Simultaneous Information and Power Transfer with Transmitters with Hardware Impairments

Simultaneous Information and Power Transfer with Transmitters with Hardware Impairments

Ayça Özçelikkale, Tomas McKelvey, Mats Viberg

***Abstract*—We investigate the performance of a communication system with simultaneous wireless information and power transfer capabilities under non-ideal transmitter hardware. We adopt an experimentally validated additive noise model in which the level of the noise at an antenna is proportional to the signal power at that antenna. We consider the linear precoder design problem and focus on the problem of minimizing the mean-square error under energy harvesting constraints. This set-up, in general, constitutes a non-convex formulation. For the single antenna information user case, we provide a tight convex relaxation, i.e. a convex formulation from which an optimal solution for the original problem can be constructed. For the general case, we propose a block coordinate descent technique to solve the resulting non-convex problem. Our numerical results illustrate the effect of hardware impairments on the system.**

I. Introduction

An attractive alternative to the traditional battery limited or grid dependent communication systems is the simultaneous wireless information and power transfer (SWIPT) framework. Here the two tasks, information and power transfer is performed simultaneously in a wireless medium. These two tasks typically require different optimal transmission strategies, hence novel transmission strategies have to be designed in order to be able to perform these tasks efficiently [1–3].

Wireless power transfer capabilities introduce great flexibility in terms of communication systems design, especially in scenarios where the transmitters have a relatively large number of antennas. On the other hand, for such systems to be utilized widely, the hardware used in each antenna component should be low cost, especially in massive multiple-input multiple-output (MIMO) systems [4]. The downside of inexpensive hardware is the fact that various impairments start to become prominent, including phase-noise, IQ-imbalance and amplifier non-linearities [4–6]. The impact of some of these distortions can be partially compensated by using compensation algorithms at the receiver or calibration methods at the transmitter, but nevertheless residual transmitter impairments still remains effective [4–6].

Although these residual transmitter impairments are known to significantly affect the performance of communication systems [4–8], up to now, this point is typically overlooked in the case of SWIPT system designs. Given that multiple-antenna systems are particularly attractive for power transfer applications, it is important that effect of hardware impairments is understood in these SWIPT systems. Here we address this issue under a linear precoding framework.

We consider the scenario where a transmitter aims to send information to an information receiver as reliably as possible, while also satisfying the energy harvesting (EH) constraints at the energy receiver. To model the hardware impairments, we adopt an additive noise model with a special covariance structure, which is validated with the experiments [5], [6] and supported by analytical arguments [4]. This set-up, in general, constitutes a non-convex formulation. For the single antenna information user case, we provide a tight convex relaxation. For the general case, we propose a block coordinate descent technique. Our results illustrate that when the channel signal-to-noise ratio (SNR) is high, significant gains can be obtained by the proposed hardware impairment aware designs.

The rest of the paper is organized as follows. In Sec. II, the system model is described. The precoder optimization problem is investigated in Sec. III. In Sec. IV, the performance of our designs is illustrated. The paper is concluded in Sec. V.

Notation: The complex conjugate transpose of a matrix A is denoted by A^{H}. The ith row jth column element is denoted by $[A]_{ij}$. The Frobenius norm is denoted by $||A||= (\mathrm{tr}[AA^{\mathrm{H}}])^{1/2}$. I_n denotes the identity matrix with $I_n \in \mathbb{C}^{n\times n}$. An optimal value of an optimization variable A is denoted by A^*.

