MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer: Weighted MMSE Approaches

Changick Song, Cong Ling, Jaehyun Park, and Bruno Clerckx

Abstract—We consider simultaneous wireless information and power transfer (SWIPT) in MIMO Broadcast networks where one energy harvesting (EH) user and one information decoding (ID) user share the same time and frequency resource. In contrast to previous SWIPT systems based on the information rate, this paper addresses the problem in terms of the weighted minimum mean squared error (WMMSE) criterion. First, we formulate the WMMSE-SWIPT problem which minimizes the weighted sum-MSE of the message signal arrived at the ID user, while satisfying the requirement on the energy that can be harvested from the signal at the EH user. Then, we propose the optimal precoder structure of the problem and identify the best possible MSE-energy tradeoff region through the alternative update of the linear precoder at the transmitter with the linear receiver at the ID user. From the derived solution, several interesting observations are made compared to the conventional SWIPT designs.

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 significant energy source for devices that are capable of harvesting RF energy to power themselves. For this reason, *simultaneous wireless information and power transfer* (SWIPT) has appeared as promising technologies in conjunction with energy harvesting (EH) devices for providing energy constrained wireless networks with convenient and perpetual energy supplies [1] [2].

In SWIPT systems, efficient transmitter designs have been active research area over the last few years as in [3]–[7] and references therein. For example, various beamforming strategies have been proposed in multi-antenna networks such as interference [3], relaying [4], and broadcast channels (BC) [5] [6] to maximize the rate at the information decoding (ID) users and the harvested energy at the EH users simultaneously. Especially in [7], the authors completely characterized the rate-energy tradeoff region by suggesting the optimum precoder for multiple-input multiple-output (MIMO) two-user BC SWIPT where one EH user and one ID user share the same spectrum resource.

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

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

C. Ling is with the Dept. of Electrical and Electronic Eng., Imperial College, London, UK (e-mail: c.ling@imperial.ac.uk).

B. Clerckx is with the Dept. of Electrical and Electronic Eng., Imperial College, London, UK and the School of Electrical Eng., Korea University, Seoul, Korea (e-mail: b.clerckx@imperial.ac.uk).

While such information rate based designs may be useful for evaluating the system performance, it may not be realistically implementable due to high encoding and decoding complexity. As a low complexity alternative, linear receivers such as the minimum mean squared error (MMSE) have also been widely proposed in conventional communication networks where the information transfer was the only concern [8]–[10]. Surprisingly, however, there is no earlier contribution on the linear receiver issue in the context of SWIPT systems. As an initial step, our work develops a new design method of the transmitter and ID receiver in terms of the MMSE for MIMO BC SWIPT where an EH user and an ID user are located separately, and identify the best possible MSE-energy tradeoff region.

To get a better handle on the stream-wise error performance, we formulate a weighted MMSE (WMMSE) SWIPT problem which minimizes the weighted sum errors for the ID user while satisfying the predetermined EH constraint for the EH user. The weight matrix makes the problem generally non-convex; thus more challenging than the equal weight cases. To solve the problem, we first focus on the transmitter side for a given ID receiver and derive the solution as a semi-closed-form that can be obtained by simple bisection methods. It is then observed that when the ID user has a single antenna, the derived precoder reduces to the optimal beamformer that maximizes the signal-to-noise ratio (SNR) corresponding to the maximum information rate scheme in [7]. It is particularly interesting to observe that while the proposed solution is attainable by a simple bisection method, the previous design in [7] still needs to compute complicated ellipsoid methods to arrive at the same point.

Meanwhile, for a given the precoder, the optimal ID receiver is simply the Wiener filter [11]. Thus, we can easily optimize the transceiver through the alternative update process. Due to the non-convexity of the problem, our solution may not ensure the global optimality, but multiple initial points may take us closer to the optimal point. Thus, the resulting MSE-energy region will carry a significant meaning as an MMSE counterpart of the rate-energy region in [7]. Simulation results show that our solution exhibits improved bit error rate (BER) performance and convergence rate compared to the conventional rate-based designs. In addition, when we set the weight matrix to be equal to the eigenvalue matrix of the ID user channel, interestingly it is observed that the proposed precoder achieves almost every boundary points of the rate-energy region suggested in [7].

