mathbf{u}+\rho\lambda_3|^2+\sum_{k=1}^{N_t}\|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k-j\mathbf{C}_2\mathbf{A}_k)-\mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})+\rho\boldsymbol{\lambda}_{4,k}\|^2$$
$$+\sum_{k=1}^{M_t}\|\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k-j\mathbf{D}_2\mathbf{B}_k)-\mathbf{v}^H(\mathbf{b}_k^T\mathbf{p})+\rho\boldsymbol{\lambda}_{5,k}\|^2+|t-\tilde{t}+\rho\lambda_6|^2\big)$$
$$\text{s.t. } (9b)-(9e). \tag{14}$$

In order to handle the term $2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|$, we will use the following lemma, which follows from the fact that $\|\mathbf{x}\|=\sqrt{\|\mathbf{x}\|^2}$ is a concave function of $\|\mathbf{x}\|^2$.

Lemma 2: For real vectors $\mathbf{x}$ and $\tilde{\mathbf{x}}$, the following inequality holds true:

$$\|\mathbf{x}\|\leq\frac{1}{2\|\tilde{\mathbf{x}}\|}\|\mathbf{x}\|^2+\frac{1}{2}\|\tilde{\mathbf{x}}\|,\quad\forall\tilde{\mathbf{x}}\neq\mathbf{0},\mathbf{x}, \tag{20}$$

with equality satisfied at $\mathbf{x}=\tilde{\mathbf{x}}$.

By using Lemma 2 we obtain the following bound

$$2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|\leq\frac{\rho\lambda_1}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|}\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2+\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|. \tag{21}$$

In the case of $\lambda_1<0$, we can find a locally tight bounds for $2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|$ and $2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|$ by following the same approach. Since $\|\tilde{\mathbf{u}}_{eq}\|=\sqrt{N_tN_r}$, by using Lemma 1 we have

$$-2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|=\frac{-2\rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|\|\tilde{\mathbf{u}}_{eq}\|}{\sqrt{N_tN_r}}\geq\frac{-2\rho\lambda_1\tilde{\mathbf{u}}_{eq}^T\tilde{\boldsymbol{\Pi}}\tilde{\mathbf{u}}_{eq}}{\sqrt{N_tN_r}} \tag{22}$$

where

$$\tilde{\boldsymbol{\Pi}}\triangleq\frac{\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}^i\tilde{\mathbf{u}}_{eq}^{iT}}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|\|\tilde{\mathbf{u}}_{eq}^i\|}. \tag{23}$$

By applying Lemma 2, we obtain

$$-2\rho\lambda_1 t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|\leq\frac{-\rho\lambda_1\ t}{\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|}\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2-\rho\lambda_1\ t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|. \tag{24}$$

Thus, in the case $\lambda_1\geq 0$, employing (16), (18) and (21), a locally tight lower bound on the objective function of problem (14) around the current point $\{\tilde{\mathbf{p}}_{eq}^i,\tilde{\mathbf{u}}_{eq}^i\}$ is expressed as (25), as shown at the bottom of the next page.

In the case $\lambda<0$, by using (16), (22) and (24), we obtain a locally tight lower bound on the objective function as (26), as shown at the bottom of the next page.

Besides, we note that constraint (9c) can be seen as a difference of convex (DC) functions, which we seek to transform into a convex constraint. Based on the CCCP concept [35]–[37], by linearizing the term $|x|^2$ in the ith iteration around the current point x^i, the following approximated convex constraint is obtained:

$$\bar{\sigma}t^2\leq x^{i*}x+x^*x^i-x^{i*}x^i. \tag{27}$$

Summarizing the above developments, in the case $\lambda_1\geq 0$, by maximizing the locally tight lower bound, f_1, and applying (27) we obtain a more tractable problem as follows:

$$\max_{\substack{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f,\\ \bar{\mathbf{p}},x,\tilde{\mathbf{u}}_{eq},\tilde{\mathbf{p}}_{eq},t,\tilde{t}\geq 0}} \; f_1$$
$$\text{s.t. } (9b),(27),(9d),(9e). \tag{28}$$

Similarly, when $\lambda_1<0$, problem (11) can be transformed as

$$\max_{\substack{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{p},\sigma_f,\\ \bar{\mathbf{p}},x,\tilde{\mathbf{u}}_{eq},\tilde{\mathbf{p}}_{eq},t,\tilde{t}\geq 0}} \; f_2$$
$$\text{s.t. } (9b),(27),(9d),(9e). \tag{29}$$

The proposed CCCP algorithm successively maximizes either one of these lower bounds on the objective function by updating one block of variables while fixing the others, that is, by solving problem (28) (when $\lambda_1\geq 0$) or (29) (when $\lambda_1<0$) in a block coordinate descent fashion.

In each iteration of the CCCP algorithm, we perform four updating steps according to the block structure of the optimization variables. The detailed derivation of each updating step is given in Appendix B for the case when $\lambda_1\geq 0$, i.e. problem (28).[5] The proposed CCCP algorithm, to be used

[5]The developments can be easily extended to the case $\lambda_1<0$, i.e. problem (29), but we omit the details due to space limitations.

TABLE I

PROPOSED CCCP ALGORITHM IN THE INNER LOOP ($\lambda_1 \geq 0$, EQUATION NUMBER REFERS TO APPENDIX B)

1. Define the tolerance of accuracy ϵ_1 and the maximum number of iterations $N_{\max}$. Initialize the algorithm with a feasible point. Set the iteration number $i = 0$.
2. **repeat**
3. – Update $\{x, t\}$ and $\bar{\mathbf{p}}$ by solving problem (38) and (45).
4. – Update $\tilde{\mathbf{u}}_{eq}$, $\mathbf{w}$ and $\mathbf{p}$ by calculating (54) and solving problems (56) and (58).
5. – Update $\tilde{\mathbf{p}}_{eq}$, $\mathbf{u}$ and $\tilde{t}$ by calculating (60) and solving problems (62) and (63).
6. – Update $\mathbf{v}$ and σ_f by solving (64) and (66).
7. – Update the iteration number: $i = i + 1$.
8. **Until** the difference between successive values of the objective function is less than ϵ_1 or the maximum number of iterations is reached, i.e., $i \geq N_{\max}$.

in the inner loop of the PDD optimization framework when $\lambda_1 \geq 0$, is summarized in Table I.

C. Proposed PDD-Based Algorithm

In the outer loop of the PDD technique, the dual variables $\{\lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_{4,k}, \boldsymbol{\lambda}_{5,k}, \lambda_6\}^m$ can be updated by using the following expressions:

$$\begin{aligned}
\lambda_1^{m+1} &= \lambda_1^m + \frac{1}{\rho^m}\left(\|\mathbf{A}^{\frac{1}{2}}\tilde{\mathbf{u}}\| - t\|\mathbf{B}^{\frac{1}{2}}\tilde{\mathbf{p}}\|\right),\\
\boldsymbol{\lambda}_{\mathbf{2}}^{m+1} &= \boldsymbol{\lambda}_{\mathbf{2}}^m + \frac{1}{\rho^m}(\bar{\mathbf{p}} - \sigma_f \mathbf{p}),\\
\lambda_3^{m+1} &= \lambda_3^m + \frac{1}{\rho^m}\left(x - \mathbf{w}^H\mathbf{H}_2\mathbf{u}\right),\\
\boldsymbol{\lambda}_{4,k}^{m+1} &= \boldsymbol{\lambda}_{4,k}^m + \frac{1}{\rho^m}\left(\tilde{\mathbf{u}}^H\mathbf{A}_k - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})\right),\ \forall k \in \{1,\ldots,N_t\},\\
\boldsymbol{\lambda}_{5,n}^{m+1} &= \boldsymbol{\lambda}_{5,n}^m + \frac{1}{\rho^m}\left(\tilde{\mathbf{p}}^H\mathbf{B}_n - \mathbf{v}^H(\mathbf{b}_n^T\mathbf{p})\right),\ \forall n \in \{1,\ldots,M_t\},\\
\lambda_6^{m+1} &= \lambda_6^m + \frac{1}{\rho^m}\left(t - \tilde{t}\right),
\end{aligned} \tag{30}$$

where superscript m denotes the outer iteration number. To control the termination of the algorithm, we define a constraint violation indicator h as (31), as shown at the bottom of this page.

The proposed PDD-based joint beamforming and jamming algorithm is summarized in Table II. Based on the discussion in Appendix A, we can obtain that the proposed PDD-based algorithm converges to a stationary solution of problem (8).

D. Computational Complexity

In this section, we analyze the computational complexity of the proposed PDD-based algorithm. The number of complex multiplications is used to evaluate the complexity.

