Generalized Precoder Designs Based on Weighted MMSE Criterion for Energy Harvesting Constrained MIMO and Multi-User MIMO Channels

Changick Song, *Member, IEEE*, Jaehyun Park, *Member, IEEE*, Bruno Clerckx, *Member, IEEE*, Inkyu Lee, *Fellow, IEEE*, and Kyoung-Jae Lee, *Member, IEEE*

***Abstract*—This paper studies precoder designs for simultaneous wireless information and power transfer (SWIPT) in multi-input multi-output (MIMO) channels, where a transmitter sends information to information decoding (ID) users while satisfying the minimum energy requirement of energy harvesting users. In contrast to the previous designs focused only on maximum information rate (MIR), we propose a more general and simpler solution using the weighted minimum mean squared error (WMMSE) criterion. To solve the SWIPT-WMMSE problem which is generally non-convex, we suggest two different design schemes, separate and joint designs. Interestingly, it is shown that the joint design achieves optimal performance with a single initial point and a few iterations, while the separate design needs a large number of iterations and initial points to approach the optimum. Based on the observation, we propose a simple closed-form solution, which is shown to achieve near optimal performance with reduced complexity. The derived solution can be adopted in various pragmatic applications of MIMO communications, such as the MMSE, quality-of-service, equal error designs, as well as the MIR by adjusting the weight matrix. We also confirm that our design strategies are a great use for managing co-channel interference in multiple ID-user scenarios. Finally, simulation results demonstrate the efficiency of the proposed MIMO-SWIPT framework.**

***Index Terms*—SWIPT, MIMO, WMMSE, Multi-user, Closed-form solution.**

I. Introduction

In RECENT years, it has been recognized that radio frequency (RF) signals that transport information can at the same time be exploited as a significant energy source for devices to harvest RF energy. For this reason, *simultaneous wireless information and power transfer* (SWIPT) has appeared as promising technologies for the energy constrained wireless networks by providing mobile devices with convenient and perpetual energy supplies [1] [2].

In multiple-input multiple-output (MIMO) SWIPT systems, there exists a substantial tradeoff between link performance of information decoding (ID) users and the harvested energy of energy harvesting (EH) users according to the beam patterns at the transmitter. For the reason, efficient precoder designs have been an active research area over the last few years [3]–[11]. Various precoding strategies have been developed on interference [3] [11], relaying [4], and broadcast channels (BC) [5]–[9] to achieve the maximum information rate (MIR) of the ID-users with minimum energy requirement of the EH-users. In particular, for MIMO-SWIPT systems with a single ID-user, the authors in [7] completely characterized the rate-energy tradeoff region by developing the optimum precoder.

While such an MIR based design is useful for evaluating the system performance, it may be unrealistic especially for low-cost devices such as sensors or relays due to high encoding and decoding complexity. Further, for systems with finite length codewords, the minimum mean squared error (MMSE) or zero-forcing (ZF) criteria are often preferred rather than the MIR, because the channel gains are equally distributed across the subchannels, thereby achieving a good quality-of-service (QoS) in each subchannel. Therefore, joint transceiver designs between the precoder and the linear MMSE receiver

Manuscript received October 23, 2015; revised April 11, 2016 and August 5, 2016; accepted September 6, 2016. Date of publication September 15, 2016; date of current version December 8, 2016. This work was supported in part by the National Research Foundation (NRF) Funded by the Ministry of Science, ICT and Future Planning, Korean Government, under Grant NRF-2015R1C1A1A02036927 and Grant 2014R1A2A1A10049769 and in part by the Infrastructure and Transportation Technology Promotion Research Program Funded by the Ministry of Land Infrastructure and Transport, Korean Government, under Grant 16CTAP-C098206-02. The work of K.-J. Lee was supported by the Basic Science Research Program through the NRF Funded by MSIP, Korean Government, under Grant NRF-2016R1C1B2011921. This paper was presented at the IEEE Global Communications Conference Workshops, Austin, TX, USA, December 2014. The associate editor coordinating the review of this paper and approving it for publication was A. Kwasinski. *(Corresponding author: Kyoung-Jae Lee.)*

C. Song is with the Department of Information and Communications Engineering, Korea National University of Transportation, Chungju 27469, South Korea (e-mail: c.song@ut.ac.kr).

J. Park is with the Department of Electronics Engineering, Pukyong National University, Busan 48513, South Korea (e-mail: jaehyun@pknu.ac.kr).

B. Clerckx is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K., and also with the School of Electrical Engineering, Korea University, Seoul 02841, South Korea (e-mail: b.clerckx@imperial.ac.uk).

I. Lee is with the School of Electrical Engineering, Korea University, Seoul 02841, South Korea (e-mail: inkyu@korea.ac.kr).

K.-J. Lee is with the Department of Electronics and Control Engineering, Hanbat National University, Daejeon 34158, South Korea (e-mail: kyoungjae@hanbat.ac.kr).

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/TWC.2016.2609904

have widely been investigated in conventional MIMO systems [12]–[16]. However, such an MMSE metric has appeared only in a small number of works in SWIPT systems [17]–[19].

In this paper, we provide a generalized framework for precoder designs in MIMO SWIPT systems using the weighted MMSE (WMMSE) criterion. The proposed solution is flexibly applied to various pragmatic applications of MIMO channels such as QoS, MMSE, and equal error designs as well as the MIR by simply adjusting the weight matrix. Accordingly, the conventional SWIPT-MIR [7] and non-SWIPT WMMSE [12] schemes are casted into special cases of our solution.

As the SWIPT-WMMSE problem is generally non-convex, the strong duality based approaches which have been often adopted for precoder designs in MIMO-SWIPT systems [7] [19] may be no longer valid. To solve the problem, we first provide a separate transceiver design which finds a solution by alternatively updating the precoder at the transmitter and the receiver at the ID-user. When the ID-user has a single antenna, the separate design produces the optimal transmit beamformer in terms of both the WMMSE and the MIR. One interesting observation here is that the proposed separate design is attainable by a simple bisection method, while the MIR scheme in [7] needs to adopt complicated optimization tools to arrive at the same point. Meanwhile, in case of an ID-user with multiple antennas, the separate design may require many iterations and initial points to approach the optimal performance.

Next, we propose a joint SWIPT-WMMSE design which identifies the optimal solution with a single initial point. A careful examination of the Karush Kuhn Tucker (KKT) conditions reveals that the optimal precoding structure consists of two Lagrange multipliers, each of which is easily computed using the bisection process. It is shown that unlike the separate design, the joint design fulfils the optimum with a few iterations and a single initial point. Based on the observation, we further suggest a near optimal closed-form solution which is attainable even without iterations.

Our WMMSE solution is flexibly applied to various applications. With equal weights across the subchannels, our solution renders the best possible MSE-energy tradeoff region which is an MMSE counterpart of the rate-energy region [7]. It is shown that the sum-MSE gain of the proposed method is connected to a great advantage in the bit-error rate (BER) performance. By adjusting the weight matrix, it is also possible to obtain a QoS design which satisfies different target SNRs across the subchannels. For the QoS design, however, some feasibility conditions need to be considered. It is particularly interesting to observe that all boundary points of the optimal rate-energy region are achievable with our solution by setting the weight matrix as an eigenvalue matrix of an effective ID-user channel.

In the last part of the paper, we consider a scenario where the transmitter supports multiple ID-users at the same time and frequency, and suggest new multi-user precoder designs in the context of the SWIPT such as the block diagonalization (SWIPT-BD), the regularized BD (SWIPT-RBD), and the sum-rate maximization (SWIPT-SRM) schemes. It is shown that the proposed separate and joint designs for the single

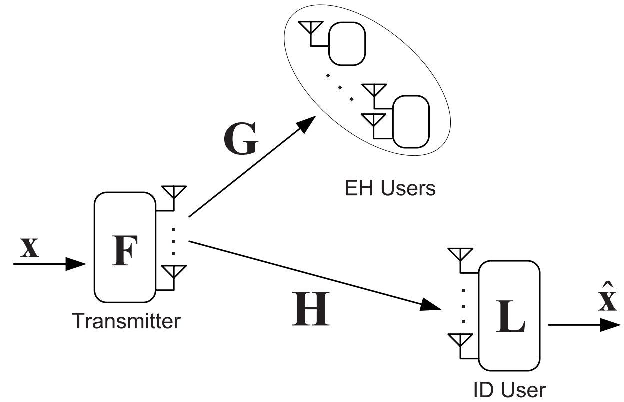

Fig. 1. MIMO broadcast systems for joint wireless information and energy transfer with linear transceivers.

ID-user are great use for designing the multiuser precoders which outperform the existing methods in [8] and [9]. Therefore, we can conclude that our work is not merely restricted to the case of a single ID-user, but provides a generalized framework for precoder designs in MIMO-SWIPT embracing both the single and multiple ID-user scenarios. Finally, simulation results demonstrate the efficiency of our designs in various circumstances.

Notations: Normal letters represent scalar quantities, boldface letters indicate vectors and boldface uppercase letters designate matrices. We use $\mathbb{R}_+$ and $\mathbb{C}$ to denote a set of positive real and complex numbers, respectively. The superscripts $(\cdot)^T$, $(\cdot)^H$, and $\mathbb{E}[\cdot]$ stand for the transpose, the conjugate transpose, and the expectation operators, respectively. $\mathbf{I}_N$ is an $N \times N$ identity matrix. In addition, $\mathrm{Tr}\,(\mathbf{A})$, $\det(\mathbf{A})$, $\|\mathbf{A}\|_2$, and $(\mathbf{A})_+$ indicate the trace, the determinant, the matrix 2-norm, and an element-wise $\max(\cdot, 0)$ operation of a matrix $\mathbf{A}$, respectively. The notation $\mathrm{blkdiag}[\mathbf{A}_1, \ldots, \mathbf{A}_K]$ and $\mathrm{diag}[\mathbf{A}]$ stand for a blockwise diagonal matrix with matrices $\mathbf{A}_1, \ldots, \mathbf{A}_K$ and a diagonal matrix with diagonal elements of $\mathbf{A}$ on its main diagonal, respectively.

II. System Model and Problem Formulation

As shown in Figure 1, we consider wireless multi-antenna broadcast channels where one transmitter equipped with N_T antennas transmits radio signals to EH users with total N_{E} antennas and an ID-user with N_{I} antennas.[1] Note that the multiple ID-user cases will be discussed in Section V. Throughout the paper, we assume quasi-static flat fading channels so that the baseband channels from the transmitter to the EH and ID users are simply represented by complex matrices $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{I}} \times N_T}$ and $\mathbf{G} \in \mathbb{C}^{N_{\mathrm{E}} \times N_T}$, respectively. It is also assumed that the transmitter knows all channel state information (CSI) $\mathbf{H}$ and $\mathbf{G}$, and each user obtains the corresponding CSI. Considering the spatial multiplexing scheme, $N_S \le \min(N_T, N_{\mathrm{I}})$ data-streams are transmitted simultaneously using the input signal vector $\mathbf{x} \in \mathbb{C}^{N_S \times 1}$ with $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{N_S}$.

[1] Basically, we assume that ID and EH users are physically separated, but the result can be easily applied to the co-located case by applying the power (or time) splitting methods as in [7].

With these assumptions, the estimated signal $\hat{\mathbf{x}} \in \mathbb{C}^{N_S \times 1}$ can be modeled by

$$\hat{\mathbf{x}} = \mathbf{L}(\mathbf{HFx} + \mathbf{n}), \tag{1}$$

where $\mathbf{L} \in \mathbb{C}^{N_S \times N_I}$ represents the linear receiver at the ID-user, $\mathbf{F} \in \mathbb{C}^{N_T \times N_S}$ is the linear precoder at the transmitter, which is subject to a power constraint $\mathrm{Tr}(\mathbf{FF}^H) \le P_T$, and $\mathbf{n} \in \mathbb{C}^{N_I \times 1}$ indicates the receiver noise with $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_I})$.

We define an estimation error at the ID-user as $\mathbf{e} = \gamma^{-1}\hat{\mathbf{x}} - \mathbf{x}$ whose weighted sum-MSE equals $\mathcal{M}(\gamma, \mathbf{F}, \mathbf{L}) = \mathrm{Tr}(\mathbf{WR}_e)$ with an error covariance matrix $\mathbf{R}_e \triangleq \mathbb{E}[\mathbf{ee}^H]$. Here, $\mathbf{W} \in \mathbb{R}_+^{N_S \times N_S}$ denotes a non-negative diagonal weight matrix and the scaling parameter $\gamma \in \mathbb{R}_+$ plays a role of scalar MMSE receiver at the ID-user [20], which is helpful for simplifying derivations.[2] Note that without consideration of the EH users, the WMMSE optimum precoder which we denote by $\mathbf{F}_{\mathrm{ID}}$ has been well investigated in literature [12] [13].

In the meantime, the total sum harvested energy E at the EH users is expressed as

$$E = \delta \mathbb{E}[\|\mathbf{GFx}\|^2] = \delta \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) \tag{2}$$

where δ is a constant accounting for the harvesting efficiency, which is assumed to be $\delta = 1$ unless stated otherwise. Let $\sqrt{g_1}$ and $\mathbf{v}_{g,1} \in \mathbb{C}^{N_T \times 1}$ denote the largest singular value of $\mathbf{G}$ and the corresponding right singular vector, respectively. Then, it is known that the energy beamformer $\mathbf{F}_{\mathrm{EH}} = \sqrt{P_T}[\mathbf{v}_{g,1}\ \mathbf{0}_{N_T \times (N_S-1)}]$ achieves the maximum energy $E_{\max} = P_T g_1$ [3] [7] and the corresponding sum-MSE equals $\mathcal{M}_{EH} = \mathcal{M}(\gamma, \mathbf{F}_{\mathrm{EH}}, \mathbf{L})$. In contrast, the sum-MSE and the energy attainable by the MMSE precoder $\mathbf{F}_{\mathrm{ID}}$ are denoted by $\mathcal{M}_{\min}$ and E_{ID}, respectively.