Notations: Normal letters represent scalar quantities, boldface letters indicate vectors and boldface uppercase letters
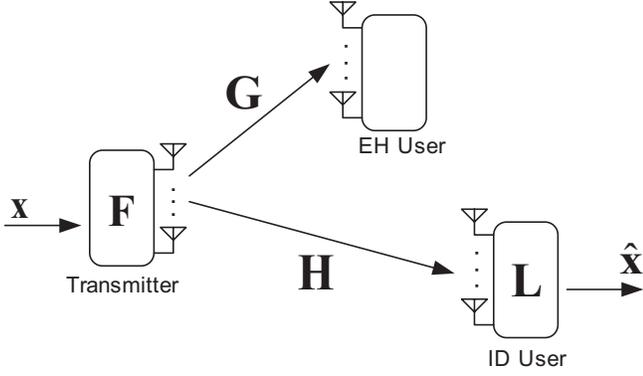
Fig. 1. A MIMO broadcast system for joint wireless information and energy transfer with linear transceiver

designate matrices. The superscripts $(\cdot)^T$, $(\cdot)^H$, and $\mathcal{E}[\cdot]$ stand for the transpose, the conjugate transpose, and the expectation operators, respectively. $\mathbf{I}_N$ is an $N \times N$ identity matrix. In addition, $\mathrm{Tr}\,(\mathbf{A})$, $|\mathbf{A}|$, and $\|\mathbf{A}\|_2$ indicate the trace, the determinant, and the matrix 2-norm of a matrix $\mathbf{A}$, respectively.

II. SYSTEM MODEL & PROBLEM FORMULATION

As shown in Figure 1, we consider a wireless multi-antenna broadcast system where one transmitter equipped with N_T antennas transmits radio signals to one EH user with N_{EH} antennas and one ID user with N_{ID} antennas at the same time and frequency. Each user is chosen from the "near-far" based scheduling at the transmitter, i.e., the closest user to the transmitter is scheduled as a EH user and the other one, further away, as an ID user. We assume that the two users are physically separated, but it can be easily extended to the co-located cases by applying a power splitting method in [7]. In this paper, we assume quasi-static flat fading channels so that the base-band channels from the transmitter to the EH and ID receivers are simply represented by complex matrices $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{ID}} \times N_T}$ and $\mathbf{G} \in \mathbb{C}^{N_{\mathrm{EH}} \times N_T}$, respectively. It is also assumed that the transmitter knows all channel state information (CSI) $\mathbf{H}$ and $\mathbf{G}$, and each user knows the corresponding CSI. Considering the spatial multiplexing scheme, we transmit N_S data-streams simultaneously using the input signal vector $\mathbf{x} \in \mathbb{C}^{N_S \times 1}$ with $\mathcal{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{N_S}$.

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

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

where $\mathbf{n} \in \mathbb{C}^{N_{\mathrm{ID}} \times 1}$ indicates the receiver noise with $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{\mathrm{ID}}})$, $\mathbf{L} \in \mathbb{C}^{N_S \times N_{\mathrm{ID}}}$ represents the linear receiver at the ID user, and $\mathbf{F} \in \mathbb{C}^{N_T \times N_S}$ is the linear precoder at the transmitter, which is subject to the power constraint $\mathrm{Tr}(\mathbf{F}\mathbf{F}^H) \leq P_T$. Then, the estimation error at the ID user can be defined as $\mathbf{e} = \gamma^{-1}\hat{\mathbf{x}} - \mathbf{x}$ and the resulting weighted sum-MSE is given by $\mathrm{MSE} = \mathrm{Tr}(\mathbf{W}\mathcal{E}[\mathbf{e}\mathbf{e}^H])$. Here, the scaling parameter γ^{-1} plays a role of automatic gain control [12] which helps simplifying the derivation and $\mathbf{W}$ stands for a real diagonal weight matrix. Note that without consideration

of the EH user, the MMSE optimum transmitter design which we denote by $\mathbf{F}_{\mathrm{ID}}$ is well investigated in literature [8].

In the meantime, the total harvested energy at the EH receiver, defined by Q, can be expressed as

$$Q = \delta\mathcal{E}[\|\mathbf{G}\mathbf{F}\mathbf{x}\|^2] = \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}) \tag{2}$$

where δ is a constant accounting for the harvesting efficiency and we assume $\delta = 1$ for simplicity 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 we denote the achievable sum-MSE with this beamformer by M_{EH}. Meanwhile, the sum-MSE and the energy attainable by the MMSE precoder $\mathbf{F}_{\mathrm{ID}}$ are denoted by $M_{\min}$ and E_{ID}, respectively.

Now, consider the case where both the EH and ID users co-exist. Then, the achievable MSE-energy region is

$$\mathcal{R}_{M,E} \triangleq \Big\{(M, E) : M \geq \mathrm{Tr}(\mathbf{W}\mathcal{E}[\mathbf{e}\mathbf{e}^H]),$$
$$E \leq \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}), \ \mathrm{Tr}(\mathbf{F}\mathbf{F}^H) \leq P_T\Big\},$$

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

$$\min_{\gamma, \mathbf{F}, \mathbf{L}} \quad \mathrm{Tr}(\mathbf{W}\mathcal{E}[\mathbf{e}\mathbf{e}^H])$$
$$s.t. \quad \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}) \geq \bar{E}, \ \mathrm{Tr}(\mathbf{F}\mathbf{F}^H) \leq P_T, \tag{3}$$

where $E_{\mathrm{ID}} \leq \bar{E} \leq E_{\max}$. This problem is convex over the transmitter (receiver) when the receiver (transmitter) and γ are given, and strictly quasi-convex with respect to γ when both $\mathbf{F}$ and $\mathbf{L}$ are given. However, it is generally non-convex on $\{\gamma, \mathbf{F}, \mathbf{L}\}$. Our goal of the paper is to find efficient solutions to problem (3) and identify the best possible MSE-energy region $\mathcal{R}_{M,E}$.