II. System Model

A. Channel Model

In our narrow-band and stationary scenario, the multi-antenna transmitter transfers data to the information receiver (IR) as well as power to the energy harvesting receiver (ER) as

$$y_I = H_I x + w_I \tag{1}$$

$$y_E = H_E x + w_E \tag{2}$$

where y_I and y_E denote the signals received by IR and ER, respectively. Here $H_I \in \mathbb{C}^{n_r\times n_t}$ and $H_E \in \mathbb{C}^{n_e\times n_t}$ represent the channel gains from the transmitter to the IR and the ER where the number of antennas at the transmitter, the IR and the ER are denoted by n_t, n_r and n_e. Zero-mean complex proper Gaussian random variables $w_I \in \mathbb{C}^{n_r\times 1} \sim \mathcal{CN}(0, K_{w_I})$, $K_{w_I} = \mathbb{E}[w_I w_I^{\mathrm{H}}]$ and $w_E \in \mathbb{C}^{n_e\times 1}$, $w_E \sim \mathcal{CN}(0, K_{w_E})$, $K_{w_E} = \mathbb{E}[w_E w_E^{\mathrm{H}}]$ denote the noise at the IR's and ER's channel, respectively.

B. Precoding at the Transmitter with Non-Ideal Hardware

With an ideal transmitter, the channel input with linear precoding x can be expressed as $x = A_o s$ [9]. Here the zero mean complex proper Gaussian random vector $s \in \mathbb{C}^{n_s}$, $s \sim \mathcal{CN}(0, K_s)$, $K_s = I$ denotes the data and $A_o \in \mathbb{C}^{n_t\times n_s}$

A. Özçelikkale, T. McKelvey and M. Viberg are with Dep. of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden e-mails: {ayca.ozcelikkale, tomas.mckelvey, mats.viberg}@chalmers.se.
A. Özçelikkale acknowledges the support of EU Marie Skłodowska-Curie Fellowship.

denotes the linear precoder. We consider the hardware impairments at the transmitter as follows [4–6]

$$x = A_o s + v, \tag{3}$$

where $v \in \mathbb{C}^{n_t}$, $v \sim \mathcal{CN}(0, K_v)$ denotes the residual hardware impairments that remain effective after utilizing impairment compensation algorithms [4], [5]. The Gaussian assumption on v is supported by experiments (see for instance [5, Fig.7]) as well as by the central limit theorem and the fact that this term models the overall effect of various different hardware impairments [4–6]. The covariance of v is given as [4–6]

$$K_v = \alpha_v \operatorname{diag}(A_o A_o^{\mathrm{H}}), \tag{4}$$

where $\operatorname{diag}(M)$ denotes the diagonal matrix formed with $[M]_{11}$, ..., $[M]_{nn}$ as the elements on the main diagonal with $M \in \mathbb{C}^{n\times n}$,. Hence the level of noise at an antenna is proportional to the signal power at that antenna [4–6]. The associated model has been used to study the performance of various communication scenarios under hardware impairments [4–8].

The constant $\alpha_v \geq 0$ indicates the quality of the hardware. As α_v increases, the quality of the hardware decreases. Here v is modelled as statistically independent of the unknown signal s due to usage of compensation algorithms [4], [5]. We note that in contrast to w_I and w_E, statistics of v depend on the precoder A_o which will be optimized.

C. Signal Recovery at the IR

Upon receiving y_I, the information receiver forms an estimate of s. The mean-square error can be expressed as

$$\varepsilon(A_o, B) = \mathbb{E}[||s - By_I||^2], \tag{5}$$

where B represents the linear estimator adopted by the IR. Here the expectation is over the relevant signals and the noise, i.e. s, w and v. By standard arguments, the optimum B can be found as

$$B = A_o^{\mathrm{H}} H_I^{\mathrm{H}} \left(H_I A_o A_o^{\mathrm{H}} H_I^{\mathrm{H}} + K_{\bar{w}_I}\right)^{-1}. \tag{6}$$

where

$$K_{\bar{w}_I} = \alpha_v H_I \operatorname{diag}(A_o A_o^{\mathrm{H}}) H_I^{\mathrm{H}} + K_{w_I} \tag{7}$$

denotes the covariance of the effective noise at the receiver, i.e. $\bar{w}_I = H_I v + w_I$.