Now, consider the case where both the EH and ID users operate at the same time and frequency. Then, we can draw an MSE-energy tradeoff region as

$$\mathcal{R}_{\mathcal{M},E} \triangleq \big\{(\mathcal{M}, E) : \mathcal{M} \ge \mathrm{Tr}(\mathbf{WR}_e),\ E \le \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}),\ \mathrm{Tr}(\mathbf{FF}^H) \le P_T\big\}$$

and each boundary point of $\mathcal{R}_{\mathcal{M},E}$ is identified by solving the following SWIPT-WMMSE optimization problem

$$\begin{aligned}&\min_{\gamma,\mathbf{F},\mathbf{L}}\ \ \mathrm{Tr}(\mathbf{WR}_e)\\ &s.t. \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) \ge \bar{E},\ \ \mathrm{Tr}(\mathbf{FF}^H) \le P_T\end{aligned} \tag{3}$$

where $E_{\mathrm{ID}} \le \bar{E} \le E_{\max}$.

The above problem is convex over a precoder (or a receiver) when a receiver (or a precoder) and γ are fixed, and strictly quasi-convex with respect to γ when both $\mathbf{F}$ and $\mathbf{L}$ are given. However, it is jointly non-convex over $\{\gamma, \mathbf{F}, \mathbf{L}\}$, and therefore the conventional strong duality based approaches [7] [19] may be no longer valid. Our goal in the first part of the paper is to offer an efficient way to find a solution of problem (3) in a single ID-user scenario and determine the best possible MSE-energy region $\mathcal{R}_{\mathcal{M},E}$.

[2] γ is important especially when we consider the precoder $\mathbf{F}$ only for a fixed receiver $\mathbf{L}$, since in this case γ is the only way to control the noise at the ID-user.

III. Separate Transceiver Designs

In this section, we solve problem (3) based on the KKT conditions. First, we propose a method which computes the optimum precoder with respect to the WMMSE, and then suggest an optimization method which alternatively updates the transmitter and the ID receiver. It is shown that the proposed design exhibits faster and monotonic convergence over the conventional methods. Several interesting observations are also made in some special cases.

Utilizing the expression of $\mathbf{e}$ and the assumptions made in the previous section, we formulate the Lagrangian as

$$\begin{aligned}\mathcal{L}(\lambda, \mu, \gamma, \mathbf{F}, \mathbf{L}) = \mathrm{Tr}\big(\mathbf{WR}_e\big) &- \lambda\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) - \bar{E}\big)\\ &+ \mu\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\big),\end{aligned} \tag{4}$$

where $\mathbf{R}_e = (\gamma^{-1}\mathbf{LHF} - \mathbf{I}_{N_S})(\gamma^{-1}\mathbf{LHF} - \mathbf{I}_{N_S})^H + \gamma^{-2}\mathbf{LL}^H$. Then, the following KKT conditions provide necessary conditions for optimality:

$$\mathbf{LHFF}^H\mathbf{H}^H + \mathbf{L} = \gamma\,\mathbf{F}^H\mathbf{H}^H \tag{5}$$

$$\mathbf{H}^H\mathbf{L}^H\mathbf{WLHF} - \bar{\lambda}\mathbf{G}^H\mathbf{GF} + \bar{\mu}\mathbf{F} = \gamma\,\mathbf{H}^H\mathbf{L}^H\mathbf{W} \tag{6}$$

$$\mathrm{Tr}\big(\mathbf{WLHFF}^H\mathbf{H}^H\mathbf{L}^H + \mathbf{WLL}^H\big) = \gamma\,\mathrm{Tr}\big(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W}\big) \tag{7}$$

$$\begin{aligned}&\lambda \ge 0; \quad \mu \ge 0,\\ &\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) \ge \bar{E},\\ &\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) \le P_T\end{aligned} \tag{8}$$

$$\bar{\mu}\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\big) = 0 \tag{9}$$

$$\bar{\lambda}\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) - \bar{E}\big) = 0, \tag{10}$$

where we define $\bar{\lambda} \triangleq \lambda\gamma^2$, $\bar{\mu} \triangleq \mu\gamma^2$, equations (5)-(7) are derived from the zero gradient conditions by using some rules of differentiation [21], and (9) and (10) come from the complement slackness conditions.

Let us define two matrices that will be used for our derivations as

$$\mathbf{Y} = \mathbf{H}^H\mathbf{L}^H\mathbf{WLH} + \frac{\mathrm{Tr}(\mathbf{WLL}^H)}{P_T}\mathbf{I}_{N_T}$$

and

$$\mathbf{Z} = \mathbf{G}^H\mathbf{G} - \frac{\bar{E}}{P_T}\mathbf{I}_{N_T}. \tag{11}$$

Here, $\mathbf{Y}$ is a positive definite matrix, but $\mathbf{Z}$ is in general positive indefinite which makes the analysis challenging. Then, each of the transmitter and the receiver can be optimized separately as in the following.

Theorem 1: For a given receiver $\mathbf{L}$, *a solution of problem (3) is attained by*

$$\mathbf{F} = \gamma\bar{\mathbf{F}}(\bar{\lambda}) = \gamma\,\big(\mathbf{Y} - \bar{\lambda}\mathbf{Z}\big)^{-1}\mathbf{H}^H\mathbf{L}^H\mathbf{W}, \tag{12}$$

where $\gamma = \sqrt{\frac{P_T}{Tr(\bar{\mathbf{F}}(\bar{\lambda})\bar{\mathbf{F}}(\bar{\lambda})^H)}}$. *Denoting* $J(\bar{\lambda}) = Tr\big(\bar{\mathbf{F}}(\bar{\lambda})^H\mathbf{Z}\bar{\mathbf{F}}(\bar{\lambda})\big)$, *we have* $\bar{\lambda} = 0$ *if* $J(0) \ge 0$. *Otherwise,* $\bar{\lambda}$ *can be chosen to satisfy* $J(\bar{\lambda}) = 0$ *using line search methods over the range* $0 < \bar{\lambda} < 1/\kappa$ *where* $\kappa \triangleq \|\mathbf{ZY}^{-1}\|_2^2$.

Proof: Let us first consider the Lagrange dual function $g(\lambda, \mu) = \min_{\gamma,\mathbf{F}} \mathcal{L}(\lambda, \mu, \gamma, \mathbf{F}, \mathbf{L})$ with fixed λ, μ, and $\mathbf{L}$.

Then, ignoring constant terms, the problem of minimizing $\mathcal{L}$ over $\mathbf{F}$ and γ is equivalently

$$\min_{\gamma,\mathbf{F}} \ \gamma^{-2}\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{K}\mathbf{F}) - 2\gamma\,\mathrm{Tr}(\Re(\mathbf{WLHF}))\big), \tag{13}$$

where $\mathbf{K} \triangleq \mathbf{H}^H\mathbf{L}^H\mathbf{WLH} - \bar{\lambda}\mathbf{G}^H\mathbf{G} + \bar{\mu}\mathbf{I}_{N_T}$. Now, suppose that at least one eigenvalue of $\mathbf{K}$ is non-positive real with corresponding eigenvector $\mathbf{v} \in \mathbb{C}^{N_T\times 1}$. Then, it can be shown that problem (13) becomes unbounded below with $\mathbf{F} = [\mathbf{v}\ \mathbf{0}_{N_T\times(N_S-1)}]$ as $\gamma \to 0^+$. Therefore, to obtain a bounded optimal value of problem (3), the optimal $\bar{\lambda}$ and $\bar{\mu}$ must be chosen to satisfy $\mathbf{K} \succ 0$.

Assuming $\mathbf{K} \succ 0$, we now check from the KKT condition (6) that the optimal precoder $\mathbf{F}(\bar{\lambda},\bar{\mu}) = \gamma\mathbf{K}^{-1}\mathbf{H}^H\mathbf{L}^H\mathbf{W}$ can be expressed as a function of two lagrange multipliers $\bar{\lambda}$ and $\bar{\mu}$. Also, by leveraging the result in (6), it follows that

$$\begin{aligned}\gamma\,\mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W}) &= \mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{WLHF}\\ &\quad -\bar{\lambda}\mathbf{F}^H\mathbf{G}^H\mathbf{GF} + \bar{\mu}\mathbf{F}^H\mathbf{F})\\ &= \mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{WLHF}) - \bar{\lambda}\bar{E} + \bar{\mu}P_T\end{aligned}$$

where the second equality is due to slackness conditions (9) and (10). Because $\gamma\,\mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W}) = \mathrm{Tr}(\mathbf{WLHFF}^H\mathbf{H}^H\mathbf{L}^H) + \mathrm{Tr}(\mathbf{WLL}^H)$ in (7), we can show that

$$\bar{\mu} = \frac{1}{P_T}\big(\bar{\lambda}\bar{E} + \mathrm{Tr}(\mathbf{WLL}^H)\big). \tag{14}$$

This relationship allows us to reduce the precoder $\mathbf{F}(\bar{\lambda},\bar{\mu})$ to a function with a single parameter $\bar{\lambda}$. Also, since $\bar{\mu} > 0$, we have $\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) = P_T$ which implies that $\gamma = \big(P_T/\mathrm{Tr}(\bar{\mathbf{F}}(\bar{\lambda})\bar{\mathbf{F}}(\bar{\lambda})^H)\big)^{1/2}$ from (9).

Before we proceed with our proof further, let us present the following two useful lemmas.

Lemma 1: The optimal $\bar{\lambda}$ takes a value over the range $0 \le \bar{\lambda} < 1/\kappa$. (Proof: Appendix A)

Lemma 2: $J(\bar{\lambda})$ is a monotonic increasing function for $0 \le \bar{\lambda} < 1/\kappa$ and we have $\lim_{\bar{\lambda}\to 1/\kappa} J(\bar{\lambda}) > 0$. (Proof: Appendix B)

With the aid of Lemma 1 and 2, the optimal $\bar{\lambda}$ can be obtained as follows: First, it is true from (10) that $\bar{\lambda}\big(\mathrm{Tr}(\bar{\mathbf{F}}(\bar{\lambda})^H\mathbf{G}^H\mathbf{G}\bar{\mathbf{F}}(\bar{\lambda})) - \gamma^{-2}\bar{E}\big) = 0$, i.e., $\bar{\lambda}J(\bar{\lambda}) = 0$. The result implies that if $J(0) \ge 0$, we must have $\bar{\lambda} = 0$, because otherwise $\bar{\lambda}J(\bar{\lambda})$ becomes greater than zero, which violates the slackness condition. In contrast, if $J(0) < 0$, one can always find a unique $\bar{\lambda} \ne 0$ such that $J(\bar{\lambda}) = 0$ through line search methods over the range $0 < \bar{\lambda} < 1/\kappa$, since $J(\bar{\lambda})$ monotonically increases with $J(1/\kappa) > 0$. Finally, as the solution satisfying the KKT necessary conditions (6)-(10) is unique, the resulting precoder (12) is also sufficient for optimality, and the proof is completed. ■

Lemma 3: When γ and $\mathbf{F}$ are given, a solution of problem (3) is expressed as

$$\mathbf{L} = \gamma\bar{\mathbf{L}} = \gamma\mathbf{F}^H\mathbf{H}^H(\mathbf{HFF}^H\mathbf{H}^H + \mathbf{I}_{N_I})^{-1}. \tag{15}$$

Proof: The proof is trivial, and thus omitted. ■

Based on the results in Theorem 1 and Lemma 3, problem (3) can be solved in an alternating fashion between the transmitter and the receiver as summarized in Algorithm 1. The sum-MSE monotonically decreases in each updating process of the transmitter and the receiver. Therefore, the inner iteration of Algorithm 1 guarantees the convergence to a local minimum. However, due to joint non-convexity of problem (3), N_G different initial points are employed so that the resulting local minimum gets closer to the global minimum. This requires additional outer loop iterations.

Algorithm 1 Separate Transceiver Design

for $n = 1 : N_G$ **do**
  Initialize $\bar{\mathbf{F}}^{(n)}$ and $\gamma^{(n)} = \big(P_T/\mathrm{Tr}(\bar{\mathbf{F}}^{(n)}(\bar{\mathbf{F}}^{(n)})^H)\big)^{1/2}$
  repeat
    Compute $\mathbf{L}^{(n)}$ using (15) with given $\bar{\mathbf{F}}^{(n)}$ and $\gamma^{(n)}$
    Update $\mathbf{F}^{(n)} = \gamma_j\bar{\mathbf{F}}^{(n)}$ using (12) with given $\mathbf{L}^{(n)}$
  until convergence
end for
Select the best solution among $\big\{(\mathbf{F}^{(n)},\mathbf{L}^{(n)})$ for $n = 1,\ldots,N_G\big\}$

Next, let us examine the optimal beamforming solutions for the special case with $N_\mathrm{I} = N_S = 1$. In this case, the transmit and receive matrices $\mathbf{F}$ and $\mathbf{L}$ reduce to a column vector $\mathbf{f} \in \mathbb{C}^{N_T\times 1}$ and a scalar value $l \in \mathbb{C}$, respectively, and the weight matrix $\mathbf{W}$ can be ignored without loss of generality. Then, we obtain the following corollary.