III. WEIGHTED MMSE TRANSCEIVER DESIGNS

In this section, we solve the problem (3) based on the Karush-Kuhn Tucker (KKT) conditions. First, we propose the optimum precoder structure in terms of the WMMSE, and then suggest alternative optimization process between the transmitter and the ID receiver. Several interesting observations will also be made when the ID user is equipped with a single antenna.

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

$$\mathcal{L}(\lambda, \mu, \gamma, \mathbf{F}, \mathbf{L})$$
$$= \mathrm{Tr}\big(\mathbf{W}(\gamma^{-1}\mathbf{L}\mathbf{H}\mathbf{F} - \mathbf{I}_{N_S})(\gamma^{-1}\mathbf{L}\mathbf{H}\mathbf{F} - \mathbf{I}_{N_S})^H\big)$$
$$+ \mathrm{Tr}(\gamma^{-2}\mathbf{W}\mathbf{L}\mathbf{L}^H) - \lambda\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{G}\mathbf{F}) - \bar{E}\big)$$
$$+ \mu\big(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\big). \tag{4}$$

Then, the following KKT conditions provide necessary conditions for optimality:

$$\mathbf{LHFF}^H\mathbf{H}^H + \mathbf{L} = \gamma\mathbf{F}^H\mathbf{H}^H \quad (5)$$

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

$$\mathrm{Tr}\left(\mathbf{WLHFF}^H\mathbf{H}^H\mathbf{L}^H + \mathbf{WLL}^H\right)$$
$$= \gamma\mathrm{Tr}\left(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W}\right) \quad (7)$$

$$\lambda \geq 0; \mu \geq 0; \mathrm{Tr}(\mathbf{F}^H\mathbf{G}^H\mathbf{GF}) \geq \bar{E}; \mathrm{Tr}(\mathbf{F}^H\mathbf{F}) \leq P_T \quad (8)$$

$$\bar{\mu}\left(\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) - P_T\right) = 0 \quad (9)$$

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

where $\bar{\lambda} \triangleq \lambda\gamma^2$, $\bar{\mu} \triangleq \mu\gamma^2$, and equations (5)-(7) are derived from the zero gradient conditions by using some rules of differentiation [13] and (9) and (10) come from the complement slackness conditions.

Now, let us define two matrices that will be used for our derivations as

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

$$\mathbf{Z} = \left(\mathbf{G}^H\mathbf{G} - \frac{\bar{E}}{P_T}\mathbf{I}_{N_T}\right).$$

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

Theorem 1: When $\mathbf{L}$ is given arbitrarily, the solutions of problem (3) are attained by

$$\hat{\mathbf{F}} = \hat{\gamma}\bar{\mathbf{F}}(\bar{\lambda}) = \hat{\gamma}\left(\mathbf{Y} - \bar{\lambda}\mathbf{Z}\right)^{-1}\mathbf{H}^H\mathbf{L}^H\mathbf{W}, \quad (11)$$

where $\hat{\gamma} = \sqrt{\frac{P_T}{Tr(\bar{\mathbf{F}}(\bar{\lambda})\bar{\mathbf{F}}(\bar{\lambda})^H)}}$. Defining $J(x) = Tr\left(\bar{\mathbf{F}}(x)^H\mathbf{Z}\bar{\mathbf{F}}(x)\right)$, if $J(0) \geq 0$, we have $\bar{\lambda} = 0$ and otherwise, a unique optimal $\bar{\lambda}$ satisfying $J(\bar{\lambda}) = 0$ can always be found over the range $0 < \bar{\lambda} < 1/\kappa$ by simple line search methods 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

$$\min_{\gamma,\mathbf{F}} \gamma^{-2}\left(\mathrm{Tr}(\mathbf{F}^H\mathbf{KF}) - 2\gamma\mathrm{Tr}(\Re(\mathbf{WLHF}))\right), \quad (12)$$

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 a non-positive real number with corresponding eigenvector $\mathbf{v} \in \mathbb{C}^{N_T \times 1}$. Then, it is easy to show that problem (12) 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 always be chosen such that $\mathbf{K} \succ 0$.