We note that due to the Gaussian distribution and the statistical independence assumptions on the relevant signals, $B\,y_I$ gives the minimum mean-square error (MMSE) estimation of s. The resulting MMSE can be expressed as

$$\begin{aligned}\varepsilon(A_o) &= n_s - \operatorname{tr}[A_o^{\mathrm{H}} H_I^{\mathrm{H}} (H_I A_o A_o^{\mathrm{H}} H_I^{\mathrm{H}} + K_{\bar{w}_I})^{-1} H_I A_o] \\ &= \operatorname{tr}[(I + A_o^{\mathrm{H}} H_I^{\mathrm{H}} (K_{\bar{w}_I})^{-1} H_I A_o)^{-1}], \end{aligned} \tag{8a}$$

where we have used Sherman-Morrison-Woodbury identity.

D. Energy Harvesting at the ER

The energy harvested at the ER can be expressed as [1]

$$\mathcal{J}(A_o) = \kappa \mathbb{E}[||y_E||_2^2] \tag{9}$$

where $\kappa \in [0, 1]$ is the loss factor that accounts for the possible loss in the energy conversion process. Hence

$$\mathcal{J}(A_o) = \kappa \operatorname{tr}[H_E \left(A_o A_o^{\mathrm{H}} + \alpha_v \operatorname{diag}(A_o A_o^{\mathrm{H}})\right) H_E{}^{\mathrm{H}}] + \kappa \operatorname{tr}[K_{w_E}].$$

An energy harvesting constraint in the form of $\mathcal{J}(A_o) \geq \gamma$ is imposed by the energy receiver. Since the EH constraints can be scaled accordingly, we set $\kappa = 1$ in the rest of the article without loss of generality.

III. Linear Precoder Design

Our aim is to find the precoder design that minimizes the MMSE while satisfying the EH constraint. We are interested in the following precoder design problem

$$\begin{aligned} (P1) \qquad & \min_{A_o} \quad \varepsilon(A_o) && \text{(10a)} \\ & \text{s.t.} \quad \mathcal{J}(A_o) \geq \gamma && \text{(10b)} \\ & \qquad \mathcal{P}(A_o) \leq P && \text{(10c)} \end{aligned}$$

where $\varepsilon(A_o)$ is as defined in (8). Here $\mathcal{P}(A_o)$ is the transmit power at the output of the transmitter given by

$$\begin{aligned} \mathcal{P}(A_o) &= \mathbb{E}[||A_o s + v||^2] && (11) \\ &= \operatorname{tr}[A_o A_o^{\mathrm{H}}] + \alpha_v \operatorname{tr}[\operatorname{diag}(A_o A_o^{\mathrm{H}})] && (12) \\ &= (1 + \alpha_v) \operatorname{tr}[A_o A_o^{\mathrm{H}}] && (13) \end{aligned}$$

As seen in (8a) when there are hardware impairments, the precoder not only affects the signal but also the statistics of the effective noise $\bar{w}_I$. Although the power used for the precoder increases the effective noise level, it is optimal to use all of the available power:

Lemma 3.1: *Let* $\sigma_{w_I} > 0$. *For an optimal solution of* A_o^* *of* (10), $\mathcal{P}(A_o^*) = P$.

Proof: Let us consider a fixed feasible A_o where $\mathcal{P}(A_o) = P_0 < P$. Let us form a new solution by scaling A_o, i.e. tA_o, $t \geq 1$. Now (8a) is a decreasing function of t under $\sigma_{w_I} > 0$. Since $\mathcal{J}(A_o)$ is also an increasing function of t, feasible solutions with a smaller objective function value can be obtained by increasing t. Hence for an optimum A_o^*, $\mathcal{P}(A_o^*) = P$. □

We note that under $\sigma_{w_I} = 0$, a solution with $\mathcal{P}(A_o) < P$ can be optimal. This can be seen, for instance, by considering the scenario with $\gamma = 0$, $n_t = n_r = n_s = 1$. Here the error is given by $(1 + |H_I|^2|A_o|^2/(\alpha_v|A_o|^2))^{-1}$ where any solution with $\mathcal{P}(A_o) > 0$ gives the same error value.