Corollary 1: In the case of the MISO channel ($\mathbf{H} \equiv \mathbf{h}^H \in \mathbb{C}^{N_T\times 1}$), the MMSE optimal beamformer is independent of the receiver l and is given by

$$\mathbf{f} = \gamma\bar{\mathbf{f}}(\bar{\lambda}) = \frac{\gamma\mathbf{A}^{-1}\mathbf{h}}{1 + \mathbf{h}^H\mathbf{A}^{-1}\mathbf{h}}, \tag{16}$$

where $\gamma = \sqrt{P_T}/\|\bar{\mathbf{f}}(\bar{\lambda})\|$ and $\mathbf{A} \triangleq P_T^{-1}\mathbf{I}_{N_T} - \bar{\lambda}\mathbf{Z}$. Let us denote $J(\bar{\lambda}) = \bar{\mathbf{f}}(\bar{\lambda})^H\mathbf{Z}\bar{\mathbf{f}}(\bar{\lambda})$. If $J(0) \ge 0$, we set $\bar{\lambda} = 0$ and otherwise, we find a unique $\bar{\lambda}$ satisfying $J(\bar{\lambda}) = 0$ through the line search over $0 < \bar{\lambda} < 1/\zeta$ with $\zeta = \|\mathbf{Z}(\mathbf{hh}^H + \frac{1}{P_T}\mathbf{I}_{N_T})^{-1}\|_2^2$.

Proof: Assuming $N_S = N_\mathrm{I} = 1$ and following the same argument in Theorem 1, we obtain the solution as

$$\mathbf{f} = \gamma\left(\mathbf{hh}^H + \frac{1}{P_T}\mathbf{I}_{N_T} - \nu\mathbf{Z}\right)^{-1}\mathbf{h}, \tag{17}$$

where $\nu \triangleq \bar{\lambda}/|l|^2$. Since ν can be chosen to satisfy $\nu J(\nu) = 0$ over the range $0 \le \nu < 1/\zeta$, a particular choice of l does not affect the beamformer $\mathbf{f}$. Thus, by invoking some matrix inversion lemma, (17) is equivalently expressed as (16), and the proof is completed. ■

In fact, when we deal with a MISO channel, the MMSE precoder also maximizes the end-to-end SNR [13] [22]. Therefore, our beamforming strategy in Corollary 1 is essentially equivalent to the solution in [7, Corollary 3.1] which was developed for the rate maximization. Note that while the previous approach utilizes the optimization tools such as the sub-gradient method, our method achieves the maximum rate by a simple bisection method, and thus provides more insights into the system design.

Finally, we would like to note that although a joint design method that will be presented in the subsequent section may yield an easier way to achieve the optimal solution, the separate design is still important, because the

WMMSE precoder in Theorem 1 can be applied to various multiuser MIMO scenarios. The details will be studied in Section V.

IV. Joint Optimal Designs

Although the separate design shows good performance, several initial points and iterations are required, which results in large computational complexity. In this section, we propose a joint design method to achieve the optimal solution with a single initial point. Then, we suggest a simple closed-form solution with a slight approximation. Note that our design strategy includes conventional SWIPT rate maximization [7] and non-SWIPT WMMSE schemes [12] [13] as special cases, and thus is more general.

For a given receiver structure in (15), the error covariance matrix in (4) equals

$$\begin{aligned}\mathbf{R}_e &= \bar{\mathbf{L}}(\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H + \mathbf{I}_{N_I})\bar{\mathbf{L}}^H - \bar{\mathbf{L}}\mathbf{H}\mathbf{F} - \mathbf{F}^H\mathbf{H}^H\bar{\mathbf{L}}^H + \mathbf{I}_{N_S}\\ &= \mathbf{I}_{N_S} - \mathbf{F}^H\mathbf{H}^H(\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H + \mathbf{I}_{N_I})^{-1}\mathbf{H}\mathbf{F}\\ &= (\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F} + \mathbf{I})^{-1}. \end{aligned} \tag{18}$$

Accordingly, from the joint optimization perspective, the Lagrangian $\mathcal{L}$ in (4) is given by

$$\begin{aligned}\mathcal{L} &= \mathrm{Tr}\big(\mathbf{W}(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F} + \mathbf{I})^{-1}\big)\\ &\quad -\lambda\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}) - \bar{E}\big) + \mu\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\big).\end{aligned} \tag{19}$$

In addition, ignoring constant terms with respect to $\mathbf{F}$, the Lagrange dual, which is defined by $g(\tilde{\lambda}, \mu) = \min_{\mathbf{F}} \mathcal{L}(\lambda, \mu, \mathbf{F})$ becomes

$$\begin{aligned}\min_{\mathbf{F}}\ & \mathrm{Tr}\big(\mathbf{W}(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F} + \mathbf{I})^{-1}\big)\\ &+\mathrm{Tr}\big(\mathbf{F}^H(\mu\mathbf{I}_{N_T} - \lambda\mathbf{G}^H\mathbf{G})\mathbf{F}\big).\end{aligned} \tag{20}$$

We see from (20) that the first trace term is positive for any matrix $\mathbf{F}$. Thus, to obtain a bounded optimal value, it must hold $\mu\mathbf{I}_{N_T} \succ \lambda\mathbf{G}^H\mathbf{G}$, which implies that the feasibility condition of problem (3) is now $\mu > \lambda g_1$. Let us define $\beta = \big(\mathrm{Tr}(\mathbf{W}\bar{\mathbf{L}}\bar{\mathbf{L}}^H)\big)^{1/2}$. Then, from (14), we have $\mu = \frac{\lambda\bar{E}+\beta^2}{P_T} > \lambda g_1$, which is equal to $0 \le \tilde{\lambda} < \frac{1}{P_T g_1 - \bar{E}}$ where $\tilde{\lambda} \triangleq \lambda/\beta^2$. We point out that problem (20) is generally non-convex with respect to $\mathbf{F}$ due to non-equal elements of weight matrix $\mathbf{W}$. Thus, the conventional strong duality based optimization methods [19] may be invalid.

To solve the problem, let us first define a positive definite matrix $\mathbf{B} = P_T^{-1}\mathbf{I} - \tilde{\lambda}\mathbf{Z}$. Then, the following eigenvalue decomposition holds as

$$\mathbf{B}^{-\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{B}^{-\frac{1}{2}} = \begin{bmatrix}\mathbf{V} & \tilde{\mathbf{V}}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}^H\\ \tilde{\mathbf{V}}^H\end{bmatrix}, \tag{21}$$

where $\mathbf{V} \in \mathbb{C}^{N_T\times N_S}$ and $\tilde{\mathbf{V}} \in \mathbb{C}^{N_T\times(N_T-N_S)}$ represent orthogonal matrices which form a basis for the range space and the null space of $\mathbf{B}^{-1/2}\mathbf{H}^H\mathbf{H}\mathbf{B}^{-1/2}$, respectively, and $\mathbf{\Lambda}$ is a diagonal matrix containing N_S non-zero positive real eigenvalues arranged in a decreasing order. Then, we are ready to show the following theorem.

Theorem 2: From a jointly optimal perspective with receiver $\mathbf{L}$, a solution of problem (3) is given as

$$\mathbf{F}(\tilde{\lambda}, \beta) = \mathbf{B}^{-\frac{1}{2}}\mathbf{V}\mathbf{\Phi}_F, \tag{22}$$

where $\mathbf{\Phi}_F$ is defined by $\mathbf{\Phi}_F \triangleq (\beta^{-1}\mathbf{W}^{1/2}\mathbf{\Lambda}^{-1/2} - \mathbf{\Lambda}^{-1})_+^{1/2}$ and $(\tilde{\lambda}, \beta)$ satisfies the following two conditions, i) $J_1(\tilde{\lambda}, \beta) = 1$ and ii) $\tilde{\lambda}J_2(\tilde{\lambda}, \beta) = 0$ where $J_1(\tilde{\lambda}, \beta) \triangleq Tr(\mathbf{\Phi}_F\mathbf{\Phi}_F^H)$ and $J_2(\tilde{\lambda}, \beta) \triangleq Tr(\mathbf{F}^H\mathbf{Z}\mathbf{F})$ with $0 \le \tilde{\lambda} < \frac{1}{P_T g_1 - \bar{E}}$ and $0 < \beta \le Tr(\mathbf{\Lambda}^{-1/2}\mathbf{W}^{1/2})$.

Proof: Applying the previous results $\tilde{\lambda} = \lambda/\beta^2$ and $\mu = \frac{\lambda\bar{E}+\beta^2}{P_T} > \lambda g_1$, the original Lagrangian in (4) can be modified as

$$\begin{aligned}\mathcal{L} &= \mathrm{Tr}\big(\mathbf{W}\mathbf{R}_e\big) - \lambda\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}) - \bar{E}\big)\\ &\quad+\frac{\lambda\bar{E}+\beta^2}{P_T}\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\big)\\ &= \mathrm{Tr}\big(\mathbf{W}\mathbf{R}_e\big) - \tilde{\lambda}\beta^2\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F})\\ &\quad+\frac{\lambda\bar{E}+\beta^2}{P_T}\mathrm{Tr}\big(\mathbf{F}^H\mathbf{F}\big) - \beta^2\\ &= \mathrm{Tr}\big(\mathbf{W}\mathbf{R}_e\big) - \tilde{\lambda}\beta^2\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) + \beta^2\left(\frac{\mathrm{Tr}(\mathbf{F}^H\mathbf{F})}{P_T} - 1\right).\end{aligned} \tag{23}$$

Then, since $\beta > 0$, the corresponding KKT conditions are derived as

$$\mathbf{L}\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H + \mathbf{L} = \gamma\,\mathbf{F}^H\mathbf{H}^H \tag{24}$$

$$(\mathbf{H}^H\mathbf{L}^H\mathbf{W}\mathbf{L}\mathbf{H} + \gamma^2\beta^2\mathbf{B})\mathbf{F} = \gamma\,\mathbf{H}^H\mathbf{L}^H\mathbf{W} \tag{25}$$

$$\mathrm{Tr}\big(\mathbf{W}\mathbf{L}\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H + \mathbf{W}\mathbf{L}\mathbf{L}^H\big) = \gamma\,\mathrm{Tr}\big(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W}\big) \tag{26}$$

$$\tilde{\lambda} \ge 0,\ \ \beta > 0,$$

$$\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) \ge 0,$$

$$\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) = P_T \tag{27}$$

$$\tilde{\lambda}\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) = 0. \tag{28}$$

Now, we examine the optimal structure of $\mathbf{F}$. Since the matrix $\mathbf{B}$ in (25) is invertible, setting $\mathbf{F} = \mathbf{B}^{-1/2}\check{\mathbf{F}}$ for any matrix $\check{\mathbf{F}}$ will not impose any restriction in our optimization process. Then, consider a general expressions for $\check{\mathbf{F}}$ as $\check{\mathbf{F}} = \check{\mathbf{F}}_\parallel + \check{\mathbf{F}}_\perp$ where $\check{\mathbf{F}}_\parallel$ and $\check{\mathbf{F}}_\perp$ represent two complementary components whose columns are parallel and orthogonal to the range space of $\mathbf{H}\mathbf{B}^{-1/2}$, respectively. Similarly, we obtain $\mathbf{L} = \mathbf{L}_\parallel + \mathbf{L}_\perp$ where $\mathbf{L}_\parallel$ and $\mathbf{L}_\perp$ indicate the complementary components whose rows are parallel and orthogonal to the column space of $\mathbf{H}\mathbf{B}^{-1/2}$, respectively. Then, by post-multiplying $\mathbf{L}_\perp^H$ to (24) and pre-multiplying $\mathbf{F}_\perp^H = \check{\mathbf{F}}_\perp^H\mathbf{B}^{-1/2}$ to (25), we can show that only the case of $\mathbf{L}_\perp = \mathbf{F}_\perp = \mathbf{0}$ satisfies the conditions. Thus, we have $\mathbf{F} = \mathbf{B}^{-1/2}\check{\mathbf{F}}_\parallel$ and $\mathbf{L} = \mathbf{L}_\parallel$ where $\check{\mathbf{F}}_\parallel = \mathbf{V}\mathbf{\Phi}_F$ and $\mathbf{L}_\parallel = \gamma\,\mathbf{\Phi}_L\mathbf{V}^H\mathbf{B}^{-1/2}\mathbf{H}^H$ for any matrices $\mathbf{\Phi}_F \in \mathbb{C}^{N_S\times N_S}$ and $\mathbf{\Phi}_L \in \mathbb{C}^{N_S\times N_S}$.