Let us focus on the subproblem with respect to $\bar{\mathbf{p}}$. Notwithstanding the computation of the invariant terms, the complexity for evaluating the expression of $\bar{\mathbf{p}}$ in (47) is dominated by two parts. The first part applies the matrix inverse based on Gauss-Jordan elimination, with complexity $O(M_t^3)$. The second part utilizes the bisection method to search the Lagrangian parameter $\tilde{\mu}$, with complexity $\xi \triangleq \log_2(\frac{\theta_{0,\mathbf{s}}}{\theta_{\mathbf{s}}})$, where $\theta_{0,\mathbf{s}}$ is the initial size of the search interval and $\theta_{\mathbf{s}}$ is the tolerance. In each iteration, the complexity for calculating the value of (52) is $O(1)$ since only scalars are involved. Thus the complexity for updating $\bar{\mathbf{p}}$ is $O(M_t^3 + \xi)$. Overall, the complexity for solving this subproblem is given by $O(I_1 I_2 (M_t^3 + \xi))$, where I_1 and I_2

$$\begin{aligned}
f_1 \triangleq\ & t - \frac{1}{2\rho}\big(\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2 + t^2\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2 - 2t\tilde{\mathbf{u}}_{eq}^T\boldsymbol{\Omega}\tilde{\mathbf{p}}_{eq} + \frac{\rho\lambda_1}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|}\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2 + \rho\lambda_1\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\| - \frac{2\rho\lambda_1\ t\tilde{\mathbf{p}}_{eq}^T\boldsymbol{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_t N_r}}\\
&+ \|\bar{\mathbf{p}} - \sigma_f\mathbf{p} + \rho\boldsymbol{\lambda}_2\|^2 + |x - \mathbf{w}^H\mathbf{H}_2\mathbf{u} + \rho\lambda_3|^2 + \sum_{k=1}^{N_t}\|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k - j\mathbf{C}_2\mathbf{A}_k) - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u}) + \rho\boldsymbol{\lambda}_{4,k}\|^2\\
&+ \sum_{k=1}^{M_t}\|\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k - j\mathbf{D}_2\mathbf{B}_k) - \mathbf{v}^H(\mathbf{b}_k^T\mathbf{p}) + \rho\boldsymbol{\lambda}_{5,k}\|^2 + |t - \tilde{t} + \rho\lambda_6|^2\big).
\end{aligned} \tag{25}$$

$$\begin{aligned}
f_2 \triangleq\ & t - \frac{1}{2\rho}\big(\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2 + t^2\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2 - 2t\tilde{\mathbf{u}}_{eq}^T\boldsymbol{\Omega}\tilde{\mathbf{p}}_{eq} + \frac{2\rho\lambda_1\tilde{\mathbf{u}}_{eq}^T\tilde{\boldsymbol{\Pi}}\tilde{\mathbf{u}}_{eq}}{\sqrt{N_t N_r}} - \frac{\rho\lambda_1\ t}{\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|}\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2 - \rho\lambda_1\ t\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}^i\|\\
&+ \|\bar{\mathbf{p}} - \sigma_f\mathbf{p} + \rho\boldsymbol{\lambda}_2\|^2 + |x - \mathbf{w}^H\mathbf{H}_2\mathbf{u} + \rho\lambda_3|^2 + \sum_{k=1}^{N_t}\|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k - j\mathbf{C}_2\mathbf{A}_k) - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u}) + \rho\boldsymbol{\lambda}_{4,k}\|^2\\
&+ \sum_{k=1}^{M_t}\|\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k - j\mathbf{D}_2\mathbf{B}_k) - \mathbf{v}^H(\mathbf{b}_k^T\mathbf{p}) + \rho\boldsymbol{\lambda}_{5,k}\|^2 + |t - \tilde{t} + \rho\lambda_6|^2\big).
\end{aligned} \tag{26}$$

$$\begin{aligned}
h = \max\Big\{&|\|\mathbf{A}^{\frac{1}{2}}\tilde{\mathbf{u}}\| - t\|\mathbf{B}^{\frac{1}{2}}\tilde{\mathbf{p}}\||, \|\bar{\mathbf{p}} - \sigma_f\mathbf{p}\|, |x - \mathbf{w}^H\mathbf{H}_2\mathbf{u}|, \|\tilde{\mathbf{u}}^H\mathbf{A}_k - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})\|, \|\tilde{\mathbf{p}}^H\mathbf{B}_n - \mathbf{v}^H(\mathbf{b}_n^T\mathbf{p})\|, |t - \tilde{t}|\Big\}\\
&\forall k \in \{1,\ldots,N_t\},\quad \forall n \in \{1,\ldots,M_t\}
\end{aligned} \tag{31}$$

TABLE II

PROPOSED PDD METHOD FOR JOINT BEAMFORMING AND JAMMING DESIGN

1. Initialize dual variables $\boldsymbol{\lambda}^0 = \{\lambda_1, \boldsymbol{\lambda}_2, \lambda_3, \boldsymbol{\lambda}_{4,k}, \boldsymbol{\lambda}_{5,k}, \lambda_6\}^0$, primal variables $\mathbf{z}^0 = \{x, t, \tilde{t}, \bar{\mathbf{p}}, \tilde{\mathbf{u}}_{eq}, \mathbf{w}, \mathbf{p}, \tilde{_{eq}}, \mathbf{u}, \mathbf{v}, \sigma_f\}^0$, $\rho^0 > 0$, and set ϵ_2, $0 < c < 1$, $\eta^0 > 0$.
2. **repeat**
3. Apply the proposed CCCP algorithm in the inner loop to update the primal variables, i.e. obtain $\mathbf{z}^m$.
4. Calculate the constraint violation $h(\mathbf{z}^m)$ based on (30).
5. **if** $h(\mathbf{z}^m) \leq \eta^m$
6. Update dual variables based on (30).
7. $\rho^{m+1} = \rho^m$
8. **else**
9. $\boldsymbol{\lambda}^{m+1} = \boldsymbol{\lambda}^m$
10. $\rho^{m+1} = c\rho^m$
11. **end**
12. $\eta^{m+1} = 0.9h(\mathbf{z}^m)$
13. $m = m + 1$
14. **until** $h < \epsilon_2$.

denote the number of iterations for the outer and inner loops, respectively.

Following a similar approach, we can obtain the computational complexity for the other subproblems. For simplicity, we assume that the ratio of the initial maximum interval size to the tolerance of the bisection search, or equivalently its logarithm ξ, is the same for all subproblems. In particular, note that the proposed one-iteration BCD-type algorithm shown in Table IV has a complexity of $O(J_{in}^2)$, where J_{in} is the number of input variables. Finally, the computational complexity of the whole PDD algorithm can be expressed as $O(I_1 I_2 \times (\xi + N_t^3 N_r^3 + M_t^3 N_r^3))$.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed PDD-based joint beamforming and jamming algorithm by means of computer simulations. As in [38], we employ a standard 60GHz mmWave channel model with uniform linear antenna array configuration. The channel matrix $\mathbf{H}_1$ can be expressed as

$$\mathbf{H}_1 = \sqrt{\frac{N_t N_r}{L}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_r(\phi_l^r)\mathbf{a}_t(\phi_l^t) \tag{32}$$

where L is the number of distinguishable paths, α_l is the complex gain of the l-th path, $\mathbf{a}_r(\phi_l^r)$ and $\mathbf{a}_t(\phi_l^t)$ are the receive and transmit array response vectors at the azimuth angle of $\phi_l^r \in [0, 2\pi)$ and $\phi_l^t \in [0, 2\pi)$, respectively. The generic expression for the response vector is given by

$$\mathbf{a}(\theta) = \frac{1}{N}\left[1, e^{jkd\pi\sin(\theta)}, \ldots, e^{jkd(N-1)\pi\sin(\theta)}\right]^T \tag{33}$$

where $k = 2\pi/\lambda$, λ is the wavelength, and d is the antenna spacing. Channel matrices $\mathbf{H}_2$ and $\mathbf{H}_3$ can be generated in a similar manner. The residual self-intefernce channel $\mathbf{H}_4$ is modeled as the product of a degradation and near-field attenuation factor, according to the simplified model in [39].

The number of transmit and receive antennas of the suspicious transmission link are set to $N_t = N_r = 4$. We assume that the SC uses the same number of transmit and receive antennas, i.e., $M_t = M_r = M$. For simplicity, we set all the noise power terms to $\sigma^2 = 1$, while the input SNR of the suspicious link is defined as $\frac{\sigma_h^2}{\sigma^2}$. Referring to the channel model (32), the number of independent paths is set to $L = 20$ while the complex gains α_l are generated according to a $\mathcal{CN}(0, 1)$ distribution. The normalized power level of the residual self-interference channel is set as $\varrho_{41} \triangleq \|\mathbf{H}_4\|^2/\|\mathbf{H}_1\|^2 = -30$dB. All results are obtained by averaging over 1000 independent Monte Carlo runs. The tolerance parameters for the proposed algorithm are chosen as $\epsilon_1 = \epsilon_2 = 10^{-4}$ while $N_{max} = 50$. The parameter c is set to be $c = 0.6$. We consider the following algorithms for comparison:

1) The proposed PDD-based joint beamforming and jamming algorithm in Table II.
2) Passive eavesdropping: the SC does not transmit jamming signals and the analog beamforming vectors are designed based on the channel matching method [13].
3) Constant-power jamming: the jamming signal's power level is constant and the channel matching based analog beamforming vectors are applied.
4) On-off jamming: the SC does not jam if the monitoring channel condition is good. Otherwise, it sends jamming signals with minimum power for successful overhearing. The channel matching based analog beamforming vectors are applied.