Assuming $\mathbf{K} \succ 0$, we now obtain from KKT condition (6)

that $\hat{\mathbf{F}} = \gamma\mathbf{K}^{-1}\mathbf{H}^H\mathbf{L}^H\mathbf{W}$. Also, from (6)-(8), it follows

$$\gamma\mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{W})$$
$$= \mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{WLHF} - \bar{\lambda}\mathbf{F}^H\mathbf{G}^H\mathbf{GF} + \bar{\mu}\mathbf{F}^H\mathbf{F})$$
$$= \mathrm{Tr}(\mathbf{F}^H\mathbf{H}^H\mathbf{L}^H\mathbf{WLHF}) - \bar{\lambda}\bar{E} + \bar{\mu}P_T$$
$$= \mathrm{Tr}(\mathbf{WLHFF}^H\mathbf{H}^H\mathbf{L}^H) + \mathrm{Tr}(\mathbf{WLL}^H),$$

where the second equality is due to (8) and the last equality follows from (7). Comparing the last two, we can show that $\bar{\mu} = \frac{\bar{\lambda}\bar{E}+\mathrm{Tr}(\mathbf{WLL}^H)}{P_T}$, which allows us to reduce $J(\bar{\lambda}, \bar{\mu})$ to a simple one-dimensional dual function $J(\bar{\lambda})$. Since $\bar{\mu} > 0$, we also have $\mathrm{Tr}(\mathbf{F}^H\mathbf{F}) = P_T$, i.e., $\hat{\gamma} = \left(P_T/\mathrm{Tr}(\bar{\mathbf{F}}(\bar{\lambda})\bar{\mathbf{F}}(\bar{\lambda})^H)\right)^{1/2}$ from (9).

Before we further proceed with our proof, let us show the following two useful lemmas. The proofs can be found in Appendix.

Lemma 1: The optimal $\bar{\lambda}$ takes a value over the range $0 \leq \bar{\lambda} < 1/\kappa$.

Lemma 2: $J(x)$ is a monotonic increasing function for $0 \leq x < 1/\kappa$ and $\lim_{x \to 1/\kappa} J(x) > 0$.

With the aid of Lemma 1 and 2, the optimal dual variable $\bar{\lambda}$ which we denote by $\hat{\lambda}$ can be obtained as follows: We first see from the KKT condition (10) that $\hat{\lambda}\left(\mathrm{Tr}(\gamma^2\bar{\mathbf{F}}(\hat{\lambda})^H\mathbf{G}^H\mathbf{G}\bar{\mathbf{F}}(\hat{\lambda})) - \bar{E}\right) = 0$ must be satisfied, which implies that $\hat{\lambda}J(\hat{\lambda}) = 0$. Therefore, if $J(0) \geq 0$, we attain $\hat{\lambda} = 0$, because $J(\bar{\lambda}) > 0$ for $\bar{\lambda} > 0$. In contrast, if $J(0) < 0$, one can find a unique $\hat{\lambda} \neq 0$ satisfying $J(\hat{\lambda}) = 0$ by simple line search methods over $0 < \bar{\lambda} < 1/\kappa$, since $J(1/\kappa) > 0$. It is worthwhile noting that $J(\hat{\lambda})$ must always be 0 or positive, because otherwise, the target energy $\bar{E}$ is not attainable at the EH receiver. Finally, as the solution satisfying the KKT conditions (6)-(10) is unique, the resulting precoder (11) is also sufficient for the optimality, and the proof is completed. ∎

Theorem 2: When γ and $\mathbf{F}$ are given, the solution of problem (3) is expressed as

$$\hat{\mathbf{L}} = \gamma\mathbf{F}^H\mathbf{H}^H(\mathbf{HFF}^H\mathbf{H}^H + \mathbf{I}_{N_{\mathrm{ID}}})^{-1} \quad (13)$$

Proof: The proof is trivial; thus omitted. ∎

Algorithm 1 WMMSE Optimization Method

---

for $j = 1 : N_G$ **do**
    Initialize $\bar{\mathbf{F}}_j$ and $\gamma_j = \left(P_T/\mathrm{Tr}(\bar{\mathbf{F}}_j\bar{\mathbf{F}}_j^H)\right)^{1/2}$
    repeat
        Compute $\mathbf{L}_j$ using (13) with given $\bar{\mathbf{F}}_j$ and γ_j
        Update $\mathbf{F}_j = \gamma_j\bar{\mathbf{F}}_j$ using (11) with given $\mathbf{L}_j$
    until $\mathbf{F}_j$ converge to the prescribed accuracy
end for
Select the best solution among $\{(\mathbf{F}_j, \mathbf{L}_j)|j = 1, \dots, N_G\}$, which shows the minimum sum-MSE

Based on results in Theorem 1 and 2, the 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 by each transmit and receive processing update. Therefore, the inner-iteration of Algorithm

1 guarantees the convergence at least to a local minimum. However, due to joint non-convexity of problem (3), numerous initial points which we denote by N_G are employed so that the resulting local minimum gets closer to the global minimum. This requires additional outer loop iterations.