The remaining problem is to find $\mathbf{\Phi}_F$ and $\mathbf{\Phi}_L$. First, by examining (24) and (25), it follows that

$$\mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_F\mathbf{\Phi}_F^H\mathbf{\Lambda}\mathbf{\Phi}_L^H + \mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_L^H = \mathbf{\Phi}_F^H\mathbf{\Lambda}\mathbf{\Phi}_L^H \tag{29}$$

$$\mathbf{\Phi}_F^H\mathbf{\Lambda}\mathbf{\Phi}_L\mathbf{W}\mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_F + \beta^2\mathbf{\Phi}_F\mathbf{\Phi}_F^H = \mathbf{\Phi}_F^H\mathbf{\Lambda}\mathbf{\Phi}_L^H\mathbf{W}. \tag{30}$$

In (29), we see that $\mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_F$ is Hermitian, since all other terms are Hermitian. Using the same argument, $\mathbf{\Phi}_F^H\mathbf{\Lambda}\mathbf{\Phi}_L^H\mathbf{W}$ in (30) is also Hermitian. Accordingly, without loss of generality, we set $\mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_F = \mathbf{D}_1$, $\mathbf{\Phi}_L\mathbf{\Lambda}\mathbf{\Phi}_L^H = \mathbf{D}_2$, and $\mathbf{\Phi}_F\mathbf{\Phi}_F^H = \mathbf{D}_3$ for $N_S \times N_S$ positive real diagonal matrices $\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3 \succeq \mathbf{0}$ (see [12, Lemma 2] for more details). Substituting the result into (29) and (30), we obtain $\mathbf{D}_1^2 + \mathbf{D}_2 = \mathbf{D}_1$ and $\mathbf{W}\mathbf{D}_1^2 + \beta^2\mathbf{D}_3 = \mathbf{W}\mathbf{D}_1$. Also, from the structure of $\mathbf{D}_1$, it is true that $\mathbf{\Lambda}\mathbf{D}_3\mathbf{D}_2 = \mathbf{D}_1^2$. Thus, combining three equations with some mathematical manipulations, we have $\mathbf{D}_3 = \mathbf{\Phi}_F\mathbf{\Phi}_F^H = (\beta^{-1}\mathbf{W}^{1/2}\mathbf{\Lambda}^{-1/2} - \mathbf{\Lambda}^{-1})_+$.

Let us now apply the result $\mathbf{F} = \mathbf{B}^{-1/2}\mathbf{V}\mathbf{\Phi}_F$ to (23). Then, the Lagrangian becomes

$$\mathcal{L} = \mathrm{Tr}\big(\mathbf{W}\mathbf{R}_e\big) - \tilde{\lambda}\beta^2\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) + \beta^2\Big(\frac{\mathrm{Tr}(\mathbf{F}^H\mathbf{F})}{P_T} - 1\Big) \quad (31)$$

$$= \mathrm{Tr}\big(\mathbf{W}\mathbf{R}_e\big) + \beta^2\big(\mathrm{Tr}(\mathbf{\Phi}_F\mathbf{\Phi}_F^H) - 1\big). \quad (32)$$

This implies that if we find $\tilde{\lambda} \geq 0$ and $\beta > 0$ such that $\tilde{\lambda}\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) = 0$, $\mathrm{Tr}(\mathbf{F}^H\mathbf{Z}\mathbf{F}) \geq 0$, and $\mathrm{Tr}(\mathbf{\Phi}_F\mathbf{\Phi}_F^H) = 1$, the resulting solution $\mathbf{F} = \mathbf{B}^{-1/2}\mathbf{V}\mathbf{\Phi}_F$ fulfils all KKT conditions (24)-(28), and thus it is optimal. Meanwhile, since the optimal β satisfies $\mathrm{Tr}(\mathbf{\Phi}_F\mathbf{\Phi}_F^H) = 1$, we have $\mathrm{Tr}(\beta^{-1}\mathbf{W}^{1/2}\mathbf{\Lambda}^{-1/2}) \geq 1$, which is equivalent to $0 < \beta \leq \mathrm{Tr}(\mathbf{W}^{-1/2}\mathbf{\Lambda}^{1/2})$, and the proof is completed. ■

Theorem 2 illustrates that the complicated matrix optimization problem in (3) can be reduced to a simple problem of finding two positive scalars $\tilde{\lambda}$ and β which satisfy the KKT conditions (27) and (28). It turns out that the reduced problem can be easily solved with the aid of Lemma 4 given below.

Lemma 4: For $\tilde{\lambda} \geq 0$ and $\beta > 0$, $J_1(\tilde{\lambda}, \beta)$ is monotonically decreasing with respect to β for a given $\tilde{\lambda}$, while $J_2(\tilde{\lambda}, \beta)$ is monotonically increasing with respect to $\tilde{\lambda}$ for a given β.

Proof: See Appendix C. ■

First, let us define $\hat{\beta}$ as the optimal β for a given $\tilde{\lambda} = 0$, i.e.,

$$\hat{\beta} \triangleq \arg_\beta(J_1(0, \beta) = 1). \quad (33)$$

Then, if $J_2(0, \hat{\beta}) \geq 0$, we have $\tilde{\lambda} = 0$ because otherwise, $\tilde{\lambda}J_2(\tilde{\lambda}, \hat{\beta})$ becomes non-zero positive, which contradicts the slackness condition (28). In contrast, if $J_2(0, \hat{\beta}) < 0$, one can always obtain a unique $\tilde{\lambda} \neq 0$ such that $J_2(\tilde{\lambda}, \hat{\beta}) = 0$ through line search methods over the range $0 < \tilde{\lambda} < 1/(P_T g_1 - \bar{E})$, because $J_2(\tilde{\lambda}, \hat{\beta})$ is monotonic increasing with respect to $\tilde{\lambda}$ and $\lim_{\tilde{\lambda}\to 1/(P_T g_1 - \bar{E})} J_2(\tilde{\lambda}, \hat{\beta}) > 0$. Similarly, for a given $\tilde{\lambda}$, a unique β which satisfies $J_1(\tilde{\lambda}, \hat{\beta}) = 1$ can be found through the line search over $0 < \beta < \mathrm{Tr}(\mathbf{W}^{-1/2}\mathbf{\Lambda}^{1/2})$, since $J_1(\tilde{\lambda}, \beta)$ monotonically decreases with respect to β and we have $\lim_{\beta\to\mathrm{Tr}(\mathbf{W}^{-1/2}\mathbf{\Lambda}^{1/2})} J_1(\tilde{\lambda}, \beta) < 1 < \lim_{\beta\to 0^+} J_1(\tilde{\lambda}, \beta)$. Therefore, a solution is easily obtained by alternatively updating $\tilde{\lambda}$ and β using the bisection methods. The detailed iterative process is summarized in Algorithm 2. Note that as problem (20) is non-convex with respect to $\mathbf{F}$, there might be a myriad of suboptimal and optimal solutions other than (22), which also meet the KKT conditions. Nevertheless, the proof of Theorem 2 reveals that (22) represents at least one of the optimal solutions since all possible solutions satisfying the KKT conditions are equivalently transformed into the proposed form without losing the MSE optimality.

Algorithm 2 Joint Optimal Design

```
Set λ̃ = 0 and β = β̂
if J_2(0, β̂) < 0 then
  repeat
    Find β which satisfies J_1(λ̃, β) = 1 for a given λ̃
    Update λ̃ to satisfy J_2(λ̃, β) = 0 for a given β
  until λ̃ converges to a prescribed accuracy
end if
Compute F̂(λ̃, β) based on (22)
```

Algorithm 3 Closed-Form Design

```
Set λ̃ = 0 and β = β̂
if J_2(0, β̂) < 0 then
  Find λ̃ which satisfies J_2(λ̃, β̂) = 0
end if
Compute F(λ̃, β̂) based on (34)
```

A. Closed-Form Design

As will be shown later in Section VI, Algorithm 2 exhibits a faster convergence than Algorithm 1 even with a single initial point. Furthermore, in the majority of cases, one or two iterations are sufficient to achieve a converged performance. Motivated by this observation, in what follows, we provide a simple closed-form solution[3] which requires only two step bisection process as

$$\mathbf{F}(\tilde{\lambda}, \hat{\beta}) = \gamma\,\mathbf{B}^{-\frac{1}{2}}\mathbf{V}(\hat{\beta}^{-1}\mathbf{W}^{1/2}\mathbf{\Lambda}^{-1/2} - \mathbf{\Lambda}^{-1})_+^{1/2}. \quad (34)$$

Suppose that $\hat{\beta}$ has been determined as in (33). Then, if $J_2(0, \hat{\beta}) \geq 0$, we attain $\tilde{\lambda} = 0$. Otherwise, we find $\tilde{\lambda}$ such that $J_2(\tilde{\lambda}, \hat{\beta}) = 0$ utilizing the bisection method. Finally, we obtain a solution by normalizing the transmit power using γ so that $\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) = P_T$ is satisfied. It turns out that the resulting solution exhibits near optimal performance, because the KKT optimality conditions (26)-(28) still hold despite the modification in the structure of $\mathbf{F}$ due to γ. The case of $\tilde{\lambda} = 0$ is associated with the non-SWIPT WMMSE design in conventional MIMO channels. Accordingly, our closed-form solution also generalizes the previous works in [12] and [13]. The whole process to obtain (34) is presented in Algorithm 3.

B. Applications

In this section, we verify that the proposed SWIPT-WMMSE precoder in Theorem 2 can be applied to different applications by adjusting the weight matrix $\mathbf{W}$.

1) Fixed-Weight MMSE (FWM) Design: The MMSE design which minimizes a sum of symbol estimation error has widely been used in both academia and industry [23]. This is because the MMSE criterion yields good performance for practical systems such as the systems with finite-length codewords or imperfect CSI. In particular, the weighted MMSE provides a convenient way to handle the error in each subchannel. Therefore, the proposed solutions in (22) and (34) with a

[3]Here, we have adopted the term "closed-form" in the sense that no iteration is needed among different variables, e.g., β, $\tilde{\lambda}$, and γ.

fixed weight $\mathbf{W}$ offers practical alternatives to the max-rate design [7].

2) Equal Error (EE) Design: While the MMSE design minimizes the sum-MSE, it may result in unequal errors across subchannels. In the fixed-rate transmission schemes where N_S symbol streams are transmitted with identical modulation and coding schemes, the subchannel with the worst MSE would dominate the performance. In this case, it is desirable to make all MSEs equal while minimizing the sum-MSE. This can be done by setting $\mathbf{W} = \mathbf{I}_{N_S}$ and applying discrete Fourier transform (DFT) matrix $\mathbf{T} \in \mathbb{C}^{N_S \times N_S}$ as

$$\mathbf{F} = \mathbf{B}^{-\frac{1}{2}} \mathbf{V} \mathbf{\Phi}_F \mathbf{T}, \tag{35}$$

since the DFT matrix (or a Hadamard matrix) can make all diagonal elements of the error covariance matrix in (18) have the same values while maintaining their sum. We point out that since the DFT matrix $\mathbf{T}$ is unitary ($\mathbf{T}^H\mathbf{T} = \mathbf{T}\mathbf{T}^H = \mathbf{I}_{N_S}$), it does not affect the KKT optimality conditions (24)-(28), and thus (35) maintains optimality in terms of the MMSE. Such a precoder in (35) is called "*BER-based design [13]*" as the BER is minimized by maximizing the minimum SNR.

3) QoS Design: The QoS design may be useful when different types of information such as video and audio with various target SNRs are transmitted across the subchannels to an ID user. Defining a matrix $\mathbf{\Gamma} = \mathbf{\Phi}_F^2 \mathbf{\Lambda}$ in which the i-th diagonal element stands for the external SNR of the i-th subchannel, the QoS design boils down to finding the weight matrix as [12]

$$\mathbf{\Gamma} = \left(\frac{1}{\beta} \mathbf{W}^{1/2} \mathbf{\Lambda}^{1/2} - \mathbf{I}_{N_S} \right)_+ = \alpha \mathbf{D} \tag{36}$$

where $\mathbf{D}$ is a diagonal matrix representing relative target SNRs across subchannels and $\alpha > 0$ indicates a correction term to obtain true target SNRs meeting both the power and EH constraints. Then, the QoS weight matrix is given by $\mathbf{W} = (\mathbf{I}_{N_S} + \alpha \mathbf{D}) \mathbf{\Lambda}^{-1/2} \beta$.

Now, let us examine the feasibility condition of the QoS design. According to circumstances, the QoS design may be unavailable in SWIPT systems due to the following two conflicting constraints as

$$\mathrm{Tr}(\mathbf{F}\mathbf{F}^H) \leq P_T \quad \text{and} \quad \mathrm{Tr}(\mathbf{F}^H \mathbf{G}^H \mathbf{G} \mathbf{F}) \geq \bar{E}. \tag{37}$$

where $\mathbf{F} = \alpha \mathbf{B}^{-1/2} \mathbf{V} \mathbf{\Lambda}^{-1/2} \mathbf{D}$. In other words, to achieve the desired relative SNRs in $\mathbf{D}$, there must exist a positive real value α satisfying the above two conditions simultaneously. Thus, we obtain the following feasibility condition.

Lemma 5: To achieve a valid QoS design in SWIPT systems, the relative SNRs in $\mathbf{D}$ *must satisfy*

$$Tr(\mathbf{V}_h^H \mathbf{Z} \mathbf{V}_h \mathbf{\Lambda}_h^{-1} \mathbf{D}) \geq 0 \tag{38}$$

where a diagonal matrix $\mathbf{\Lambda}_h \in \mathbb{R}_+^{N_S \times N_S}$ *and an orthonormal matrix* $\mathbf{V}_h \in \mathbb{C}^{N_T \times N_S}$ *are acquired from the eigenvalue decomposition* $\mathbf{H}^H \mathbf{H} = \mathbf{V}_h \mathbf{\Lambda}_h \mathbf{V}_h^H$.

Proof: See Appendix D. ∎

Interestingly, the left-hand side of equation (38) is independent of $\tilde{\lambda}$ for any diagonal matrix $\mathbf{D}$. Therefore, it is important to check whether the relative SNRs in $\mathbf{D}$ satisfies (38) or not. Once we find such $\mathbf{D}$, we get a maximum value of $\alpha = P_T / \mathrm{Tr}(\mathbf{\Lambda}_h^{-1} \mathbf{D})$ and plugging this into (36) yields

$$\mathbf{W} = \left(\mathbf{I}_{N_S} + \frac{P_T \mathbf{D}}{\mathrm{Tr}(\mathbf{\Lambda}_h^{-1} \mathbf{D})} \right) \mathbf{\Lambda}^{-1/2} \beta.$$

Note that if we set $\mathbf{D} = \mathbf{I}$ and it is feasible, our solution reduces to a SWIPT-ZF design as in [22].