In a first series of experiments, we study the convergence of the proposed PDD-based algorithm versus the number of outer iterations m. Here we set $M = 4$, $P_t = 20$dB and SNR $= 20$dB while the monitoring channel condition is given by $\varrho_{21} \triangleq \|\mathbf{H}_2\|^2/\|\mathbf{H}_1\|^2 = -20$dB. In Fig. 2 (a), we plot the value of the constraint violation indicator in (31) versus iteration number for the proposed algorithm. We find that the convergence of the penalty terms is very fast with h decreasing to a value of 10^{-4} after 40 iterations, which supports our claim that the proposed PDD-based algorithm can effectively tackle the equality constraints. Fig. 2 (b) shows the monitoring rate performance at the SC versus the number of iterations for the proposed algorithm. It can be observed from this figure that the latter converges within 35 iterations.

Fig. 3 and Fig. 4 show the monitoring rate performance versus SNR of the different algorithms under study for two different monitoring channel characterized by $\varrho_{21} = -20$dB and 0dB, respectively; for these experiments we set $M = 4$

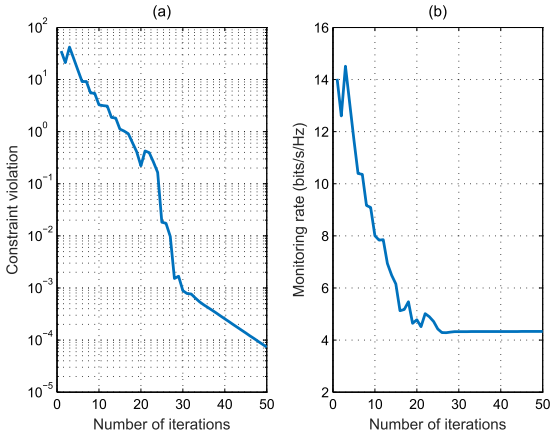

Fig. 2. Convergence performance of the proposed PDD-based algorithm: (a) constraint violation indicator versus the number of iterations; (b) monitoring rate versus the number of iterations.

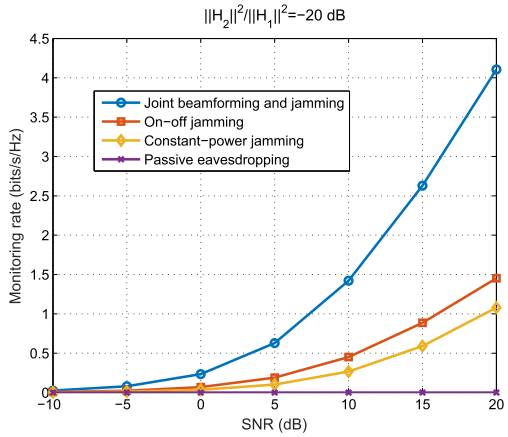

Fig. 3. Monitoring rate versus SNR ($\varrho_{21} = -20$dB).

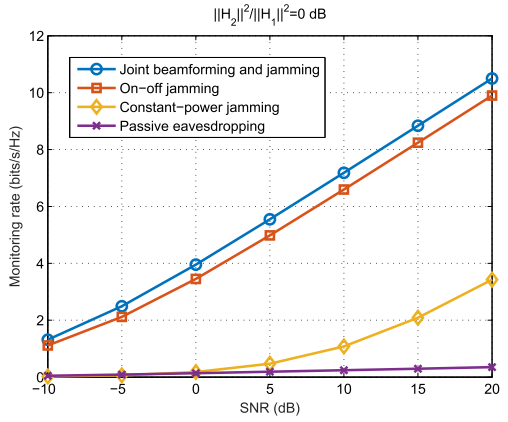

Fig. 4. Monitoring rate versus SNR ($\varrho_{21} = 0$dB).

and $P_t = 20$dB. For both cases of channel condition, the proposed algorithm achieves the best performance, followed by the on-off jamming, constant-power jamming, and passive eavesdropping algorithms, where the latter fails under worst monitoring channel conditions. Moreover, a better monitoring channel provides higher monitoring rate performance while the relative gain resulting from the proposed algorithm becomes larger in worse monitoring channel scenarios.

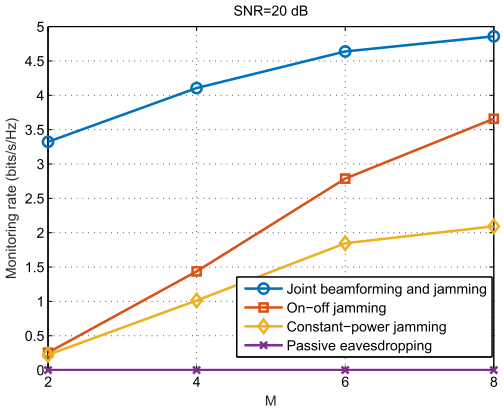

Fig. 5. Monitoring rate versus number of antennas M.

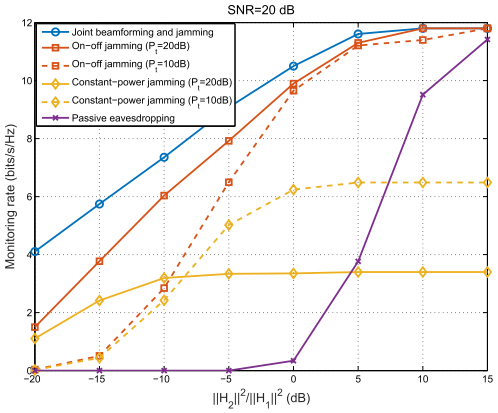

Fig. 6. Monitoring rate versus ϱ_{21}.

Fig. 5 shows the monitoring rate performance of the four analyzed algorithms versus the number of SC transmit/receive antennas M. In this simulation, we set SNR= 20dB, $P_t = 10$dB and the monitoring channel condition is given by $\varrho_{21} = -20$dB. From the figure, the monitoring rate performance improves as the number of antennas increases. Moreover, the proposed PDD-based algorithm achieves the best performance.

Next we compare monitoring rate performance of the four algorithms for different conditions of monitoring channel, as measured by ϱ_{21}, for the case $M = 4$ and SNR $= 20$dB. In Fig. 6, the results show that the proposed design algorithm achieves the best performance. The relative gain becomes larger with a decrease in the quality of the monitoring channel. The on-off jamming algorithm with a larger SC transmit power budget provides better performance while the passive eavesdropping algorithm fails in bad monitoring channels but works under good conditions. Moreover, when the monitoring channel gain is much larger than that of the suspicious channel (e.g., $\varrho_{21} = 15$dB), except the constant power jamming, all the algorithms generate similar performance because in this case, the SC does not need to send jamming signals and the performance becomes less sensitive to the design of the analog beamforming vectors.

Fig. 7 illustrates the jamming power versus ϱ_{21} for the proposed and on-off jamming algorithms. From the results,

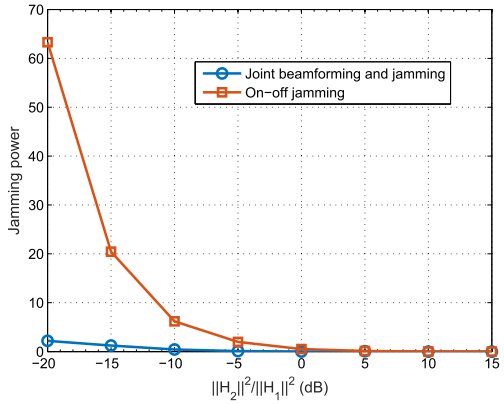

Fig. 7. Jamming power versus ϱ_{21}.

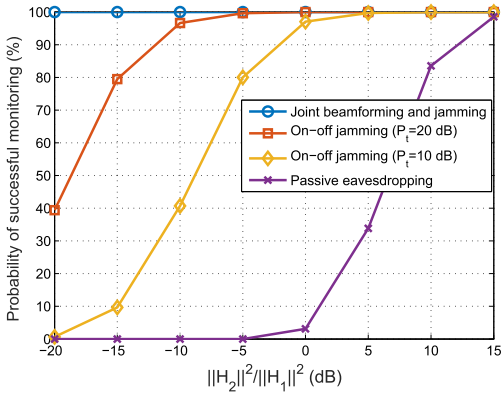

Fig. 8. Probability of successful monitoring versus ϱ_{21}.

it is seen that these algorithms do not attempt to jam the suspicious receiver when the monitoring channel power is large. It can also be seen that when the use of jamming is necessary (i.e. weak monitoring channel), the proposed PDD-based algorithm leads to a significant reduction in the SC transmit power. These results demonstrate the effectiveness of the proposed joint design to handle different scenarios of monitoring channel.

Fig. 8 compares the probability of successful monitoring achieved at the SC by the proposed algorithm, the passive eavesdropping algorithm and the on-off jamming algorithm (note that the on-off jamming and constant-power jamming algorithms share the same probability of successful monitoring) versus monitoring channel condition ϱ_{21}. In this experiment, a design algorithm is considered unsuccessful for a given channel realization as long as any one of the underlying constraints is not satisfy. From this figure, we observe that the rate of success of the various algorithms under study increases with ϱ_{21}. However, the proposed algorithm significantly outperforms the other ones by a significant margin, which shows the effectiveness of the proposed PDD approach for joint optimal design.