Next, let us examine the optimal beamforming solutions with $N_{ID} = N_S = 1$ for the special cases where the ID receiver or both the ID and EH receivers have a single antenna. In this case, the transmit matrix $\mathbf{F}$ reduces to a column vector $\mathbf{f} \in \mathbb{C}^{N_T \times 1}$, the receive matrix $\mathbf{L}$ is equivalent to a scalar value $l \in \mathbb{C}$, and the weight matrix $\mathbf{W}$ can be ignored without loss of generality.

Corollary 1: In the case of MISO channel from the transmitter to the ID receiver, i.e., $\mathbf{H} \equiv \mathbf{h}^H$ with $\mathbf{h} \in \mathbb{C}^{N_T \times 1}$, the MMSE optimal beamformer, which is independent of the receiver l, is given by

$$\hat{\mathbf{f}} = \hat{\gamma}\bar{\mathbf{f}} = \hat{\gamma}\mathbf{A}^{-1}\mathbf{h}\Big(\frac{1}{1 + \mathbf{h}^H\mathbf{A}^{-1}\mathbf{h}}\Big), \tag{14}$$

where $\hat{\gamma} = \sqrt{P_T/\|\bar{\mathbf{f}}\|^2}$ and $\mathbf{A} \triangleq P_T^{-1}\mathbf{I}_{N_T} - \bar{\lambda}\mathbf{Z}$ where $\bar{\lambda}$ is obtained as follows: If $J(0) \geq 0$, we set $\bar{\lambda} = 0$ and otherwise, we find a unique $\bar{\lambda}$ satisfying $J(\bar{\lambda}) = 0$ by simple line search methods over $0 < \bar{\lambda} < 1/\zeta$ with $\zeta = \|\mathbf{Z}(\mathbf{h}\mathbf{h}^H + \frac{1}{P_T}\mathbf{I}_{N_T})^{-1}\|_2^2$.

Proof: Assuming $N_S = N_{ID} = 1$ and following similar approaches as in Theorem 1, the KKT conditions (5)-(10) lead to the solution

$$\hat{\mathbf{f}} = \gamma\Big(\mathbf{h}\mathbf{h}^H + \frac{1}{P_T}\mathbf{I}_{N_T} - \nu\mathbf{Z}\Big)^{-1}\mathbf{h}l^{-1}, \tag{15}$$

where $\nu \triangleq |l|^{-2}\bar{\lambda}$ and $\gamma = \sqrt{P_T/\|\bar{\mathbf{f}}\|^2}$ and ν is chosen to satisfy $\nu J(\nu) = 0$ over the range $0 \leq \nu < 1/\zeta$. Since a particular choice of l has no influence on the optimal value of ν and the resulting MSE of (15), i.e., $E = (1 + \mathbf{h}^H\mathbf{A}^{-1}\mathbf{h})^{-1}$ [11] is independent of l, setting $l = 1$ will not hurt the generality. Finally, by invoking some matrix inversion lemma, (15) is alternatively expressed as $\mathbf{f} = \gamma\mathbf{A}^{-1}\mathbf{h}(1 + \mathbf{h}^H\mathbf{A}^{-1}\mathbf{h})^{-1}$, and the proof is completed. ∎

In fact, when we deal with a single stream transmission, i.e., $N_S = 1$, the MMSE beamforming also maximizes the end-to-end SNR which leads to the maximum information rate [9] [11]. Therefore, our beamforming strategy in Corollary 1 is essentially equivalent to one in [7, Corollary 3.1] which was developed for the rate maximization. Interestingly, it is shown that while the previous approach [7] requires a complicated optimization tool such as the ellipsoid method to obtain the solution, our method enables us to achieve the maximum rate by a semi-closed-form solution that can be solved by a simple bisection method; thus, providing more insights into the system as well as complexity reduction.