4) Maximum Information Rate (MIR) Design: The following lemma proves that a specific choice of the weight matrix $\mathbf{W}$ also leads our solution to the MIR design.

Lemma 6: The joint WMMSE design in Algorithm 2 with a weight matrix $\mathbf{W} = \beta^{-2} \mathbf{\Lambda}$ *maximizes the information rate in MIMO SWIPT systems.*

Proof: See Appendix E. ∎

We now confirm from Lemma 6 that the previous max-rate design in [7] is categorized into a special case of our SWIPT-WMMSE solution with $\mathbf{W} = \beta^{-2} \mathbf{\Lambda}$. The result also implies that a near optimal rate can be achieved with a closed-form solution as proposed in (34), which provides a useful insight into the system design. We point out that an iterative sub-gradient process was adopted in [7] to obtain the solution.

V. Multiple ID-User Scenarios

In this section, we show how our methods for a single ID-user can be applied to the multiple ID-user scenarios. We consider three different multiuser precoding schemes in the context of SWIPT, namely, the SWIPT-BD, the SWIPT-RBD, and the SWIPT-SRM. While the latter gradually approaches the maximum sum-rate using iterative methods, the former is rather dedicated to a suboptimal closed-form design by focusing on the interference and noise power mitigation. For ease of presentation, we restrict our discussion to the sum-rate performance of all ID-users, although various design metrics may be also relevant by adjusting the weight matrix as in the previous sections.

A. Multiuser Signal Model

To describe the general K user signal model let us generalize some notations defined in Section II. Let N_{I_k} be the number of antennas at the k-th ID-user with $\sum_{k=1}^{K} N_{I_k} = N_I$. Then, considering full spatial multiplexing for each user, the transmitted and estimated signals $\mathbf{x}$ and $\hat{\mathbf{x}}$ in (1) can be redefined as $\mathbf{x} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_K^T]^T$ and $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \ldots, \hat{\mathbf{x}}_K^T]^T$ where $\mathbf{x}_k \in \mathbb{C}^{N_{I_k}}$ and $\hat{\mathbf{x}}_k \in \mathbb{C}^{N_{I_k}}$ represent the k-th ID-user signals with $\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^H] = \mathbf{I}_{N_{I_k}}, \forall k$.

Similarly, we have $\mathbf{F} = [\mathbf{F}_1, \ldots, \mathbf{F}_K]$, $\mathbf{L} = \mathrm{blkdiag}[\mathbf{L}_1, \ldots, \mathbf{L}_K]$, $\mathbf{W} = \mathrm{blkdiag}[\mathbf{W}_1, \ldots, \mathbf{W}_K]$, and $\mathbf{H} = [\mathbf{H}_1^T, \ldots, \mathbf{H}_K^T]^T$ where $\mathbf{F}_k \in \mathbb{C}^{N_T \times N_{I_k}}$, $\mathbf{L}_k \in \mathbb{C}^{N_{I_k} \times N_{I_k}}$, $\mathbf{W}_k \in \mathbb{C}^{N_{I_k} \times N_{I_k}}$, and $\mathbf{H}_k \in \mathbb{C}^{N_{I_k} \times N_T}$ denote the transmit/receive beamforming, weight, and channel matrices corresponding to the k-th ID user, respectively. The estimated signal vector is then given by

$$\hat{\mathbf{x}}_k = \mathbf{L}_k (\mathbf{H}_k \mathbf{F}_k \mathbf{x}_k + \sum_{i \neq k}^{K} \mathbf{H}_k \mathbf{F}_i \mathbf{x}_i + \mathbf{n}_k) \tag{39}$$

with $\mathbf{n}_k \in \mathbb{C}^{N_{I_k}}$ being the noise vector of user k with $\mathbb{E}[\mathbf{n}_k\mathbf{n}_k^H] = \mathbf{I}_{N_{I_k}}, \forall k$.

Assuming the optimal codewords at the transmitter and an information lossless receiver $\mathbf{L}_k$ at the k-th ID-user, the achievable sum-rate is computed by solving the following problem

$$\begin{aligned}\max_{\mathbf{F}} \quad & \sum_{k=1}^{K} \log_2 \det(\mathbf{F}_k^H\mathbf{H}_k^H\mathbf{R}_{v,k}^{-1}\mathbf{H}_k\mathbf{F}_k + \mathbf{I}_{N_{I_k}}) \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{F}_k^H\mathbf{F}_k) \le P_T, \quad \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{F}_k^H\mathbf{G}^H\mathbf{G}\mathbf{F}_k) \ge \bar{E}.\end{aligned} \tag{40}$$

where $\mathbf{R}_{v,k} = \sum_{i\neq k}^{K} \mathbf{H}_k\mathbf{F}_i\mathbf{F}_i^H\mathbf{H}_k^H + \mathbf{I}_{N_{I_k}}$.

B. SWIPT-BD

A simple way to improve the sum-rate in (40) is to remove the co-channel interference $\sum_{i\neq k}^{K} \mathbf{H}_k\mathbf{F}_i\mathbf{F}_i^H\mathbf{H}_k^H$. To this end, the authors in [8] have proposed a BD scheme in the context of SWIPT by combining the conventional BD beamformers in [24] with a power allocation scheme to meet the constraints in (40). However, since the scheme still relies on the same beam directions as in the case of non-SWIPT systems, some performance loss is observed. To tackle the problem, we suggest a new SWIPT-BD solution taking the beam directions to the EH users into account, thereby achieving an improved rate-energy tradeoff region.

The proposed BD scheme occurs in two steps, i.e., $\mathbf{F}_k = \mathbf{Q}_{\mathrm{BD},k}\mathbf{T}_k$ where $\mathbf{Q}_{\mathrm{BD},k} \in \mathbb{C}^{N_T\times N_{I_k}}$ denotes a unitary pre-beamformer taking responsibility for the co-channel interference mitigation and $\mathbf{T}_k \in \mathbb{C}^{N_{I_k}\times N_{I_k}}$ represents a post-beamformer for intra-user interference management. First, $\mathbf{Q}_{\mathrm{BD},k}, \forall k$ can be achieved by applying the QR-decomposition to the channel inversion matrix of $\mathbf{H}$ which is given as follows.

$$\mathbf{H}_{\mathrm{CI}}^{\dagger} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1} = [\mathbf{H}_{m_1}^{\dagger}, \mathbf{H}_{m_2}^{\dagger}, \ldots, \mathbf{H}_{m_K}^{\dagger}].$$

Specifically, applying the QR-decomposition to the k-th submatrix $\mathbf{H}_{m_k}^{\dagger} \in \mathbb{C}^{N_T\times N_{I_k}}$, we attain $\mathbf{Q}_{\mathrm{BD},k}$ being composed of N_{I_k} orthonormal basis vectors of $\mathbf{H}_{m_k}^{\dagger}$ so that the resulting $\mathbf{Q}_{\mathrm{BD},k}$ fulfils zero-interference condition $\mathbf{H}_i\mathbf{Q}_{\mathrm{BD},k} = \mathbf{0}, \forall i \neq k$ [24]. Thus, we now have K independent single-user channels $\hat{\mathbf{x}}_{\mathrm{BD},k} = \mathbf{L}_k(\widetilde{\mathbf{H}}_k\mathbf{T}_k\mathbf{x}_k + \mathbf{n}_k), \forall k$ instead of (39) where $\widetilde{\mathbf{H}}_k = \mathbf{H}_k\mathbf{Q}_{\mathrm{BD},k}$, The harvested energy at the EH-users in (2) is also modified as $\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{T}_k^H\mathbf{G}_k^H\mathbf{G}_k\mathbf{T}_k)$ where $\mathbf{G}_k = \mathbf{G}\mathbf{Q}_{\mathrm{BD},k}$ denotes the k-th effective sub-channel to the EH-users.

Remark 1: Due to the structural constraint of the BD scheme, the maximum achievable energy of SWIPT-BD is given by $E_{BD,\max} = P_T g_{BD}$ *where* $g_{BD} \triangleq \max_k \|\mathbf{G}_k^H\mathbf{G}_k\|_2^2$, *which is typically smaller than the maximum energy* $E_{\max} = P_T g_1$ *with* $g_1 = \|\mathbf{G}^H\mathbf{G}\|_2^2$ *for the single ID user scenarios. Thus, the SWIPT-BD may be infeasible in a high EH requirement greater than* $E_{BD,\max}$. *In this case, one may instead adopt a user selection scheme which supports only a selected user among multiple ID-users. Hence, the best possible rate-energy region of SWIPT-BD is obtained by taking time sharing between the single user selection scheme and the SWIPT-BD.*

In what follows, we focus on the feasible range $\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{T}_k^H\mathbf{G}_k^H\mathbf{G}_k\mathbf{T}_k) \le E_{\mathrm{BD},\max}$. The SWIPT-BD optimization problem is then formulated as

$$\begin{aligned}\min_{(\mathbf{T}_k,\mathbf{L}_k)\forall k,\gamma} \quad & \sum_{k=1}^{K} \mathrm{Tr}\Big(\mathbf{W}_k\mathbb{E}\Big[\big(\frac{1}{\gamma}\hat{\mathbf{x}}_{\mathrm{BD},k} - \mathbf{x}_k\big)\big(\frac{1}{\gamma}\hat{\mathbf{x}}_{\mathrm{BD},k} - \mathbf{x}_k\big)^H\Big]\Big) \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{T}_k^H\mathbf{T}_k) \le P_T, \ \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{T}_k^H\mathbf{G}_k^H\mathbf{G}_k\mathbf{T}_k) \ge \bar{E},\end{aligned} \tag{41}$$

where $0 \le \bar{E} \le E_{\mathrm{BD},\max}$. Define $\mathbf{B}_k = P_T^{-1}\mathbf{I}_{N_{I_k}} - \lambda\mathbf{Z}_k$ for $\lambda \ge 0$ and $\mathbf{Z}_k = \mathbf{Q}_{\mathrm{BD},k}^H\mathbf{Z}\mathbf{Q}_{\mathrm{BD},k}$, and an eigenvalue decomposition $\mathbf{B}_k^{-\frac{1}{2}}\widetilde{\mathbf{H}}_k^H\widetilde{\mathbf{H}}_k\mathbf{B}_k^{-\frac{1}{2}} = \mathbf{V}_k\mathbf{\Lambda}_k\mathbf{V}_k^H$ with an orthonormal matrix $\mathbf{V}_k \in \mathbb{C}^{N_{I_k}\times N_{I_k}}$ and a diagonal matrix $\mathbf{\Lambda}_k \in \mathbb{C}^{N_{I_k}\times N_{I_k}}$ with non-zero eigenvalues. Then, the solution of (41) is obtained as in the following theorem.

Theorem 3: The joint optimal solution of problem (41) is given by

$$\begin{aligned}\mathbf{T}_k(\lambda,\beta) &= \mathbf{B}_k^{-\frac{1}{2}}\mathbf{V}_k\mathbf{\Phi}_{T,k}, \ \forall k \\ \mathbf{L}_k &= \gamma\,\mathbf{F}_k^H\mathbf{H}_k^H(\mathbf{H}_k\mathbf{F}_k\mathbf{F}_k^H\mathbf{H}_k^H + \mathbf{I}_{N_{I_k}})^{-1}, \ \forall k\end{aligned} \tag{42}$$

where $\mathbf{\Phi}_{T,k}$ *is defined by* $\mathbf{\Phi}_{T,k} \triangleq (\beta^{-1}\mathbf{W}_k^{1/2}\mathbf{\Lambda}_k^{-1/2} - \mathbf{\Lambda}_k^{-1})_+^{1/2}$ *and* (λ, β) *satisfies the following two conditions, i)* $J_1(\tilde{\lambda}, \beta) = 1$ *and ii)* $\lambda J_2(\lambda, \beta) = 0$ *where* $J_1(\lambda, \beta) \triangleq \sum_{k=1}^{K} Tr(\mathbf{\Phi}_{T,k}\mathbf{\Phi}_{T,k}^H)$ *and* $J_2(\lambda, \beta) \triangleq \sum_{k=1}^{K} Tr(\mathbf{T}_k^H\mathbf{Z}_k\mathbf{T}_k)$ *with* $0 \le \lambda < \frac{1}{P_T g_{BD} - \bar{E}}$ *and* $0 < \beta \le \sum_{k=1}^{K} Tr(\mathbf{\Lambda}_k^{-1/2}\mathbf{W}_k^{1/2})$.

Proof: The proof is similar to Theorem 2. Details are omitted. ■

Combining $\mathbf{T}_k$ in (42) with $\mathbf{Q}_{\mathrm{BD},k}$, we finally obtain a SWIPT-BD solution

$$\mathbf{F}_{\mathrm{BD},k} = \mathbf{Q}_{\mathrm{BD},k}\mathbf{B}_k^{-\frac{1}{2}}\mathbf{V}_k\mathbf{\Phi}_{T,k}, \forall k. \tag{43}$$

Here, two positive scalars λ and β can be computed exploiting either algorithm 2 or 3. As shown in Lemma 6, setting $\mathbf{W}_k = \beta^{-2}\mathbf{\Lambda}_k, \forall k$ gives rise to a rate maximizing SWIPT-BD scheme. Other design criteria can also be fulfilled by adjusting the weight matrix.