V. Conclusion

In this paper, we have investigated the joint beamforming and jamming design for a mmWave information surveillance system. In order to tackle this challenging problem, where the design variables are highly coupled through the objectives and constraint functions, a novel PDD algorithm has been proposed to jointly optimize the analog transmit and receive beamforming vectors of the suspicious transmission link, the analog jamming and monitoring beamforming vectors at the SC and the jamming signal power level, under transmit power, successful monitoring and self-interference power constraints at the SC, along with unit modulus constraint on the analog beamforming vectors. We have studied the complexity of the proposed PDD-based algorithm and claimed that it converges to a stationary solution of the original problem. Our simulation results have demonstrated that the proposed algorithm achieves better performance than the competing beamforming and jamming algorithms used as benchmarks. Finally, the proposed PDD algorithmic framework can be also applied to solve other highly coupled optimization problems appearing in the design of wireless communication systems as well as other compelling applications in science and engineering.

Appendix A
Proposed PDD Optimization Framework

In this appendix, we present our proposed PDD-based design algorithm within the general framework of [31] and [32] and discuss its convergence behavior. The PDD method is an algorithmic framework that can be applied to the minimization of a nonconvex nonsmooth function subject to nonconvex coupling constraints, by blending the benefits of the penalty and augmented Lagrangian methods [33]. Consider the following problem

$$\begin{aligned} &\min_{\mathbf{x}\in\mathcal{X}} \ f(\mathbf{x}) \\ &\text{s.t. } \mathbf{h}(\mathbf{x}) = \mathbf{0}, \\ &\quad \mathbf{g}_i(\mathbf{x}_i) \leq \mathbf{0}, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{34}$$

where $f(\mathbf{x})$ is a scalar continuously differentiable function, $\mathbf{g}_i(\mathbf{x}_i) \in \mathbb{R}^{q_i \times 1}$ is a vector of q_i continuously differentiable but *possibly nonconvex* functions while $\mathbf{h}(\mathbf{x}) \in \mathbb{R}^{p\times 1}$ is a vector of p continuously differentiable functions. The feasible set $\mathcal{X}$ is the Cartesian product of n simple closed convex sets: $\mathcal{X} \triangleq \prod_{i=1}^{n} \mathcal{X}_i$ with $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$ and $\sum_{i=1}^{n} n_i = N$. Accordingly the optimization variable $\mathbf{x} \in \mathbb{R}^N$ can be decomposed as $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ with $\mathbf{x}_i \in \mathcal{X}_i, i = 1, 2, \ldots, n$.

If the coupling constraints, $\mathbf{h}(\mathbf{x}) = 0$, is absent, the classical CCCP algorithms [35]–[37] can be applied for solving problem (34). This observation motivates us to dualize the difficult coupling constraints with appropriate penalty, and use coordinate-decomposition to perform fast computation, hence the new optimization framework is referred to as the penalty dual decomposition method. In particular, the PDD method applied to problem (34) takes the form of a double-loop algorithm, where the inner loop serves to optimize the nonconvex augmented Lagrangian problem via a block coordinate descent method (i.e., to some predefined accuracy using a few inner iterations), while the outer loop serves to update the dual variable and a penalty parameter.

The detailed steps of the PDD method are presented in Table III, where the notation 'optimize $(P(\rho^m, \boldsymbol{\lambda}^m)$,

TABLE III
PDD METHOD FOR PROBLEM (34)

1 initialize $\mathbf{x}^0, \rho^0 > 0, \boldsymbol{\lambda}^0$, and set $0 < c < 1$, $m = 1$
2 **repeat**
3 $\quad \mathbf{x}^m = \text{optimize}(P(\rho^m, \boldsymbol{\lambda}^m), \mathbf{x}^{m-1}, \epsilon^m)$
4 $\quad$ **if** $\|\mathbf{h}(\mathbf{x}^m)\|_\infty \leq \eta^m$
5 $\quad\quad \boldsymbol{\lambda}^{m+1} = \boldsymbol{\lambda}^m + \frac{1}{\rho^m}\mathbf{h}(\mathbf{x}^m)$
6 $\quad\quad \rho^{m+1} = \rho^m$
7 $\quad$ **else**
8 $\quad\quad \boldsymbol{\lambda}^{m+1} = \boldsymbol{\lambda}^m$
9 $\quad\quad \rho^{m+1} = c\rho^m$
10 $\quad$ **end**
11 $\quad m = m + 1$
12 **until** some termination criterion is met

$\mathbf{x}^{m-1}, \epsilon^m)$' represents an optimization oracle that has the following meaning: starting from $\mathbf{x}^{m-1}$, it invokes an iterative optimization algorithm to solve to some accuracy ϵ^m the following problem:

$$P(\rho^m, \boldsymbol{\lambda}^m): \quad \min_{\mathbf{x}\in\mathcal{X}} \ \{\mathcal{L}_m(\mathbf{x}) \triangleq f(\mathbf{x}) + \boldsymbol{\lambda}^{mT}\mathbf{h}(\mathbf{x}) + \frac{1}{2\rho^m}\|\mathbf{h}(\mathbf{x})\|^2\}, \quad (35)$$

where $\mathcal{L}_m(\mathbf{x})$ is the augmented Lagrange function with dual variable $\boldsymbol{\lambda}^m$ and penalty parameter ρ^m.

For simplicity, let us define $\mathbf{g}(\mathbf{x}) \triangleq (\mathbf{g}_i(\mathbf{x}_i))_i$. We have the following theorem regarding the convergence of the PDD method.

Theorem 1: Let $\{\mathbf{x}^m, \boldsymbol{\nu}^m\}$ be the sequence generated by the algorithm proposed in Table III, where $\boldsymbol{\nu}^m = (\boldsymbol{\nu}_i^m)_i$ denotes the Lagrange multipliers associated with the constraints $\mathbf{g}_i(\mathbf{x}_i) \leq 0$, $\forall i$. The termination criterion for the oracle involved in this algorithm is

$$\left\|\nabla_{\mathbf{x}}\mathcal{L}_m(\mathbf{x}^m) + \nabla\mathbf{g}(\mathbf{x}^m)^T\boldsymbol{\nu}^m\right\|_\infty \leq \epsilon^m, \quad \forall m \quad (36)$$

with $\epsilon^m, \eta^m \to 0$ as $m \to \infty$. Suppose that $\mathbf{x}^\star$ is a limit point of the sequence $\{\mathbf{x}^m\}$ and *Robinson's condition*[6] [40, Ch. 3] holds for problem (34) at $\mathbf{x}^\star$. Then $\mathbf{x}^\star$ is a stationary point of (34), i.e., it satisfies the KKT condition.

A general proof of this theorem can be found in [31] and [32]. It is seen that the key to the PDD method is the implementation of the optimization oracle in Step 3. Typically, to exploit the separable structure of the constraints of problem $P(\rho^m, \boldsymbol{\lambda}^m)$, the oracle can be taken as the CCCP algorithm [35]–[37], where each time one block of variables is chosen to be optimized while fixing the others, by minimizing a locally tight upper bound of the objective function. Further, a simple way to set η^m is to make it explicitly related to the constraint violation of the last iteration or the current minimum constraint violation. For example, we set $\eta^m = 0.9\|\mathbf{h}(\mathbf{x}^{m-1})\|_\infty$ in the simulation. Moreover, since the penalty term $\|\mathbf{h}(\mathbf{x})\|_\infty$ vanishes eventually, a practical choice of the termination criterion for the PDD method can be $\|\mathbf{h}(\mathbf{x}^m)\|_\infty \leq \epsilon_o$, where ϵ_o denotes a prescribed small constant.

[6]Robinson's condition is a type of constraint qualification condition used for KKT analysis and the assumption is a standard one, made in many previous works on constrained optimization, e.g., [40]–[42].

APPENDIX B
DERIVATION OF UPDATING STEPS IN THE CCCP ALGORITHM

In **Step 1**, we optimize $\{x, t\}$ and $\bar{\mathbf{p}}$ in parallel by fixing the other variables. Note that in this case problem (28) can be decomposed into two independent subproblems. The first subproblem with respect to $\{x, t\}$ is expressed as (37), as shown at the top of the next page. Introducing a new variable $\boldsymbol{\beta} \triangleq [t, \Re\{x\}, \Im\{x\}]^T \in \mathbb{R}^{3\times 1}$ and basis vectors $\mathbf{e}_1 \triangleq [1,0,0]^T$, $\mathbf{e}_2 \triangleq [0,1,0]^T$, and $\mathbf{e}_3 \triangleq [0,0,1]^T$, problem (37) can be equivalently expressed as a convex quadratic optimization problem shown in (38), as shown at the top of the next page. Note that (38) can be solved in closed form based on the Lagrange multipliers method. Attaching a Lagrange multiplier μ to constraint (38b), we obtain after some manipulations the Lagrange function shown in (39), as shown at the top of the next page. By applying the first order optimality condition with respect to $\boldsymbol{\beta}$, we obtain the optimal value of $\boldsymbol{\beta}$ as

$$\boldsymbol{\beta}(\mu) = \frac{1}{\rho}\mathbf{J}^{-1}\mathbf{r} \quad (40)$$

where

$$\mathbf{J} = \begin{bmatrix} \frac{1}{\rho}\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2 + 2\mu\bar{\sigma} & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ 0 & 0 & \frac{1}{\rho} \end{bmatrix} \triangleq \text{diag}\left\{\theta(\mu), \frac{1}{\rho}, \frac{1}{\rho}\right\} \quad (41)$$

and

$$\mathbf{r} = \begin{bmatrix} \rho + \tilde{\mathbf{u}}_{eq}^T\boldsymbol{\Omega}\tilde{\mathbf{p}}_{eq} + \frac{\rho\lambda_1\tilde{\mathbf{p}}_{eq}^T\boldsymbol{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_t N_r}} \\ \Re\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\} - \Re\{\rho\lambda_3\} + 2\mu\rho\Re\{x^i\} \\ \Im\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\} - \Im\{\rho\lambda_3\} + 2\mu\rho\Im\{x^i\} \end{bmatrix} \triangleq \begin{bmatrix} \pi_1 \\ \pi_2(\mu) \\ \pi_3(\mu) \end{bmatrix}. \quad (42)$$

Here $\mu \geq 0$ is chosen such that the complementarity slackness condition of constraint (38b) is satisfied.