IV. NUMERICAL RESULTS

In this section, we provide several interesting observations by comparing the numerical performance of the two MIMO BC SWIPT design criteria: WMMSE and the information rate [7]. Key system parameters are summarized in Table I. Throughout our simulation, the channels are generated from Rayleigh fading pathloss model, $\mathbf{H} = d_h^{-3/2}\widetilde{\mathbf{H}}$ and
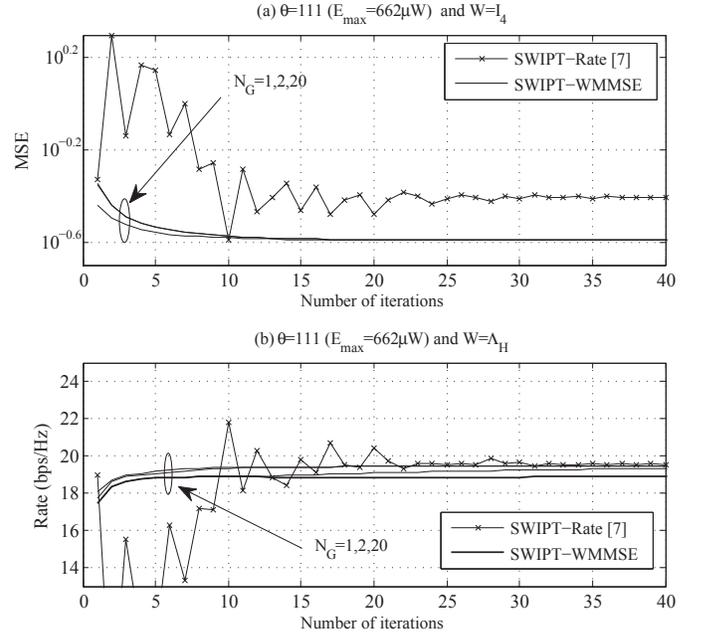

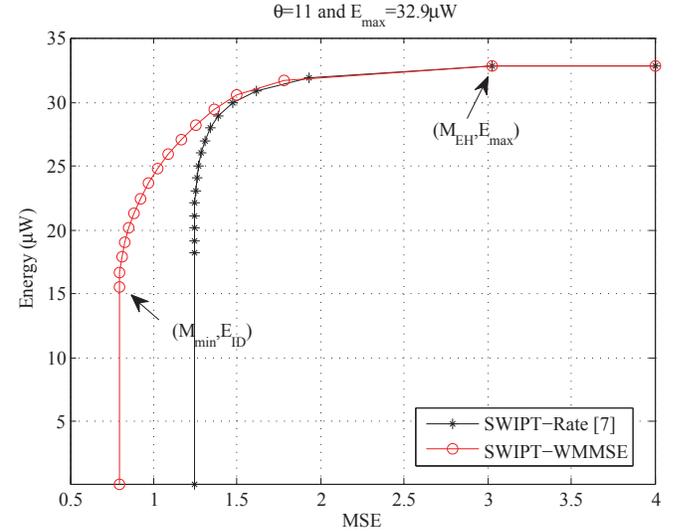Fig. 2. Sum-MSE and rate performance comparison of the two MIMO BC SWIPT methods with $\bar{E} = E_{max}/2$

Fig. 3. Achievable MSE-energy region of the two MIMO BC SWIPT methods with $P_T = 10$mW and $\mathbf{W} = \mathbf{I}_4$

$\mathbf{G} = d_g^{-3/2}\widetilde{\mathbf{G}}$ where each element of $\widetilde{\mathbf{H}}$ and $\widetilde{\mathbf{G}}$ is drawn from the i.i.d. CSCG normal distribution according to the MATLAB code with the seed number θ as in Table I. Also, we set $N_T = N_S = N_{ID} = N_{EH} = 4$ and $d_h = d_g = 10$(m) for the ease of the presentation. Let us assume that $P_T = 20$ dBm ($100mW$). Then, since the noise power is -100 dBm/Hz $\times$ 100MHz$=1\mu W$, the per-antenna SNR at the ID user becomes $20 - 30 - (-30) = 20$ dB, which is equivalent to the standard signal model with $P_T = 10^5$ in (1). Thus, 1 energy unit at the EH user in (2) amounts to $1\mu W$ in our simulation setting.

Figure 2-(a) and 2-(b) illustrate the convergence trends

TABLE I
SYSTEM PARAMETERS

Noise Power Spectral Density	-100 dBm/Hz
Energy Conversion Efficiency	50% ($\delta = 0.5$)
Pathloss Exponent	3
Signal Bandwidth	10 MHz
Distance: Tx - ID receiver	$d_h(m)$
Distance: Tx - EH receiver	$d_g(m)$
Channel Realization (MATLAB Code)	randn('state',θ); $\widetilde{\mathbf{H}} = (\text{randn}(N_{\text{ID}}, N_T) + j * \text{randn}(N_{\text{ID}}, N_T))/\text{sqrt}(2)$; $\widetilde{\mathbf{G}} = (\text{randn}(N_{\text{EH}}, N_T) + j * \text{randn}(N_{\text{EH}}, N_T))/\text{sqrt}(2)$;