We note that $\varepsilon(A_o)$ is not a convex function of A_o. This is true even for the case with $\alpha_v = 0$. Although an optimal solution can be constructed for the $\alpha_v = 0$ case when there are no EH constraints (see for instance [9]), these results do not immediately generalize to (10). We further note that although it is possible to rewrite Problem P1 in terms of a new variable $R_{A_o} = A_o A_o^{\mathrm{H}} \succeq 0$, this formulation will have a rank-constraint $\operatorname{rank}(R_{A_o}) \leq n_s$ (so that an admissible optimal $A_o \in \mathbb{C}^{n_t \times n_s}$ can be found from an optimal $R_{A_o} \in \mathbb{C}^{n_t \times n_t}$), which corresponds to a non-convex constraint when $n_s < n_t$.

A. MISO IR Channel

We now consider the scenario with multiple-input single-output IR channel, hence $n_r = 1$, $H_I \in \mathbb{C}^{1\times n_t}$. The error can be expressed as

$$\begin{aligned} \varepsilon(A_o) &= \operatorname{tr}[(I_{n_s} + A_o^{\mathrm{H}} H_I^{\mathrm{H}} (K_{\bar{w}_I})^{-1} H_I A_o)^{-1}] && \text{(14a)} \\ &= \operatorname{tr}[(1 + (K_{\bar{w}_I})^{-1} H_I A_o A_o^{\mathrm{H}} H_I^{\mathrm{H}})^{-1}] + n_s - 1, && \text{(14b)} \end{aligned}$$

where $K_{\bar{w}_I} = \alpha_v H_I \operatorname{diag}(A_o A_o^{\mathrm{H}}) H_I^{\mathrm{H}} + \sigma_{w_I}^2 \in \mathbb{R}$ and (14b) follows from the equivalence of the non-zero eigenvalues of the products of matrices $M_1 M_2$ and $M_2 M_1$ [10]. We note that minimizing the error expression in (14b) is equivalent to maximizing the following signal-to-noise ratio expression

$$f(A_o) = \frac{H_I A_o A_o^{\mathrm{H}} H_I^{\mathrm{H}}}{\alpha_v H_I \operatorname{diag}(A_o A_o^{\mathrm{H}}) H_I^{\mathrm{H}} + \sigma_{w_I}^2} \tag{15}$$

Hence equivalent to (10), we consider the following problem

$$\max_{A_o} \quad f(A_o) \tag{16}$$

subject to (10c) and (10b). Although (16) constitutes a non-convex formulation, we will provide a tight convex relaxation to it. We first introduce a new variable $K_{A_o} = A_o A_o^{\mathrm{H}}$ and rewrite the objective function as follows

$$\bar{f}(K_{A_o}) = \frac{H_I K_{A_o} H_I^{\mathrm{H}}}{\alpha_v \operatorname{tr}[K_{A_o} \operatorname{diag}(H_I^{\mathrm{H}} H_I)] + \sigma_{w_I}^2} \tag{17}$$

where we have used

$$\begin{aligned} \operatorname{tr}[M \operatorname{diag}(A_o A_o^{\mathrm{H}}) M^{\mathrm{H}}] &= \operatorname{tr}[\operatorname{diag}(A_o A_o^{\mathrm{H}}) M^{\mathrm{H}} M] && \text{(18a)} \\ &= \operatorname{tr}[A_o A_o^{\mathrm{H}} \operatorname{diag}(M^{\mathrm{H}} M)] && \text{(18b)} \\ &= \operatorname{tr}[A_o^{\mathrm{H}} \operatorname{diag}(M^{\mathrm{H}} M) A_o], && \text{(18c)} \end{aligned}$$

for $\forall M \in \mathbb{C}^{n_s \times n}$ with $n \geq 1$ is an arbitrary integer and $H_I \operatorname{diag}(K_{A_o}) H_I^{\mathrm{H}} = \operatorname{tr}[H_I \operatorname{diag}(K_{A_o}) H_I^{\mathrm{H}}]$ for $H_I \in \mathbb{C}^{1 \times n_t}$. The energy harvested can be expressed as

$$\begin{aligned} \mathcal{J}_K(K_{A_o}) &= \operatorname{tr}[H_E \left(K_{A_o} + \alpha_v \operatorname{diag}(K_{A_o})\right) H_E^{\mathrm{H}}] && \text{(19)} \\ &= \operatorname{tr}[\left(H_E^{\mathrm{H}} H_E + \alpha_v \operatorname{diag}(H_E^{\mathrm{H}} H_E)\right) K_{A_o}]. && \text{(20)} \end{aligned}$$