When $\mathbf{e}_1^T\boldsymbol{\beta}(0)\boldsymbol{\beta}(0)^T\mathbf{e}_1\bar{\sigma} - 2\Re\{x^i\}\mathbf{e}_2^T\boldsymbol{\beta}(0) - 2\Im\{x^i\}\mathbf{e}_3^T\boldsymbol{\beta}(0) + |x^i|^2 \leq 0$ is satisfied, then the optimal $\boldsymbol{\beta} = \boldsymbol{\beta}(0)$. Otherwise we must have

$$\mathbf{e}_1^T\boldsymbol{\beta}(\mu)\boldsymbol{\beta}(\mu)^T\mathbf{e}_1\bar{\sigma} - 2\Re\{x^i\}\mathbf{e}_2^T\boldsymbol{\beta}(\mu) - 2\Im\{x^i\}\mathbf{e}_3^T\boldsymbol{\beta}(\mu) + |x^i|^2 = 0. \quad (43)$$

Moreover, (43) can be rewritten as

$$\frac{\pi_1^2}{\rho^2\theta(\mu)^2}\bar{\sigma} - 2\Re\{x^i\}\pi_2(\mu) - 2\Im\{x^i\}\pi_3(\mu) + |x^i|^2 = 0 \quad (44)$$

which is a cubic equation with respect to μ. Note the optimal μ must be positive, hence (44) can be solved using a one-dimensional search (e.g., bisection method [34]). Finally, by substituting the optimal μ we obtain the solution for $\boldsymbol{\beta}$, i.e., the optimal $\{x, t\}$.

The second subproblem with respect to $\bar{\mathbf{p}}$ is given by

$$\min_{\bar{\mathbf{p}}} \ \|\bar{\mathbf{p}} - \sigma_f\mathbf{p} + \rho\boldsymbol{\lambda}_2\|^2 \quad (45a)$$

$$\text{s.t. } \|\mathbf{H}_4\bar{\mathbf{p}}\|^2 \leq \delta. \quad (45b)$$

$$\min_{x,t} \ -t+\frac{1}{2\rho}\left(t^2\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2-2t\tilde{\mathbf{u}}_{eq}^T\mathbf{\Omega}\tilde{\mathbf{p}}_{eq}-\frac{2\rho\lambda_1\ t\tilde{\mathbf{p}}_{eq}^T\mathbf{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_tN_r}}+|x-\mathbf{w}^H\mathbf{H}_2\mathbf{u}+\rho\lambda_3|^2\right) \tag{37a}$$

$$\text{s.t. } \bar{\sigma}t^2\le x^{i*}x+x^*x^i-x^{i*}x^i. \tag{37b}$$

$$\min_{\boldsymbol{\beta}} \ -\mathbf{e}_1^T\boldsymbol{\beta}+\frac{1}{2\rho}\Bigg(\mathbf{e}_1^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_1\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2-2\mathbf{e}_1^T\boldsymbol{\beta}\tilde{\mathbf{u}}_{eq}^T\mathbf{\Omega}\tilde{\mathbf{p}}_{eq}-2\mathbf{e}_1^T\boldsymbol{\beta}\frac{\rho\lambda_1\tilde{\mathbf{p}}_{eq}^T\mathbf{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_tN_r}}+|\mathbf{e}_2^T\boldsymbol{\beta}-\Re\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\}+\Re\{\rho\lambda_3\}|^2$$
$$+|\mathbf{e}_3^T\boldsymbol{\beta}-\Im\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\}+\Im\{\rho\lambda_3\}|^2\Bigg) \tag{38a}$$

$$\text{s.t. } \mathbf{e}_1^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_1\bar{\sigma}-2\Re\{x^i\}\mathbf{e}_2^T\boldsymbol{\beta}-2\Im\{x^i\}\mathbf{e}_3^T\boldsymbol{\beta}+|x^i|^2\le 0. \tag{38b}$$

$$L(\boldsymbol{\beta},\mu)\triangleq\frac{1}{2\rho}\Bigg(\mathbf{e}_1^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_1\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2+\mathbf{e}_2^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_2+\mathbf{e}_3^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_3-2\mathbf{e}_1^T\boldsymbol{\beta}\tilde{\mathbf{u}}_{eq}^T\mathbf{\Omega}\tilde{\mathbf{p}}_{eq}-2\mathbf{e}_1^T\boldsymbol{\beta}\frac{\rho\lambda_1\tilde{\mathbf{p}}_{eq}^T\mathbf{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_tN_r}}$$
$$+2\mathbf{e}_2^T\boldsymbol{\beta}\left(\Re\{\rho\lambda_3\}-\Re\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\}\right)+2\mathbf{e}_3^T\boldsymbol{\beta}\left(\Im\{\rho\lambda_3\}-\Im\{\mathbf{w}^H\mathbf{H}_2\mathbf{u}\}\right)\Bigg)-\mathbf{e}_1^T\boldsymbol{\beta}$$
$$+\mu\big(\mathbf{e}_1^T\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{e}_1\bar{\sigma}-2\Re\{x^i\}\mathbf{e}_2^T\boldsymbol{\beta}-2\Im\{x^i\}\mathbf{e}_3^T\boldsymbol{\beta}+|x^i|^2\big). \tag{39}$$

This is also a convex quadratic optimization problem, and by following the same approach, we obtain its Lagrange function

$$L(\bar{\mathbf{p}},\tilde{\mu})\triangleq\mathrm{Tr}((\bar{\mathbf{p}}-\sigma_f\mathbf{p}+\rho\boldsymbol{\lambda}_2)(\bar{\mathbf{p}}-\sigma_f\mathbf{p}+\rho\boldsymbol{\lambda}_2)^H)+\tilde{\mu}(\mathrm{Tr}(\mathbf{H}_4\bar{\mathbf{p}}\bar{\mathbf{p}}^H\mathbf{H}_4^H)-\delta) \tag{46}$$

where $\tilde{\mu}$ denotes the Lagrange multiplier of constraint (45b). Applying the first order optimality condition of $L(\bar{\mathbf{p}},\tilde{\mu})$ with respect to $\bar{\mathbf{p}}$ yields

$$\bar{\mathbf{p}}(\tilde{\mu})=(\mathbf{I}_{M_t}+\tilde{\mu}\mathbf{H}_4^H\mathbf{H}_4)^{-1}(\sigma_f\mathbf{p}-\rho\boldsymbol{\lambda}_2). \tag{47}$$

If the solution of $\bar{\mathbf{p}}$ with $\tilde{\mu}=0$ satisfies constraint (45b), the optimal $\tilde{\mu}$ is zero. Otherwise we can obtain the solution of $\tilde{\mu}$ through the slackness condition

$$\|\mathbf{H}_4\bar{\mathbf{p}}\|^2=\delta \tag{48}$$

which is equivalent to

$$\mathrm{Tr}((\mathbf{I}_{M_t}+\tilde{\mu}\bar{\mathbf{H}}_4)^{-1}\bar{\mathbf{H}}_4(\mathbf{I}_{M_t}+\tilde{\mu}\bar{\mathbf{H}}_4)^{-1}\tilde{\mathbf{X}})=\delta \tag{49}$$

where $\tilde{\mathbf{X}}\triangleq(\sigma_f\mathbf{p}-\rho\boldsymbol{\lambda}_2)(\sigma_f\mathbf{p}-\rho\boldsymbol{\lambda}_2)^H$ and $\bar{\mathbf{H}}_4\triangleq\mathbf{H}_4^H\mathbf{H}_4$. Note that

$$\mathbf{I}_{M_t}+\tilde{\mu}\bar{\mathbf{H}}_4=\bar{\mathbf{H}}_4^{\frac{1}{2}}(\bar{\mathbf{H}}_4^{-1}+\tilde{\mu}\mathbf{I}_{M_t})\bar{\mathbf{H}}_4^{\frac{1}{2}}. \tag{50}$$

Thus (49) can be rewritten as

$$\mathrm{Tr}((\bar{\mathbf{H}}_4^{-1}+\tilde{\mu}\mathbf{I}_{M_t})^{-2}\hat{\mathbf{X}})=\delta \tag{51}$$

where $\hat{\mathbf{X}}\triangleq\bar{\mathbf{H}}_4^{-\frac{1}{2}}\tilde{\mathbf{X}}\bar{\mathbf{H}}_4^{-\frac{1}{2}}$. Finally, (49) can be equivalently expressed as

$$\sum_{i=1}^{M_t}\frac{b_i}{(a_i+\tilde{\mu})^2}=\delta \tag{52}$$

where $b_i=[\bar{\mathbf{J}}^H\hat{\mathbf{X}}\bar{\mathbf{J}}]_{i,i}$, $\bar{\mathbf{J}}$ denotes a unitary matrix consisting of the eigenvectors of $\bar{\mathbf{H}}_4^{-1}$, and $\{a_k\}$ denotes the corresponding eigenvalues, this is: $\bar{\mathbf{J}}\mathrm{diag}\{a_1,a_2,\ldots,a_{M_t}\}\bar{\mathbf{J}}^H$ is the eigenvalue decomposition of $\bar{\mathbf{H}}_4^{-1}$. Since $\tilde{\mu}\ge 0$, (52) can be easily solved by using the bisection method. Finally the optimal $\bar{\mathbf{p}}$ can be obtained by substituting the optimal $\tilde{\mu}$.