of the proposed scheme with $P_T = 100$mW in terms of both the sum-MSE and the information rate with weight matrices $\mathbf{W} = \mathbf{I}_4$ and $\mathbf{W} = \mathbf{\Lambda}_H$, respectively, where $\mathbf{\Lambda}_H$ indicates the eigenvalue matrix of $\mathbf{H}^H\mathbf{H}$. In each iteration, the sum-MSE (M) and the rate (R) are computed from $M = \text{Tr}(\mathbf{W}(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F} + \mathbf{I}_{N_T})^{-1})$ and $R = \log|\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F} + \mathbf{I}_{N_T}|$ under the assumption that the ID user employs the MMSE receiver (13) and the maximum likelihood receiver, respectively. Due to the non-convexity of problem (3), the resulting performance depends on the number of initial points and iterations, especially when the weight matrix is non-identity. Nevertheless, it is seen that 20 random initial points and 10 number of iterations seem to be sufficient to achieve the converged performance for our scheme. It is somewhat natural that the proposed MMSE scheme obtains improved MSE performance over the *"SWIPT-Rate"* scheme [7]. However, it is particularly interesting to observe that our solution achieves the optimum rate as well when the weight factors are properly chosen, which implies that the rate maximizing precoder design can be regarded as a special case of our WMMSE design[1]. Extensive computer simulations confirm that $\mathbf{W} = \mathbf{\Lambda}_H$ is the best choice for maximizing the rate, but rigorous proof remains open.

Figures 3 and 4 illustrate the boundary points of the achievable MSE-energy and rate-energy regions, respectively. 20 random initial points are used for the proposed SWIPT-WMMSE scheme. It is shown that our solution achieves the outstanding MSE-energy region compared to the SWIPT-Rate scheme [7], which leads to the BER performance improvement as is shown later. In addition, Figure 4 shows that when $\mathbf{W} = \mathbf{\Lambda}_H$, the WMMSE precoder achieves almost every boundary points of the optimal rate-energy region; thus, confirming the previous observation made in Figure 2. We would like to note that although here we presented simulations results for a sample channel ($\theta = 11$), the trend does not change for all channel realizations.

Figure 5 compares the uncoded BER performance of the two MIMO BC SWIPT systems with 4QAM. We averaged over 10^5 random channel realizations and set $\bar{E} = E_{\max}/2$ in each realization so that the EH user harvests at least 50% of the maximum energy allowed for a given channel. The figure shows that the proposed MMSE scheme with $\mathbf{W} = \mathbf{I}_4$ obtains approximately 5 dB gain over the SWIPT-Rate scheme. This is
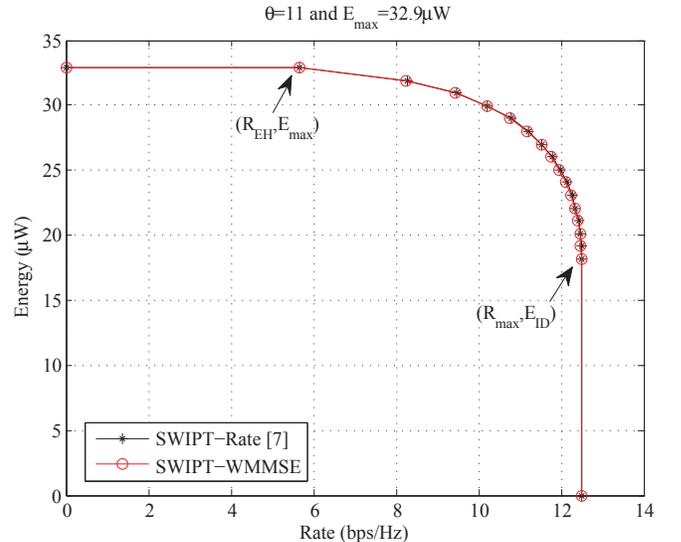
[1]Similar observation has been made in point-to-point MIMO systems without energy transfer [8]

Fig. 4. Achievable Rate-energy region of the two MIMO BC SWIPT methods with $P_T = 10$mW and $\mathbf{W} = \mathbf{\Lambda}_H$

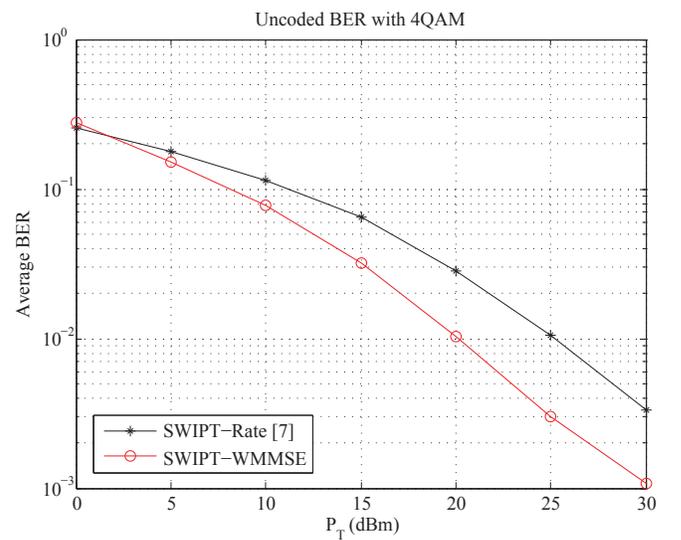

Fig. 5. BER performance comparison of the two MIMO BC SWIPT methods with $\mathbf{W} = \mathbf{I}_4$ and $\bar{E} = E_{\max}/2$

because the rate-based precoder allocates more resource to the stronger sub-channel, while the MMSE precoder neutralizes the channel gains across sub-channels. On the contrary, for the same reason, the rate-based design may exhibit better error performance when the transmitter adopts sufficiently strong channel codes. The proposed WMMSE precoder will softly bridge two extreme cases by adjusting the weight matrix $\mathbf{W}$.