The transmit power can be expressed as

$$\mathcal{P}_K(K_{A_o}) = (1 + \alpha_v) \operatorname{tr}[K_{A_o}].$$

Hence the optimization problem in (16) can be written as

$$\begin{aligned} \max_{K_{A_o} \succeq 0} \quad & \bar{f}(K_{A_o}) && \text{(21a)} \\ \text{s.t.} \quad & \mathcal{J}_K(K_{A_o}) \geq \gamma, && \text{(21b)} \\ & \mathcal{P}_K(K_{A_o}) \leq P, && \text{(21c)} \\ & \operatorname{rank}(K_{A_o}) \leq n_s. && \text{(21d)} \end{aligned}$$

We relax the rank constraint in (21d) and consider the following problem

$$\max_{K_{A_o} \succeq 0} \quad \bar{f}(K_{A_o}) \tag{22a}$$

subject to (21c) and (21b).

We will now show that this relaxation is tight, i.e. the solution of (22) provides a solution for (16). To this end, we first show (22) can be solved using convex optimization methods.

We note that the objective function of (22) is a linear-fractional function, hence it is not convex function of K_{A_o} [11]. Nevertheless, (22) can be written as a convex optimization problem using Charnes-Cooper transformation [12]. Let us define $\beta = (\operatorname{tr}[K_{A_o} H_I^{\mathrm{H}} H_I] + \sigma_{w,I}^2)^{-1}$, $G_{A_o} = \beta K_{A_o}$. Hence (22) can be equivalently written as

$$\begin{aligned} \max_{G_{A_o} \succeq 0, \beta \geq 0} \quad & \operatorname{tr}[H_I G_{A_o} H_I^{\mathrm{H}}] && \text{(23a)} \\ \text{s.t.} \quad & \mathcal{J}_K(G_{A_o}) \geq \beta\gamma, && \text{(23b)} \\ & \mathcal{P}_K(G_{A_o}) \leq \beta P, && \text{(23c)} \\ & H_I G_{A_o} H_I^{\mathrm{H}} + \beta \sigma_{w,I}^2 = 1. && \text{(23d)} \end{aligned}$$

This is a convex formulation. We obtain the following:

Lemma 3.2: *Let* (16) *be feasible. Then the optimum values for* (23) *and* (16) *are equal and can be attained. Using an optimal solution of* (23)*, an optimal rank* 1 *solution* K_{A_o} *for* (16) *can be constructed.*

Proof: We observe that for any fixed $\beta \geq 0$, (23) is a semi-definite programming (SDP) problem with two constraints. By [13, Thm 2.2], a solution G_{A_o} for (23) with rank 1 always exists. Hence the rank constraint inherent in (16) can be always satisfied. An optimal rank 1 solution for (16) from a solution of (23) can be constructed using [13, Algorithm RED].□

This result shows that by solving the convex optimization problem (23), a solution for (16) can be found. We further discuss the feasibility of (16) (equivalently feasibility of (10)) at the end of Sec. III-C.