In **Step 2**, we optimize $\tilde{\mathbf{u}}_{eq}$, $\mathbf{w}$, $\mathbf{p}$ in parallel while other variables are fixed. In this case, problem (28) can be decomposed into three independent subproblems.

The subproblem for the real vector $\tilde{\mathbf{u}}_{eq}$ is an unconstrained quadratic optimization problem, which can be expressed as

$$\min_{\tilde{\mathbf{u}}_{eq}} \ \|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2-2t\tilde{\mathbf{u}}_{eq}^T\mathbf{\Omega}\tilde{\mathbf{p}}_{eq}+\frac{\rho\lambda_1}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|}\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}\|^2+\sum_{k=1}^{N_t}\|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k-j\mathbf{C}_2\mathbf{A}_k)-\mathbf{v}^H(\mathbf{a}_k^T\mathbf{u})+\rho\boldsymbol{\lambda}_{4,k}\|^2. \tag{53}$$

By examining the first order optimality condition, we can obtain the unique optimal solution

$$\tilde{\mathbf{u}}_{eq}=\left(\left(1+\frac{\rho\lambda_1}{\|\mathbf{A}_{eq}^{\frac{1}{2}}\tilde{\mathbf{u}}_{eq}^i\|}\right)\mathbf{A}_{eq}+\mathbf{I}_{2N_tN_r}\right)^{-1}(t\mathbf{\Omega}\tilde{\mathbf{p}}_{eq}-\mathbf{s}) \tag{54}$$

where

$$\mathbf{s}=\big[\left(\Re\{\rho\boldsymbol{\lambda}_{4,1}\}-\Re\{\mathbf{a}_1^T\mathbf{u}\mathbf{v}^H\}\right),\ldots,\left(\Re\{\rho\boldsymbol{\lambda}_{4,N_t}\}-\Re\{\mathbf{a}_{N_t}^T\mathbf{u}\mathbf{v}^H\}\right),\left(\Im\{\rho\boldsymbol{\lambda}_{4,1}\}-\Im\{\mathbf{a}_1^T\mathbf{u}\mathbf{v}^H\}\right),\ldots,\left(\Im\{\rho\boldsymbol{\lambda}_{4,N_t}\}-\Im\{\mathbf{a}_{N_t}^T\mathbf{u}\mathbf{v}^H\}\right)\big]^T. \tag{55}$$

The subproblems with respect to $\mathbf{w}$ and $\mathbf{p}$ can be handled similarly since both of them can be transformed into quadratic optimization problems with unit modulus constraints. The subproblem for $\mathbf{w}$ is

$$\min_{\mathbf{w}} \ |x-\mathbf{w}^H\mathbf{H}_2\mathbf{u}+\rho\lambda_3|^2 \tag{56a}$$

$$\text{s.t. } |\mathbf{w}(i)|=1,\quad\forall i. \tag{56b}$$

By appropriate rearrangement, this can be equivalently formulated as follows

$$\min_{\mathbf{w}} \ \mathbf{w}^H\mathbf{\Gamma}_1\mathbf{w} - 2\Re\{\mathbf{w}^H\boldsymbol{\omega}_1\} \tag{57a}$$
$$\text{s.t. } |\mathbf{w}(i)| = 1, \quad \forall i, \tag{57b}$$

where $\mathbf{\Gamma}_1 \triangleq \mathbf{H}_2\mathbf{u}\mathbf{u}^H\mathbf{H}_2^H$ and $\boldsymbol{\omega}_1 \triangleq \mathbf{H}_2\mathbf{u}(x+\rho\lambda_3)^*$. Noting that the unit modulus constraints are separable, we can update $\mathbf{w}$ by using a single iteration BCD-type algorithm [43] as further explained in Appendix C. Similarly, the subproblem for $\mathbf{p}$ is given by

$$\begin{aligned}\min_{\mathbf{p}} \ & \|\bar{\mathbf{p}} - \sigma_f\mathbf{p} + \rho\boldsymbol{\lambda}_2\|^2 \\ & + \sum_{k=1}^{M_t} \|\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k - j\mathbf{D}_2\mathbf{B}_k) - \mathbf{v}^H(\mathbf{b}_k^T\mathbf{p}) + \rho\boldsymbol{\lambda}_{5,k}\|^2 \\ \text{s.t. } & |\mathbf{p}(i)| = 1, \quad \forall i. \end{aligned} \tag{58}$$

By following the same approach, we can rewrite (58) as a quadratic optimization problem with unit modulus constraints as follows,

$$\begin{aligned}\min_{\mathbf{p}} \ & \mathbf{p}^H\mathbf{\Gamma}_2\mathbf{p} - 2\Re\{\mathbf{p}^H\boldsymbol{\omega}_2\} \\ \text{s.t. } & |\mathbf{p}(i)| = 1, \quad \forall i, \end{aligned} \tag{59}$$

where $\mathbf{\Gamma}_2 \triangleq (\sigma_f^2 + \mathbf{v}^H\mathbf{v})\mathbf{I}_{M_t}$ and $\boldsymbol{\omega}_2 \triangleq (\bar{\mathbf{p}} + \rho\boldsymbol{\lambda}_2)\sigma_f^2 + \sum_{k=1}^{M_t}(\mathbf{b}_k(\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k - j\mathbf{D}_2\mathbf{B}_k) + \rho\boldsymbol{\lambda}_{5,k})\mathbf{v})$. Problem (59) can be solved by using the same approach as presented in Appendix C.

In **Step 3**, we optimize $\tilde{\mathbf{p}}_{eq}$, $\mathbf{u}$ and $\tilde{t}$ by fixing the remaining variables. In this case, problem (28) can be decomposed into three independent subproblems.

The subproblem for $\tilde{\mathbf{p}}_{eq}$ is an unconstrained quadratic optimization problem and can be addressed like that for $\tilde{\mathbf{u}}_{eq}$. Examining its first order optimality condition yields a closed form as follows,

$$\tilde{\mathbf{p}}_{eq} = \left(t^2\mathbf{B}_{eq} - \frac{2\rho\lambda_1\ t\mathbf{\Pi}}{\sqrt{M_tN_r}} + \mathbf{I}_{2M_tN_r}\right)^{-1}(t\mathbf{\Omega}^T\tilde{\mathbf{u}}_{eq} - \tilde{\mathbf{s}}) \tag{60}$$

where

$$\begin{aligned}\tilde{\mathbf{s}} = \big[\ & \big(\Re\{\rho\boldsymbol{\lambda}_{5,1}\} - \Re\{\mathbf{b}_1^T\mathbf{p}\mathbf{v}^H\}\big), \ldots, \\ & \big(\Re\{\rho\boldsymbol{\lambda}_{5,M_t}\} - \Re\{\mathbf{b}_{M_t}^T\mathbf{p}\mathbf{v}^H\}\big), \big(\Im\{\rho\boldsymbol{\lambda}_{5,1}\} - \Im\{\mathbf{b}_1^T\mathbf{p}\mathbf{v}^H\}\big), \\ & \ldots, \big(\Im\{\rho\boldsymbol{\lambda}_{5,M_t}\} - \Im\{\mathbf{b}_{M_t}^T\mathbf{p}\mathbf{v}^H\}\big)\ \big]^T. \end{aligned} \tag{61}$$

The subproblem for $\mathbf{u}$ is given by

$$\min_{\mathbf{u}} \ |x - \mathbf{w}^H\mathbf{H}_2\mathbf{u} + \rho\lambda_3|^2 \tag{62a}$$
$$+ \sum_{k=1}^{N_t} \|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k - j\mathbf{C}_2\mathbf{A}_k) - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u}) + \rho\boldsymbol{\lambda}_{4,k}\|^2 \tag{62b}$$
$$\text{s.t. } |\mathbf{u}(i)| = 1, \quad \forall i. \tag{62c}$$

Thus $\mathbf{u}$ can be optimized by following the approach presented in Appendix C.

The subproblem for $\tilde{t}$ is expressed as

$$\begin{aligned}\min_{\tilde{t}} \ & |t - \tilde{t} + \rho\lambda_6|^2 \\ \text{s.t. } & \tilde{t} \geq 0. \end{aligned} \tag{63}$$

It is readily seen that the optimal value will be $\tilde{t} = \max\{0, t + \rho\lambda_6\}$.

In **Step 4**, we optimize $\mathbf{v}$ and σ_f by fixing the other variables. To this end, we decompose problem (28) into two independent subproblems.