V. CONCLUSION

In this paper, we investigated efficient linear transceiver designs for MIMO BC SWIPT systems which minimizes the weighted sum-MSE of the ID-user while satisfying the energy harvesting constraint of the EH-user. First, we identified the optimal precoder structure as a closed-form solution that can be obtained by simple bisection methods. Then, it was shown that the derived precoder achieves the maximum SNR at the ID user with reduced complexity when the ID user has a single antenna. Second, we suggested alternative updating process to obtain the joint WMMSE solution with the ID receiver. Although the problem is non-convex, it is observed from simulation results that our solution with multiple initial points approaches the best possible MSE-energy tradeoff region. It was also interesting to observe that the proposed WMMSE precoder with $\mathbf{W} = \mathbf{\Lambda}_H$ achieves almost every boundary points of the optimal rate-energy region.

VI. ACKNOWLEDGEMENT

APPENDIX

A. Proof of Lemma 1

Using $\bar{\mu} = \frac{\bar{\lambda}\bar{E} + \text{Tr}(\mathbf{W}\mathbf{L}\mathbf{L}^H)}{P_T}$, the feasibility condition $\mathbf{K} \succ 0$ is equivalently $\mathbf{Y} \succ \bar{\lambda}\mathbf{Z}$, which holds if and only if $\bar{\lambda} < 1/\kappa$ where κ indicates the maximum eigenvalue of a matrix $\mathbf{Z}\mathbf{Y}^{-1}$. Note that the maximum eigenvalue of a matrix is equivalent to its matrix two norm [14]. Therefore, combining with the condition $\lambda \geq 0$ in (8), we obtain Lemma 1.

B. Proof of Lemma 2

On the one hand, we first set the derivative of the function $J(x)$ with respect to x as

$$
\begin{aligned}
\frac{\partial J(x)}{\partial x} &= \text{Tr}\left(\left(\frac{\partial J(x)}{\partial \bar{\mathbf{F}}(x)}\right)^T \frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right) \\
&= \text{Tr}\left(\gamma^2 [\mathbf{Z}\bar{\mathbf{F}}(x)]^H \frac{\partial \bar{\mathbf{F}}(x)}{\partial x}\right).
\end{aligned} \quad (16)
$$

In addition, taking a derivative at both sides of the equation (6), i.e., $\mathbf{K}\bar{\mathbf{F}}(x) = \mathbf{H}^H\mathbf{L}^H\mathbf{W}$, 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 (16), it follows

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

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

On the other hand, applying the obtained precoder (11) to $J(x)$, it is rewritten by

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

This result implies that as $x \to 1/\kappa$, $J(x)$ will have an infinitely large number ($+\infty$ or $-\infty$) due to the inverse operation of a rank deficient matrix $\mathbf{I}_{N_T} - \kappa^{-1}\mathbf{Z}\mathbf{Y}^{-1}$. As shown previously, $J(x)$ is increasing function over $x < 1/\kappa$, and thus we have $\lim_{x \to 1/\kappa} J(x) \to +\infty$, and the proof is concluded.

REFERENCES

[1] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. IEEE International Symposium on Information Theory (ISIT'10)*, pp. 2363–2367, June 2010.

[2] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE International Symposium on Information Theory (ISIT'08)*, pp. 1612–1616, July 2008.

[3] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4210–4221, August 2013.

[4] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," vol. Available:http://arxiv.org/abs/1212.5406, 2012.

[5] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO Beamforming for Simultaneous Wireless Information and Power Transfer," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.

[6] H. Son and B. Clerckx, "Joint beamforming design for multi-user wireless information and power transfer," *to appear in IEEE Trans. Wireless Commun.*

[7] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 1989–2001, May 2013.

[8] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. on Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

[9] 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, pp. 2381–2401, September 2003.

[10] 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, pp. 2310–2319, July 2010.

[11] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, pp. 2700–2712, August 2005.

[12] J. Joung and Y. H. Lee, "Regularized channel diagonalization for multiuser MIMO downlink using a modified MMSE criterion," *IEEE Trans. Sig. Process.*, vol. 55, pp. 1573–1579, Apr. 2007.

[13] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentialtion: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, pp. 2740–2746, Jun. 2007.

[14] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Third Edition, The Johns Hopkins University Press, Baltimore and London, 1996.