C. SWIPT-RBD

The BD scheme only care about the interference, and thus is vulnerable to noise-limited scenarios. To tackle the problem, regularization techniques have been widely investigated in multiuser MIMO systems to suppress the noise power as well as the interference power [16], [25], [26]. In this section, we suggest a new RBD solution in the context of the SWIPT utilizing the results in Section III.

Let receivers $\mathbf{L}_k$ be arbitrarily fixed. Then, the RBD beamformer can be achieved by solving the following

WMMSE problem

$$\min_{\mathbf{F}_k,\forall k,\gamma} \sum_{k=1}^{K} \mathrm{Tr}\Big(\mathbb{E}\Big[\big(\frac{1}{\gamma}\hat{\mathbf{x}}_k - \mathbf{W}_k\mathbf{x}_k\big)\big(\frac{1}{\gamma}\hat{\mathbf{x}}_k - \mathbf{W}_k\mathbf{x}_k\big)^H\Big]\Big)$$
$$\text{s.t.} \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{F}_k^H\mathbf{F}_k) \le P_T, \ \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{F}_k^H\mathbf{G}^H\mathbf{G}\mathbf{F}_k) \ge \bar{E}_{\mathrm{BD}}. \tag{44}$$

A major difference of (44) from (41) is that the weight matrix $\mathbf{W}_k \in \mathbb{C}^{N_T \times N_{I_k}}$ here plays a role of a target channel matrix for the k-th ID-user. It is thus expected that the resulting solution of (44) minimizes the interference plus noise power while keeping the desired signal $\mathbf{L}_k\mathbf{H}_k\mathbf{F}_k\mathbf{x}_k$ in (39) as close as possible to some target signal waveform $\mathbf{W}_k\mathbf{x}_k$. Thus, by adjusting the weight matrix $\mathbf{W}_k$, one can obtain various solutions for different design criteria. Also, as the problem takes the noise into account, it resolves the noise enhancement issue of the SWIPT-BD. The solution of problem (44) is summarized below.

Theorem 4: The optimal solution of problem (44) is given by

$$\mathbf{F}_k = \gamma\,\bar{\mathbf{F}}_k(\eta) = \gamma\,\big(\bar{\mathbf{Y}} - \eta\mathbf{Z}\big)^{-1}\mathbf{H}_k^H\mathbf{L}_k^H\mathbf{W}_k, \tag{45}$$

with γ and $\bar{\mathbf{Y}}_k$ being equal to

$$\gamma = \sqrt{\frac{P_T}{\sum_{k=1}^{K} Tr(\bar{\mathbf{F}}_k(\eta)\bar{\mathbf{F}}_k(\eta)^H)}}$$
$$\text{and} \quad \bar{\mathbf{Y}} = \mathbf{H}^H\mathbf{L}^H\mathbf{L}\mathbf{H} + \frac{Tr(\mathbf{L}\mathbf{L}^H)}{P_T}\mathbf{I}_{N_I}, \tag{46}$$

respectively, and $\eta \ge 0$ denotes a positive real value satisfying the following conditions. Denoting $J(\eta) = \sum_{k=1}^{K} Tr\big(\bar{\mathbf{F}}_k(\eta)^H\mathbf{Z}\bar{\mathbf{F}}_k(\eta)\big)$, we have $\eta = 0$ if $J(0) \ge 0$. Otherwise, η is chosen such that $J(\eta) = 0$ using line search methods over the range $0 < \eta < 1/\kappa$ where $\kappa \triangleq \|\mathbf{Z}\bar{\mathbf{Y}}^{-1}\|_2^2$.

Proof: The proof simply follows from Theorem 1, and thus omitted for brevity. ■

As our goal in this section is to find an efficient BD solution, a sensible choice of the weight (target) matrix is $\mathbf{W}_k = \mathbf{L}_k^H\mathbf{H}_k\mathbf{F}_{\mathrm{BD},k}, \forall k$. Also, various information lossless receivers such as a square unitary matrix can be adopted for $\mathbf{L}_k$ without affecting the sum-rate performance in (40). Thus, we finally obtain a simple closed-form RBD beamformer as

$$\mathbf{F}_{\mathrm{RBD},k} = \gamma\left(\mathbf{H}^H\mathbf{H} + \frac{N_I}{P_T}\mathbf{I}_{N_I} - \eta\mathbf{Z}\right)^{-1}\mathbf{H}_k^H\mathbf{H}_k\mathbf{F}_{\mathrm{BD},k}, \tag{47}$$

with γ and η being similarly obtained as in Theorem 4.

D. SWIPT-SRM

Utilizing the equivalence relationship between the sum-rate maximization and WMMSE problems [27], [28], problem (40) can be transformed into an equivalent problem

$$\min_{\gamma,\mathbf{F},\mathbf{W},\mathbf{L}} \sum_{k=1}^{K}\Big\{\mathrm{Tr}(\mathbf{W}_k\mathbf{R}_{e,k}) - \log_2|\mathbf{W}_k \ln 2| - \frac{N_{I_k}}{\ln 2}\Big\}$$
$$\text{s.t.} \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{F}_k^H\mathbf{F}_k) \le P_T, \ \sum_{k=1}^{K}\mathrm{Tr}(\mathbf{F}_k^H\mathbf{G}^H\mathbf{G}\mathbf{F}_k) \ge \bar{E}, \tag{48}$$

Algorithm 4 SWIPT-SRM

for $n = 1 : N_G$ **do**
  Initialize $\mathbf{F}^{(n)}$
  repeat
    Compute $\mathbf{L}^{(n)}$ using (49) for a given $\mathbf{F}^{(n)}$
    Set $\mathbf{W}^{(n)} := \mathrm{diag}[\mathbf{W}]$ using (50) for given $\mathbf{L}^{(n)}$ and $\mathbf{F}^{(n)}$
    Update $\mathbf{F}^{(n)}$ using (51) for given $\mathbf{W}^{(n)}$ and $\mathbf{L}^{(n)}$
  until convergence
end for
Select the best solution among $\{\mathbf{F}^{(n)}$ for $n = 1, \ldots, N_G\}$

where $\mathbf{R}_{e,k} = \mathbb{E}[(\gamma^{-1}\hat{\mathbf{x}}_k - \mathbf{x}_k)(\gamma^{-1}\hat{\mathbf{x}}_k - \mathbf{x}_k)^H]$. Unlike (40), the problem is now convex with respect to each of the variables, and thus easily solved by alternatively updating each filter for given others until convergence. The detailed updating process is summarized in Algorithm 4 where the solutions used in each iteration step are computed by

$$\mathbf{L}_k = \gamma\,\mathbf{F}_k^H\mathbf{H}_k^H(\mathbf{H}_k\mathbf{F}_k\mathbf{F}_k^H\mathbf{H}_k^H + \mathbf{R}_{\upsilon,k})^{-1}, \forall k, \tag{49}$$
$$\mathbf{W}_k = \mathbf{F}_k^H\mathbf{H}_k^H\mathbf{R}_{\upsilon,k}^{-1}\mathbf{H}_k\mathbf{F}_k + \mathbf{I}_{N_{I_k}}, \forall k, \tag{50}$$

and

$$\mathbf{F} = \gamma\,\big(\mathbf{Y} - \bar{\lambda}\mathbf{Z}\big)^{-1}\mathbf{H}^H\mathbf{L}^H\mathbf{W} \tag{51}$$

with γ, $\bar{\lambda}$, $\mathbf{Y}$, and $\mathbf{Z}$ being similarly defined as in Theorem 1.[4]

VI. Numerical Results

In this section, we demonstrate the efficiency of the proposed SWIPT-WMMSE designs through the numerical results. For ease of presentation, we assume that all ID and EH users have the same distance, i.e, d_h and d_g meters from the transmitter, respectively. Considering the suburban propagation (pathloss exponent 3), the channel matrices are then generated according to the Rayleigh pathloss model, $\mathbf{H} = d_h^{-3/2}\widetilde{\mathbf{H}}$ and $\mathbf{G} = d_g^{-3/2}\widetilde{\mathbf{G}}$, in which each element of $\widetilde{\mathbf{H}}$ and $\widetilde{\mathbf{G}}$ is drawn from the i.i.d. circularly symmetric standard normal distribution. We assume that the noise power at each ID-user equals −30dBm, and therefore 1 energy unit in our standard signal model (1) amounts to 1μW. We set $d_h = d_g = 10$(m) and assume 50% energy conversion efficiency, i.e., $\delta = 0.5$ of the EH users.

A. Examples for Single ID-User Scenarios

From Figure 2 to 7, we consider a single ID-user in MIMO SWIPT systems with $N_S = N_I = N_E = N$. First, Figure 2 compares the weighted sum-MSE of various schemes in a sample channel with $P_T = 20$dBm, $E_{\max} = 0.67$mW and a weight matrix $\mathbf{W} = \mathrm{diag}[4, 3, 2, 1]$. We set the target energy $\bar{E} = 0.8E_{\max}$ so that the EH user can harvest at least 80% of the maximum energy. The proposed WMMSE designs show faster and monotonic convergence compared to

[4]We would like to mention that a similar approach has been adopted in [9] for multiuser SWIPT systems. A major difference of our work from the previous design is that we find the WMMSE beamformer $\mathbf{F}$ in (50) as a closed-form with the aid of Theorem 1, while the method in [9] requires to compute iterative sub-gradient process to find $\mathbf{F}$ in each iteration.

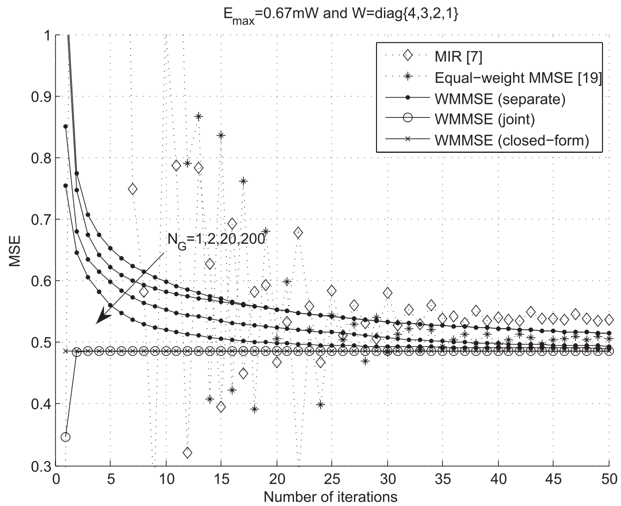

Fig. 2. Convergence trend of various schemes with $\bar{E} = 0.8E_{\max}$ and $P_T = 20$dBm in $N_T = N = 4$ systems with $K = 1$.

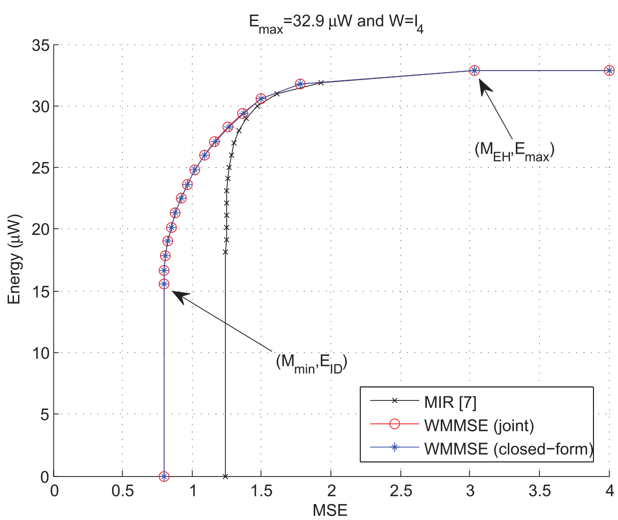

Fig. 3. MSE-energy region comparison with $P_T = 20$dBm and $\mathbf{W} = \mathbf{I}_4$ in $N_T = N = 4$ systems with $K = 1$.

the conventional sub-gradient based schemes developed for the MIR [7] and the equal weight MMSE [19] in SWIPT systems. In particular, the proposed joint and closed-form designs achieve the minimum MSE with few iterations, while dozens of iterations are still needed for the sub-gradient methods until convergence possibly with some MSE loss. It is also interesting to observe that the joint design attains the optimal performance with a single initial point, whereas the separate design requires more than 20 random initial points to approach the optimum due to the non-convexity of problem (3).

Figure 3 examines MSE-energy tradeoff regions in a sample channel with $P_T = 10$dBm, $E_{\max} = 32.9\mu$W, and a weight matrix $\mathbf{W} = \mathbf{I}_4$. We can check that the closed-form solution accomplishes the minimum MSE regardless of the target energy levels. Note that the MSE gain of our designs in Figures 2 and 3 will lead to an improved BER performance as shown in the next figure.

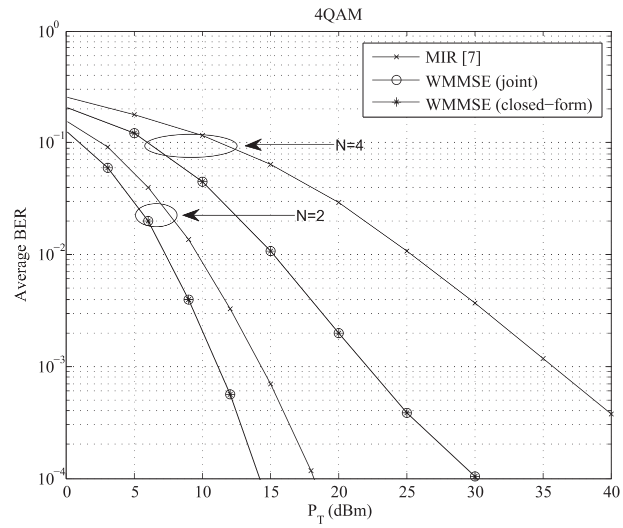

Fig. 4. Average BER performance comparison with $\bar{E} = 0.8E_{\max}$ in $N_T = 4$ systems with $K = 1$.