The corresponding subproblem for $\mathbf{v}$ is a quadratic optimization problem with unit modulus constraints, as given by

$$\min_{\mathbf{v}} \ \sum_{k=1}^{N_t} \|\tilde{\mathbf{u}}_{eq}^T(\mathbf{C}_1\mathbf{A}_k - j\mathbf{C}_2\mathbf{A}_k) - \mathbf{v}^H(\mathbf{a}_k^T\mathbf{u}) + \rho\boldsymbol{\lambda}_{4,k}\|^2$$
$$+ \sum_{k=1}^{M_t} \|\tilde{\mathbf{p}}_{eq}^T(\mathbf{D}_1\mathbf{B}_k - j\mathbf{D}_2\mathbf{B}_k) - \mathbf{v}^H(\mathbf{b}_k^T\mathbf{p}) + \rho\boldsymbol{\lambda}_{5,k}\|^2 \tag{64a}$$
$$\text{s.t. } |\mathbf{v}(i)| = 1, \forall i. \tag{64b}$$

We adopt the same method as in Appendix C to solve this problem.

The subproblem for σ_f can be expressed as

$$\begin{aligned}\min_{\sigma_f} \ & t^2\|\mathbf{B}_{eq}^{\frac{1}{2}}\tilde{\mathbf{p}}_{eq}\|^2 - 2t\tilde{\mathbf{u}}_{eq}^T\mathbf{\Omega}\tilde{\mathbf{p}}_{eq} - \frac{2\rho\lambda_1\ t\tilde{\mathbf{p}}_{eq}^T\mathbf{\Pi}\tilde{\mathbf{p}}_{eq}}{\sqrt{M_tN_r}} \\ & + \ |\bar{\mathbf{p}} - \sigma_f\mathbf{p} + \rho\boldsymbol{\lambda}_2\|^2 \\ \text{s.t. } & 0 \leq \sigma_f \leq \sqrt{\frac{P_t}{M_t}}. \end{aligned} \tag{65}$$

Note that $\mathbf{\Pi}$ is a function of $\mathbf{B}_{eq}$, while $\mathbf{B}_{eq}$ is a function of σ_f. By rearranging (65) as a function of σ_f, we find that it has the following form

$$\begin{aligned}\min_{\sigma_f} \ & a\sigma_f^2 + b\sigma_f + \frac{c_1\sigma_f^2 + c_2}{\sqrt{d_1\sigma_f^2 + d_2}} + \frac{e_1\sigma_f^2 + e_2}{\sqrt{f_1\sigma_f^2 + f_2}} \\ \text{s.t. } & 0 \leq \sigma_f \leq \sqrt{\frac{P_t}{M_t}} \end{aligned} \tag{66}$$

where

$$\begin{aligned} a &\triangleq t^2\tilde{\mathbf{p}}^H\mathbf{h}_3\mathbf{h}_3^H\tilde{\mathbf{p}} + \mathbf{p}^H\mathbf{p} > 0 \\ b &\triangleq -2\Re\{\mathbf{p}^H(\bar{\mathbf{p}} + \rho\boldsymbol{\lambda}_2)\} \\ c_1 &\triangleq -2t\tilde{\mathbf{u}}_{eq}^{iT}\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}\Re\{\tilde{\mathbf{p}}^{iH}\mathbf{h}_3\mathbf{h}_3^H\tilde{\mathbf{p}}\} \\ c_2 &\triangleq -2t\tilde{\mathbf{u}}_{eq}^{iT}\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}\Re\{\tilde{\mathbf{p}}^{iH}\tilde{\mathbf{p}}\}\frac{\sigma_1^2}{M_t} \\ d_1 &\triangleq \tilde{\mathbf{u}}_{eq}^{iT}\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}^i(\tilde{\mathbf{p}}^{iH}\mathbf{h}_3\mathbf{h}_3^H\tilde{\mathbf{p}}^i) > 0 \\ d_2 &\triangleq \tilde{\mathbf{u}}_{eq}^{iT}\mathbf{A}_{eq}\tilde{\mathbf{u}}_{eq}^i(\tilde{\mathbf{p}}^{iH}\tilde{\mathbf{p}}^i)\frac{\sigma_1^2}{M_t} > 0 \\ e_1 &\triangleq -2\rho\lambda_1\ t\frac{\tilde{\mathbf{p}}_{eq}^{iT}\tilde{\mathbf{p}}_{eq}}{\|\tilde{\mathbf{p}}_{eq}^i\|\sqrt{M_tN_r}}\Re\{\tilde{\mathbf{p}}^{iH}\mathbf{h}_3\mathbf{h}_3^H\tilde{\mathbf{p}}\} \\ e_2 &\triangleq -2\rho\lambda_1\ t\frac{\tilde{\mathbf{p}}_{eq}^{iT}\tilde{\mathbf{p}}_{eq}}{\|\tilde{\mathbf{p}}_{eq}^i\|\sqrt{M_tN_r}}\Re\{\tilde{\mathbf{p}}^{iH}\tilde{\mathbf{p}}\}\frac{\sigma_1^2}{M_t} \\ f_1 &\triangleq \tilde{\mathbf{p}}^{iH}\mathbf{h}_3\mathbf{h}_3^H\tilde{\mathbf{p}}^i > 0 \\ f_2 &\triangleq \tilde{\mathbf{p}}^{iH}\tilde{\mathbf{p}}^i\frac{\sigma_1^2}{M_t} > 0. \end{aligned} \tag{67}$$

All of these coefficients are real value, and can be easily calculated. Problem (66) can be solved by using a one-dimensional search method [34].

TABLE IV
BCD-TYPE ALGORITHM FOR PROBLEM (68)

0. set $k = 1$ and $\mathbf{q} = \mathbf{A}\mathbf{x}$
1. **repeat**
2. **for** $i \in \{1, 2, \ldots, m\}$
3. $b = \mathbf{A}(i,i)\mathbf{x}(i) - \mathbf{q}(i) + \mathbf{b}(i)$
4. $x = \frac{b}{|b|}$
5. $\mathbf{q} = \mathbf{q} + (x - \mathbf{x}(i))\mathbf{A}(:,i)$
6. $\mathbf{x}(i) = x$
7. **end**
8. $k = k + 1$
9. **until** some termination criterion is met

APPENDIX C
QUADRATIC OPTIMIZATION WITH UNIT MODULUS CONSTRAINTS

Let us consider a general problem:

$$\begin{aligned} &\min_{\mathbf{x}} \ \phi(\mathbf{x}) \triangleq \mathbf{x}^H\mathbf{A}\mathbf{x} - 2\Re\{\mathbf{x}^H\mathbf{b}\} \\ &\text{s.t. } |\mathbf{x}(i)| = 1, \quad \forall i. \end{aligned} \tag{68}$$

where $\mathbf{b} \in \mathbb{C}^{n\times 1}$ and $\mathbf{A} \in \mathbb{C}^{m\times m}$ is positive semidefinite.

We can use a BCD-type algorithm to address problem (68), that is, in each step we update one element of $\mathbf{x}$ while fixing the others. It is readily seen that $\phi(\mathbf{x})$, when considered as a function of $\mathbf{x}(i)$, can be expressed in terms of a quadratic function in the form of $\tilde{\phi}(\mathbf{x}) \triangleq a|\mathbf{x}(i)|^2 - 2\Re\{b^*\mathbf{x}(i)\}$ for some real number a and complex number b. Considering that $|\mathbf{x}(i)| = 1$, problem (68) reduces to

$$\max_{|\mathbf{x}(i)|=1} \ \Re\{b^*\mathbf{x}(i)\} \tag{69}$$

It follows that the optimum $\mathbf{x}(i)$ is given by $b/|b|$.

In what follows, we show how the complex number b can be easily obtained. On the one hand, we have

$$\frac{\partial\tilde{\phi}(\mathbf{x}(i))}{\partial\mathbf{x}^*(i)}|_{\mathbf{x}(i)=\tilde{\mathbf{x}}(i)} = a\tilde{\mathbf{x}}(i) - b. \tag{70}$$

On the other hand, we have

$$\frac{\partial\phi(\mathbf{x})}{\partial\mathbf{x}^*}|_{\mathbf{x}=\tilde{\mathbf{x}}} = \mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}. \tag{71}$$

Combining the above equations, we have $[\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}]_i = a\tilde{\mathbf{x}}(i) - b$, where $[.]_i$ denotes the ith element of a vector. By comparing the coefficient of $\tilde{\mathbf{x}}(i)$, we obtain $a = \mathbf{A}(i,i)$. It follows that

$$b = \mathbf{A}(i,i)\tilde{\mathbf{x}}(i) - [\mathbf{A}\tilde{\mathbf{x}}]_i + \mathbf{b}(i). \tag{72}$$

According to the above analysis, the entries of $\mathbf{x}$ can be recursively updated. The corresponding algorithm for updating $\mathbf{x}$ is summarized in Table IV, where the recursion in step 5 is due to the fact that $\mathbf{q}$ should be updated once $\mathbf{x}(i)$ is updated (which is done in step 6).

REFERENCES

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[2] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[3] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[4] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, no. 1, pp. 1–10, 2007.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[6] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[7] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[8] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[9] V. Venkateswaran and A. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.

[10] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[11] O. El Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. ICC*, Jun. 2012, pp. 3724–3729.

[12] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[13] J. Zhang, M. Haardt, I. Soloveychik, and A. Wiesel, "A channel matching based hybrid analog-digital strategy for massive multi-user MIMO downlink systems," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Jul. 2016, pp. 1–5.

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

[15] X. Zhai, Y. Cai, Q. Shi, M. Zhao, G. Y. Li, and B. Champagne, "Joint transceiver design with antenna selection for large-scale MU-MIMO mmWave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2085–2096, Sep. 2017.

[16] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretical security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2008.

[17] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[18] R. Liu and W. Trappe, Eds., *Securing Wireless Communications at the Physical Layer*. Norwell, MA, USA: Springer, 2009.

[19] J. Xu, L. Duan, and R. Zhang, "Surveillance and intervention of infrastructure-free mobile communications: A new wireless security paradigm," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 152–159, Aug. 2017.

[20] J. Xu, L. Duan, and R. Zhang, "Proactive eavesdropping via cognitive jamming in fading channels," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2790–2806, May 2017.

[21] J. Xu, L. Duan, and R. Zhang, "Proactive eavesdropping via jamming for rate maximization over Rayleigh fading channels," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 80–83, Feb. 2016.

[22] H. Cai, Q. Zhang, Q. Li, and J. Qin, "Proactive monitoring via jamming for rate maximization over MIMO Rayleigh fading channels," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2021–2024, Sep. 2017.

[23] C. Zhong, X. Jiang, F. Qu, and Z. Zhang, "Multi-antenna wireless legitimate surveillance systems: Design and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4585–4599, Jul. 2017.

[24] Y. Zeng and R. Zhang, "Wireless information surveillance via proactive eavesdropping with spoofing relay," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 8, pp. 1449–1461, Dec. 2016.
[25] X. Jiang, H. Lin, C. Zhong, X. Chen, and Z. Zhang, "Proactive eavesdropping in relaying systems," *IEEE Signal Process. Lett.*, vol. 24, no. 6, pp. 917–921, Jun. 2017.
[26] G. Ma, J. Xu, L. Duan, and R. Zhang, "Wireless surveillance of two-hop communications," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Jul. 2017, pp. 1–5.
[27] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM*, Hong Kong, Aug. 2013, pp. 375–386.
[28] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.
[29] E. Antonio-Rodriguez, R. Lopez-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "SINR optimization in wideband full-duplex MIMO relays under limited dynamic range," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Jun. 2014, pp. 177–180.
[30] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward MIMO full-duplex relaying with limited dynamic range," *Signal Process.*, vol. 127, pp. 266–281, Oct. 2016.
[31] Q. Shi, M. Hong, X. Fu, and T.-H. Chang. (Dec. 2017). "Penalty dual decomposition method for nonsmooth nonconvex optimization." [Online]. Available: https://arxiv.org/abs/1712.04767
[32] Q. Shi and M. Hong, "Penalty dual decomposition method with application in signal processing," in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 4059–4063.
[33] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.
[34] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[35] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Comput.*, vol. 15, no. 4, pp. 915–936, Apr. 2003.
[36] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.
[37] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2962–2973, Jun. 2012.
[38] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
[39] Z. Xiao, P. Xia, and X.-G. Xia, "Full-duplex millimeter-wave communication," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 136–143, Dec. 2017.
[40] A. Ruszczynski, *Nonlinear Optimization*. Princeton, NJ, USA: Princeton Univ. Press, 2011.
[41] Z. Lu and Y. Zhang, "Sparse approximation via penalty decomposition methods," *SIAM J. Optim.*, vol. 23, no. 4, pp. 2448–2478, 2013.
[42] Z. Lu and Y. Zhang, "Penalty decomposition methods for rank minimization," *Optim. Methods Softw.*, vol. 30, no. 3, pp. 531–558, May 2015.
[43] Q. Shi and M. Hong, "Spectral efficiency optimization for mmWave multiuser MIMO systems," *IEEE J. Sel. Topics Signal Process.*, to be published.

Yunlong Cai (S'07–M'10–SM'16) received the B.S. degree in computer science from Beijing Jiaotong University, Beijing, China, in 2004, the M.Sc. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From 2010 to 2011, he was a Post-Doctoral Fellow with the Electronics and Communications Laboratory, Conservatoire National des Arts et Metiers, Paris, France. Since 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. From 2016 to 2017, he was a Visiting Scholar with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include transceiver design for multiple-antenna systems, sensor array processing, mmWave communications, physical layer security, and cooperative and relay communications.

Cunzhuo Zhao received the B.S. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2015, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering. His research interests include algorithm design and analysis for MIMO communication systems, cooperative communications, and physical layer security.

Qingjiang Shi received the Ph.D. degree in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. From 2009 to 2010, he visited Prof. Z.-Q. (Tom) Luo's Research Group, University of Minnesota, Twin Cities, Minneapolis, MN, USA. In 2011, he was a Research Scientist with the Research and Innovation Center (Bell Labs China), Alcatel-Lucent, Shanghai, China. Since 2012, he has been with the School of Information and Science Technology, Zhejiang Sci-Tech University. From 2016 to 2017, he was a Research Fellow with Iowa State University. Since 2018, he has been a Professor with the School of Software Engineering, Tongji University. He is an Adjunct Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics. His interests include algorithm design and analysis with applications in signal processing, wireless communications, and machine learning.

He received the National Excellent Doctoral Dissertation Nomination Award in 2013, the Shanghai Excellent Doctoral Dissertation Award in 2012, and the Best Paper Award from the IEEE PIMRC'09 Conference. Since 2017, he has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Geoffrey Ye Li (S'93–M'95–SM'97–F'06) received the B.S.E. and M.S.E. degrees from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree from the Department of Electrical Engineering, Auburn University, Auburn, AL, USA, in 1994.

He was a Teaching Assistant and then a Lecturer with Southeast University, Nanjing, from 1986 to 1991, a Research and Teaching Assistant with Auburn University, from 1991 to 1994, and a Post-Doctoral Research Associate with the University of Maryland at College Park, College Park, MD, USA, from 1994 to 1996. He was with ATT Labs–Research at Red Bank, Middletown, NJ, USA, as a Senior and then a Principal Technical Staff Member from 1996 to 2000. Since 2000, he has been with the School of Electrical and Computer Engineering, Georgia Institute of Technology, as an Associate Professor and then a Full Professor. He has also been holding a Cheung Kong Scholar title at the University of Electronic Science and Technology of China since 2006. His general research interests include statistical signal processing and machine learning for wireless communications. In these areas, he has published over 200 journal papers in addition to over 40 granted patents and many conference papers. His publications have been cited over 30 000 times and he has been recognized as the Worlds Most Influential Scientific Mind, also known as a Highly-Cited Researcher, by Thomson Reuters almost every year.

Dr. Li was awarded the IEEE Fellow for his contributions to signal processing for wireless communications in 2005. He has won the 2010 IEEE ComSoc Stephen O. Rice Prize Paper Award, the 2013 IEEE VTS James Evans Avant Garde Award, the 2014 IEEE VTS Jack Neubauer Memorial Award, the 2017 IEEE ComSoc Award for Advances in Communication, and the 2017 IEEE SPS Donald G. Fink Overview Paper Award. He also received the 2015 Distinguished Faculty Achievement Award from the School of Electrical and Computer Engineering, Georgia Tech. He has organized and chaired many international conferences, including the Technical Program Vice-Chair of the IEEE ICC03, the Technical Program Co-Chair of the IEEE SPAWC11, a General Chair of the IEEE GlobalSIP14, and a Technical Program Co-Chair of the IEEE VTC16 (Spring). He has been involved in editorial activities for over 20 technical journals for the IEEE, including founding Editor-in-Chief of the *IEEE 5G Tech Focus*.

Benoit Champagne (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from the École Polytechnique de Montréal in 1983, the M.Sc. degree in physics from the Université de Montréal in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto in 1990. From 1990 to 1999, he was an Assistant and then an Associate Professor at INRS-Telecommunications, Université du Quebec, Montréal, QC, Canada. In 1999, he joined McGill University, Montréal, QC, where he is currently a Full Professor with the Department of Electrical and Computer Engineering; he also served as the Associate Chairman of Graduate Studies in the Department from 2004 to 2007. His research focuses on the study of advanced algorithms for the processing of communication signals by digital means. His interests span many areas of statistical signal processing, including detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing, where he has co-authored nearly 250 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Fonds de Recherche sur la Nature et les Technologies from the Government of Quebec, as well as some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi.

He has served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he was the Registration Chair, for the IEEE ICASSP 2004, the Co-Chair, Antenna and Propagation Track, for the IEEE VTCCFall 2004, the Co-Chair, Wide Area Cellular Communications Track, for the IEEE PIMRC 2011, the Co-Chair, Workshop on D2D Communications, for the IEEE ICC 2015, and the Publicity Chair, for the IEEE VTC-Fall 2016. He was an Associate Editor of the *EURASIP Journal on Advances in Signal Processing* from 2005 to 2007, the IEEE SIGNAL PROCESSING LETTERS from 2006 to 2008, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2010 to 2012, as well as a Guest Editor for two special issues of the *EURASIP Journal on Applied Signal Processing* published in 2007 and 2014, respectively.