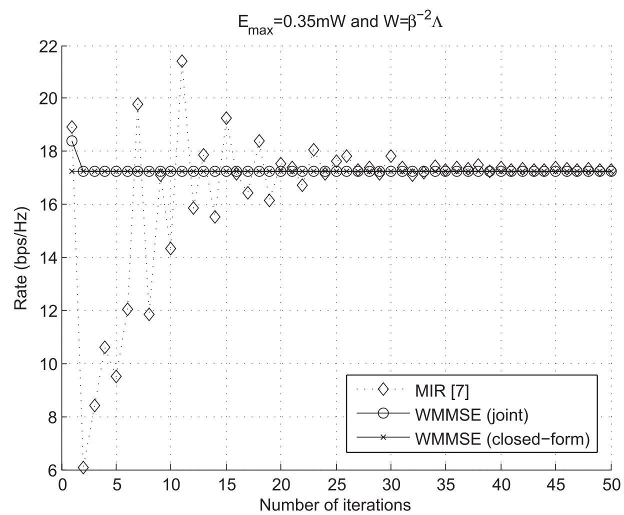

Fig. 5. Convergence trend of various schemes with $\bar{E} = 0.8E_{\max}$ and $P_T = 20$dBm in $N_T = N = 4$ systems with $K = 1$.

Figure 4 exhibits the average BER performance of the schemes. We assume the QPSK modulation in each subchannel and average the performance over 10^5 random channel realizations. In each channel instance, we set the target energy as $\bar{E} = 0.8E_{\max}$. We adopt the weight matrix $\mathbf{W} = \mathbf{I}_4$ and apply the DFT matrix $\mathbf{T}$ as in (35) for our schemes, which amounts to the equal error design in Section IV-B.2. By examining the figure, we can recognize that the proposed WMMSE designs obtain approximately 4 and 13 dB advantages over the MIR design at the BER of 10^{-3} in $N = 2$ and $N = 4$ systems, respectively. The gain is attributed to the fact that the rate-based precoder allocates more resource to the stronger subchannel while our WMMSE precoder neutralizes the channel gains across subchannels, thereby maximizing the minimum SNR.

In Figures 5 and 6, we investigate the performance of the WMMSE designs with respect to the information rate.

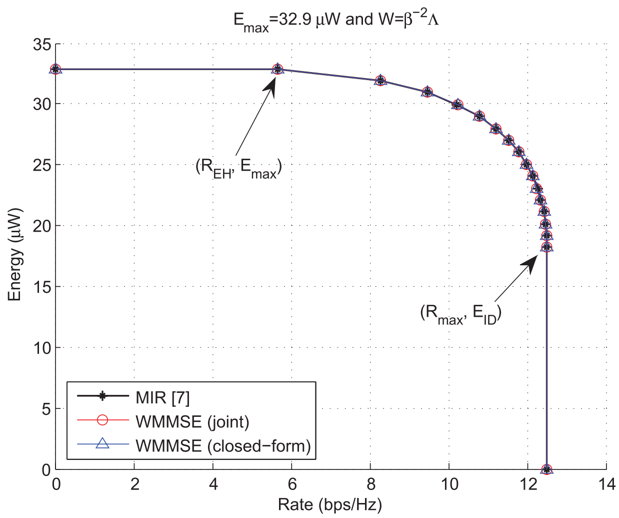

Fig. 6. Rate-energy region comparison with $P_T = 10$dBm in $N_T = N = 4$ systems with $K = 1$.

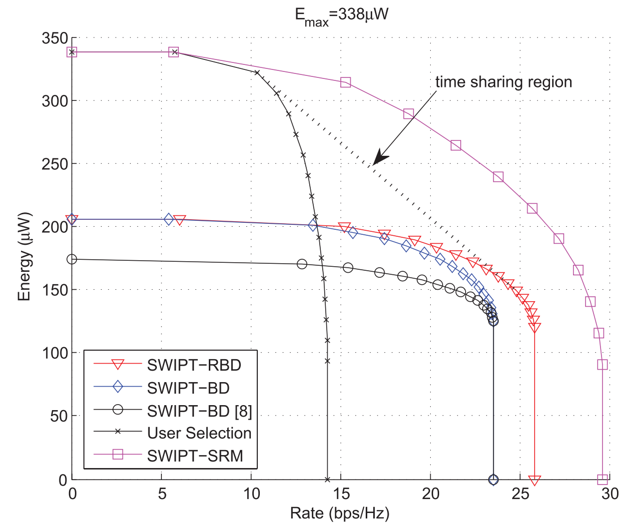

Fig. 8. Rate-energy region comparison with $P_T = 15$dBm in $\{2, 2, 2, 2\} \times 8$ with $K = 4$ and $N_I = 8$.

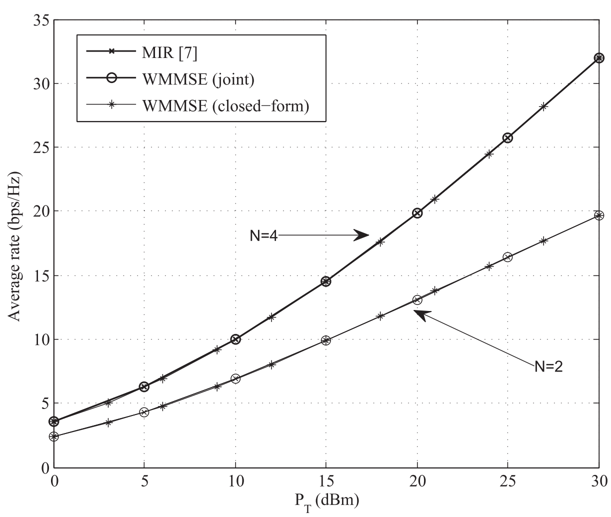

Fig. 7. Average rate performance comparison with $\bar{E} = 0.8E_{\max}$ and $\mathbf{W} = \beta^{-2}\mathbf{\Lambda}$ in $N_T = 4$ systems with $K = 1$.

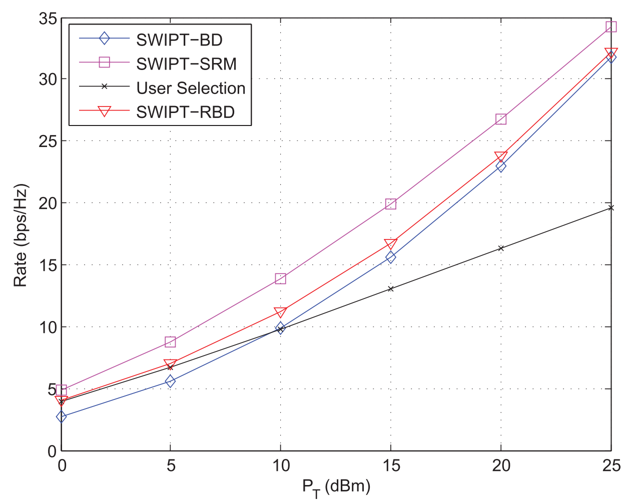

Fig. 9. Average rate performance comparison in $\{2, 2, 2\} \times 6$ with $\bar{E} = 0.8E_{\mathrm{BD,max}}$, $K = 3$, and $N_I = 6$.

In Figure 5, we present a convergence trend of various schemes in an example channel with $P_T = 20$dBm and $E_{\max} = 0.35$mW. It is interesting to observe that the proposed closed-form method achieves the maximum information rate at one sitting with $\mathbf{W} = \beta^{-2}\mathbf{\Lambda}$ while the MIR design still needs dozens of iterations to arrive at the same point. By comparing the rate-energy tradeoff regions in Figure 6, one can also confirm that our solution maintains optimality for all target energy levels.

Figure 7 illustrates the average sum-rate performance of various schemes according to the transmit power P_T to demonstrate the efficiency of our proposed schemes for all channel variations. We average the performance over 10^3 random channels and set $\bar{E} = 0.8E_{\max}$ in each channel instance. It is clear that the proposed WMMSE designs maximizes the average sum-rate without breaking the EH constraint for a given channel.

B. Examples for Multiple ID-User Scenarios

In Figures 8 and 9, we consider general K-user systems with N_T-transmit and N_{I_k}-receive antennas at the transmitter and the k-th ID-user, respectively, which we denote by $\{N_{I_1}, \ldots, N_{I_K}\} \times N_T$. Figure 8 compares the rate-energy tradeoff regions of various schemes in $\{2, 2, 2, 2\} \times 8$. It is shown that the proposed SWIPT-BD achieves an improved rate-energy region over the conventional scheme [8]. The regularization method further enhances the sum-rate performance. Although the BD based algorithms suffer from restricted energy supply for the EH-users, we may obtain a larger convex region by taking time sharing between the BD and the single user selection scheme. Note that all schemes are based on closed-form solutions except the SWIPT-SRM which exhibits the best possible rate-energy region depending on numerous iterations and random initial points ($N_G = 20$).

Figure 9 exhibits average sum-rate performance in $\{2, 2, 2\} \times 6$ systems. We set $\bar{E} = 0.8E_{\mathrm{BD,max}}$ for all schemes, which implies that the EH-users should attain energy at least 80% of the maximum achievable energy of SWIPT-BD. It is confirmed that the proposed SWIPT-(R)BD and SWIPT-SRM show improved rate performance over the single-user optimal algorithm based selection as SNR goes to high. At low SNR, some performance loss is observed for SWIPT-BD due to the noise enhancement effect. We check that the SWIPT-RBD resolves such an issue by taking the noise into account during the optimization process.

Through the simulation results so far, we confirm that the proposed WMMSE beamforming scheme is flexibly adjusted under various design criteria such as the MMSE, the MIR, and the minimum BER by selecting the weight matrix appropriately. In addition, we demonstrate that it is also a great use for multiuser MIMO scenarios with co-channel interference. Hence, our design method offers a useful framework for the precoder design in MIMO-SWIPT systems.

VII. Conclusion

In this paper, we have studied WMMSE precoder designs for MIMO-SWIPT systems. First, considering a single ID-user, we presented the separate design which requires multiple iterations and initial points to obtain a solution. Then, we proposed the joint transceiver design to achieve the optimal solution with a single initial point, based on which a new optimal closed-form solution has also been introduced with two-step bisection process. The proposed designs are adopted in many applications by adjusting the weight matrix $\mathbf{W}$ and includes conventional SWIPT-MIR and non-SWIPT WMMSE designs as special cases. We have also extended our method to the multiple ID-user scenarios by adopting the BD and SRM structures in the context of the SWIPT, and thereby providing a generalized framework for precoder designs in MIMO-SWIPT channels including both the single and multiple-user cases.

Appendix

A. Proof of Lemma 1

Applying the result in (14), it is easy to verify that the feasibility condition $\mathbf{K} \succ 0$ becomes $\mathbf{Y} \succ \bar{\lambda}\mathbf{Z}$, which is equivalent to $\bar{\lambda}^{-1}\mathbf{I} - \mathbf{Z}\mathbf{Y}^{-1} \succ \mathbf{0}$. Then, by definition of the matrix two norm κ, the feasibility condition holds as long as $\bar{\lambda}^{-1} > \kappa$ [29]. Therefore, combining with the condition $\lambda \geq 0$ in (8), we obtain Lemma 1.

B. Proof of Lemma 2

First, we set the derivative of the function $J(x)$ with respect to x as

$$\begin{aligned}\frac{\partial J(x)}{\partial x} &= \mathrm{Tr}\left(\left(\frac{\partial J(x)}{\partial \bar{\mathbf{F}}(x)}\right)^T \frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right)\\ &= \mathrm{Tr}\left(\gamma^2[\mathbf{Z}\bar{\mathbf{F}}(x)]^H \frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right). \end{aligned} \tag{52}$$

In addition, taking a derivative at both sides of equation (6), $\mathbf{K}\bar{\mathbf{F}}(x) = \mathbf{H}^H\mathbf{L}^H\mathbf{W}$, we have

$$\frac{\partial}{\partial x}\left[\mathbf{K}\bar{\mathbf{F}}(x)\right] = -\mathbf{Z}\bar{\mathbf{F}}(x) + \mathbf{K}\frac{\partial \bar{\mathbf{F}}(x)}{\partial x} = 0,$$

which gives us $\mathbf{Z}\bar{\mathbf{F}}(x) = \mathbf{K}\frac{\partial \bar{\mathbf{F}}(x)}{\partial x}$. Then, plugging this result back into (52) yields

$$\frac{\partial J(x)}{\partial x} = \mathrm{Tr}\left(\gamma^2\left(\frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right)^H \mathbf{K}\frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right),$$

which implies that as long as $\mathbf{K} \succ 0$ (or $x < 1/\kappa$), $J(x)$ is monotonically increasing with x.

On the other hand, applying (12) to $J(x)$, it is rewritten by

$$\begin{aligned}J(x) &= \mathrm{Tr}\big(\mathbf{W}^H\mathbf{L}^H\mathbf{H}^H(\mathbf{Y} - x\mathbf{Z})^{-1}\mathbf{Z}(\mathbf{Y} - x\mathbf{Z})^{-1}\mathbf{H}\mathbf{L}\mathbf{W}\big)\\ &= \mathrm{Tr}\big(\mathbf{W}^H\mathbf{L}^H\mathbf{H}^H(\mathbf{I}_{N_T} - x\mathbf{Z}\mathbf{Y}^{-1})^{-1}\\ &\qquad \times \mathbf{Y}^{-1}\mathbf{Z}(\mathbf{Y} - x\mathbf{Z})^{-1}\mathbf{H}\mathbf{L}\mathbf{W}\big).\end{aligned}$$

Since the matrix $\mathbf{I}_{N_T} - \kappa^{-1}\mathbf{Z}\mathbf{Y}^{-1}$ is rank deficient, it is seen that as $x \to 1/\kappa$, $J(x)$ will go to positive or negative infinity. Reminding that $J(x)$ is an increasing function over $x < 1/\kappa$, we have $\lim_{x\to 1/\kappa} J(x) \to +\infty$, and the proof is concluded.

C. Proof of Lemma 4

In Theorem 2, we have proved that $\mathbf{F}(\tilde{\lambda}, \beta)$ in (22) minimizes the Lagrangian $\mathcal{L}(\mathbf{F}, \tilde{\lambda}, \beta)$ in (31) for given $\tilde{\lambda}$ and β. Now, let us suppose that $J_2(\tilde{\lambda}, \beta)$ is an increasing function with respect to $\tilde{\lambda}$, i.e., $J_2(x, \beta) > J_2(y, \beta)$ for two positive real values $x \in \mathbb{R}_+$ and $y \in \mathbb{R}_+$ with $x < y$. Then, one can show that $\mathbf{F}(x, \beta)$ further reduces the Lagrangian $\mathcal{L}(\mathbf{F}, y, \beta)$ than $\mathbf{F}(y, \beta)$. This contradicts our previous result that $\mathbf{F}(y, \beta)$ is a optimal solution of $\min_{\mathbf{F}} \mathcal{L}(\mathbf{F}, y, \beta)$, and therefore $J_2(\tilde{\lambda}, \beta)$ must be non-decreasing with respect to $\tilde{\lambda}$. Using the same argument, it is also verified from (32) that an increase of β will decrease $J_1(\tilde{\lambda}, \beta)$ for a given $\tilde{\lambda}$, and thus we obtain Lemma 4.

D. Proof of Lemma 5

It simply follows from (37) that a positive real value α exists if and only if

$$\begin{aligned}&\mathrm{Tr}(\mathbf{V}^H\mathbf{B}^{-1/2}\mathbf{G}^H\mathbf{G}\mathbf{B}^{-1/2}\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{D})\\ &\geq \frac{\bar{E}}{P_T}\mathrm{Tr}(\mathbf{V}^H\mathbf{B}^{-1}\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{D})\\ &\quad \Leftrightarrow \quad \mathrm{Tr}(\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{V}^H\mathbf{B}^{-1/2}\mathbf{Z}\mathbf{B}^{-1/2}\mathbf{V}\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{D}) \geq 0\\ &\quad \Leftrightarrow \quad \mathrm{Tr}(\mathbf{H}_\dagger\mathbf{Z}\mathbf{H}_\dagger^H\mathbf{D}) \geq 0,\end{aligned}$$

where $\mathbf{H}_\dagger = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{V}^H\mathbf{B}^{-\frac{1}{2}}$. In the meantime, since we have $\mathbf{H}^H\mathbf{H} = \mathbf{B}^{1/2}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H\mathbf{B}^{1/2}$ from (21), one can show that $\mathbf{H}_\dagger^H(\mathbf{H}^H\mathbf{H})\mathbf{H}_\dagger = \mathbf{I}_{N_S}$, which implies that $\mathbf{H}_\dagger$ is independent of $\tilde{\lambda}$. Thus, setting $\tilde{\lambda} = 0$, i.e., $\mathbf{H}_\dagger = \mathbf{\Lambda}_h^{-\frac{1}{2}}\mathbf{V}_h^H$ makes no loss of generality and we obtain Lemma 5.

E. Proof of Lemma 6

By setting $\mathbf{W} = \beta^{-2}\mathbf{\Lambda}$ and applying the facts that $\lambda = \tilde{\lambda}\beta^2$ and $\mu P_T = \lambda\bar{E} + \beta^2$, the optimal precoder $\mathbf{F}$ in (22) can be modified as

$$\begin{aligned}\mathbf{F} &= \left(\frac{1}{P_T}\mathbf{I}_{N_S} - \tilde{\lambda}\mathbf{Z}\right)^{-1/2}\mathbf{V}(\beta^{-2}\mathbf{I}_{N_S} - \mathbf{\Lambda}^{-1})_+^{1/2}\\ &= \left(\mu\mathbf{I}_{N_S} - \lambda\mathbf{G}^H\mathbf{G}\right)^{-1/2}\mathbf{V}(\mathbf{I}_{N_S} - (\beta^{-2}\mathbf{\Lambda})^{-1})_+^{1/2}\\ &= \mathbf{C}^{-1/2}\mathbf{V}(\mathbf{I}_{N_S} - \mathbf{\Psi}^{-1})_+^{1/2},\end{aligned} \tag{53}$$

where $\mathbf{C} = \mu\mathbf{I}_{N_S} - \lambda\mathbf{G}^H\mathbf{G}$ and $\boldsymbol{\Psi} = \beta^{-2}\boldsymbol{\Lambda}$. Then, it is easy to verify that $\boldsymbol{\Psi}$ can be obtained from the eigenvalue matrix of the following effective channel as

$$\mathbf{C}^{-\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{C}^{-\frac{1}{2}} = \begin{bmatrix}\mathbf{V} & \tilde{\mathbf{V}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Psi} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}^H\\ \tilde{\mathbf{V}}^H\end{bmatrix}.$$

Also, it is obvious that the optimal $\tilde{\lambda}$ and β satisfying (27) and (28) fulfil the conditions (8)-(10). Thus, the resulting solution in (53) is equivalent to one in [7, Th. 1] which maximizes the information rate, and thus we conclude the lemma.

REFERENCES

[1] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. Int. Symp. Inf Theory (ISIT)*, Jun. 2010, pp. 2363–2367.

[2] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2008, pp. 1612–1616.

[3] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4210–4221, Aug. 2013.

[4] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[5] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4798–4810, Sep. 2014.

[6] H. Son and B. Clerckx, "Joint beamforming design for multi-user wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6397–6409, Nov. 2014.

[7] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[8] W. Wang, L. Li, Q. Sun, and J. Jin, "Power allocation in multiuser MIMO systems for simultaneous wireless information and power transfer," in *Proc. IEEE VTC Fall*, Sep. 2013, pp. 1–5.

[9] X. Gui, Z. Zhu, and I. Lee, "Sum rate maximizing in a multi-user MIMO system with SWIPT," in *Proc. IEEE VTC Spring*, May 2015, pp. 1–5.

[10] C. Song, C. Ling, J. Park, and B. Clerckx, "MIMO broadcasting for simultaneous wireless information and power transfer: Weighted MMSE approaches," in *Proc. Globecom Workshops*, Dec. 2014, pp. 1151–1156.

[11] H. Lee, S.-R. Lee, K.-J. Lee, H.-B. Kong, and I. Lee, "Optimal beamforming designs for wireless information and power transfer in MISO interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4810–4821, Sep. 2015.

[12] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[13] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[14] C. Song, K.-J. Lee, and I. Lee, "MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2310–2319, Jul. 2010.

[15] C. Song, K.-J. Lee, and I. Lee, "MMSE-based MIMO cooperative relaying systems: Closed-form designs and outage behavior," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1390–1401, Sep. 2012.

[16] K.-J. Lee and I. Lee, "MMSE based block diagonalization for cognitive radio MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3139–3144, Oct. 2011.

[17] C. Xing, N. Wang, J. Ni, Z. Fei, and J. Kuang, "MIMO beamforming designs with partial CSI under energy harvesting constraints," *IEEE Trans. Signal Process. Lett.*, vol. 20, pp. 363–366, Apr. 2013.

[18] Y. Chen, Z. Wen, N. C. Beaulieu, S. Wang, and J. Sun, "Joint source-relay design in a MIMO two-hop power-splitting-based relaying network," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1746–1749, Oct. 2015.

[19] A. Ozcelikkale and T. M. Duman, "Linear precoder design for simultaneous information and energy transfer over two-user MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5836–5847, Oct. 2015.

[20] J. Joung and Y. H. Lee, "Regularized channel diagonalization for multiuser MIMO downlink using a modified MMSE criterion," *IEEE Trans. Sig. Process.*, vol. 55, no. 4, pp. 1573–1579, Apr. 2007.

[21] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

[22] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[23] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution: HSPA LTE for Mobile Broadband*. New York, NY, USA: Elsevier, 2008.

[24] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[25] H. Sung, S.-R. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3489–3499, Nov. 2009.

[26] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 953–961, Mar. 2008.

[27] S. S. Christensen, R. Agarwal, E. D. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[28] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[29] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.

Changick Song (S'09–M'13) received the B.S., M.S., and Ph.D. degrees from Korea University, Seoul, South Korea, in 2007, 2009, and 2012, respectively, all in electrical engineering. He was a Visiting Researcher with the University of Southern California, Los Angeles, CA, USA, in 2009, and with the Queen's University, Kingston, ON, Canada, in 2011. From September 2012 to April 2013, he was a Research Professor with Korea University and from May 2013 to August 2014, he was a Postdoctoral Research Associate with the Communications and Signal Processing (CSP) Group at Imperial College, London, U.K. In September 2014, he joined the faculty of Korea National University of Transportation, Chungju, Korea, where he is currently an Assistant Professor with the Department of Information and Communications Engineering. His research interest includes information theory and signal processing for wireless communications and security.

Jaehyun Park (S'03–M'10) received the B.S. and Ph.D. (M.S.- Ph.D. joint program) degrees in electrical engineering from the Korea Advanced Institute of Science and Technology in 2003 and 2010, respectively. From 2010 to 2013, he was a Senior Researcher with the Electronics and Telecommunications Research Institute, where he was involved in transceiver design and spectrum sensing for cognitive radio systems. From 2013 to 2014, he was a Post-Doctoral Research Associate with the Electrical and Electronic Engineering Department, Imperial College London. He is currently an Assistant Professor with the Electronic Engineering Department, Pukyong National University, South Korea. His research interests include signal processing for wireless communications, with a focus on detection and estimation for MIMO systems and cognitive radio systems, and joint information and energy transfer.

Bruno Clerckx (M'08) received the M.S. and Ph.D. degrees in applied science from the Université catholique de Louvain (Louvain-la-Neuve, Belgium) in 2000 and 2005, respectively. He is currently a Senior Lecturer (Associate Professor) with the Electrical and Electronic Engineering Department, Imperial College London, U.K. He held a visiting research position with Stanford University, Stanford, CA, USA, in 2003 and EURECOM, Sophia-Antipolis, France, in 2004. In 2006, he held a post-doctoral position with the Université catholique de Louvain. From 2006 to 2011, he was with Samsung Electronics, Suwon, South Korea, where he actively contributed to 3GPP LTE/LTE-A and IEEE 802.16m and acted as the rapporteur for the 3GPP Coordinated Multi-Point Study Item. He is the author or co-author of two books on MIMO wireless communications and networks and numerous research papers, standard contributions, and patents. He received the Best Student Paper Award at the IEEE SCVT 2002 and several Awards from Samsung in recognition of special achievements. He served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2011 to 2015. He is currently an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Inkyu Lee (S'92–M'95–SM'01–F'15) received the B.S. degree (Hons.) in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1992 and 1995, respectively. From 1995 to 2001, he was a Member of Technical Staff with Bell Laboratories, Lucent Technologies, where he studied high-speed wireless system designs. From 2001 to 2002, he was with Agere Systems (formerly Microelectronics Group of Lucent Technologies), Murray Hill, NJ, USA, as a Distinguished Member of Technical Staff. Since 2002, he has been with Korea University, Seoul, where he is currently a Professor with the School of Electrical Engineering. Also he is the head of the BK21 Center for humanware information technology. In 2009, he was a Visiting Professor with the University of Southern California, Los Angeles, CA, USA. He has authored over 130 papers in IEEE journals and has 30 U.S. patents granted or pending. His research interests include digital communications, signal processing, and coding techniques applied for next-generation wireless systems. He has served as an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2001 to 2011 and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011. In addition, he has been a Chief Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on 4G Wireless Systems) in 2006. He was a recipient of the IT Young Engineer Award at the IEEE/IEEK Joint Award in 2006 and of the Best Paper Award at APCC in 2006, IEEE VTC in 2009, and ISPACS in 2013. He was also a recipient of the Best Research Award from the Korea Information and Communications Society in 2011 and the Best Young Engineer Award from the National Academy of Engineering in Korea (NAEK) in 2013. He has been elected as a member of NAEK in 2015. He is an IEEE Fellow.

Kyoung-Jae Lee (S'06–M'11) received the B.S. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2005 and 2011, respectively. He interned with Beceem Communications, Inc., Santa Clara, CA, USA, in 2007. He was a Visiting Student with the University of Southern California, Los Angeles, CA, USA, in 2009. He was a Research Professor with Korea University in 2011. From 2011 to 2012, he was a Post-Doctoral Fellow with the Wireless Networking and Communications Group, University of Texas at Austin, Austin, TX, USA. Since 2012, he has been with Hanbat National University, Daejeon, South Korea, where he is currently an Associate Professor with the Department of Electronics and Control Engineering. His research interests include communication theory, signal processing, and information theory applied to the next-generation wireless communications. He was a recipient of the Best Paper Award at the IEEE VTC Fall in 2009, the IEEE ComSoc APB Outstanding Paper Award in 2013, and the IEEE ComSoc APB Outstanding Young Researcher Award in 2013.