We now discuss the effect of hardware impairments on the solution of (16). In particular, due to Lemma 3.1 and Lemma 3.2, a solution in the form of $A_o = \sqrt{\theta} u$, $\theta = P/(1+\alpha_v)$ with $||u||^2 = 1$, $u \in \mathbb{C}^{n_t \times 1}$ is optimal for (16). Hence maximizing the objective function of (16), i.e. (15), is equivalent to the maximization of the following expression

$$\frac{\theta |H_I u|^2}{\alpha_v \theta \sum_{i=1}^{n_t} |[H_I]_{i1}|^2 |[u]_i|^2 + \sigma_{w_I}^2}. \tag{24}$$

Let us consider the case $\gamma = 0$. Without the hardware impairments, i.e. $\alpha_v = 0$, $u_o = H_I^{\mathrm{H}} / ||H_I||$ which distributes the power proportional to strength of the channel coefficients is optimal (since u_o is the eigenvector associated with the largest eigenvalue of $H_I^{\mathrm{H}} H_I$). With $\alpha_v > 0$, this solution is not necessarily optimal. An example scenario is the following: Let $n_t = 2$, $H_I = [h_1 \; h_2] \in \mathbb{R}^{1 \times 2}$, $u \in \mathbb{R}^{n_t \times 1}$. Hence $u_a = [h_1 \; h_2]^{\mathrm{H}} / ||H_I||$ is optimal for $\alpha_v = 0$ for $\sigma_{w_I}^2 > 0$. On the other hand, for $\alpha_v > 0$, $\sigma_{w_I}^2 = 0$ the problem corresponds to minimizing an expression in the form of $\beta + 1/\beta$, $\beta \geq 0$, $\beta = (h_1 u_1)/(h_2 u_2)$ and $\beta = 1$, hence $u_b = [h_2 \; h_1]^{\mathrm{H}} / ||H_I||$ is optimal. Hence this suggests for $\sigma_{w_I}^2 \approx 0$, $h_1 \neq h_2$, the designs without and with the hardware impairments are different. Here the solution for $\alpha_v > 0$ needs to provide the optimum trade-off between two opposing forces: sending more power along the strong channel coefficient in order to increase signal power at the IR and sending less power along this channel coefficient in order to decrease the effective noise power at the IR. We note that under certain channel conditions, EH constraints may force the approach that assume ideal hardware to provide designs that are more close to the ones designed with the awareness of non-ideal hardware. Such a case occurs, for instance for the above scenario when $H_E / ||H_E|| = [h_2 \; h_1] / ||H_I||$ where the EH constraints will favor solutions close to u_b as γ increases.

B. Precoder Design with Fixed Receiver Filter at the IR

We now consider the general scenario $n_r \geq 1$. In order to propose a design for Problem P1 in this multiple-antenna IR scenario, we first consider the case where the IR uses a fixed estimation filter:

$$(P2) \qquad \min_{A_o} \quad \mathbb{E}_S[||s - By_I||^2] \tag{25a}$$

subject to (10c) and (10b). We note that here the filter B that does not depend on A_o.

Although Problem P2 also forms a non-convex formulation, we again derive a tight convex relaxation. For a given B, the mean-square error in (5) can be written as

$$\begin{aligned}\varepsilon(A_o, B) &= ||I_{n_s} - BH_I A_o||^2 + \mathrm{tr}[B(H_I K_v H_I^{\mathrm{H}} + K_{w_I})B^{\mathrm{H}}] \\ &= \mathrm{tr}[A_0^{\mathrm{H}} H_I^{\mathrm{H}} B^{\mathrm{H}} B H_I A_0] - 2\,\mathrm{Re}[\mathrm{tr}[BH_I A_0]] \\ &\quad + n_s + \alpha_v \,\mathrm{tr}[BH_I \,\mathrm{diag}(A_o A_o^{\mathrm{H}}) H_I^{\mathrm{H}} B^{\mathrm{H}}] \\ &\quad + \mathrm{tr}[BK_{w_I} B^{\mathrm{H}}],\end{aligned} \tag{26}$$

where $\mathrm{Re}[z]$ denotes the real part of $z \in \mathbb{C}$. We note that due to (18) the terms in (26) that include $\mathrm{diag}(A_o A_o^{\mathrm{H}})$ can also be expressed as convex quadratic functions of A_o. Hence the objective function of Problem P2, i.e. (26), is a convex quadratic function of A_o. Similarly, the constraints can be written as convex quadratic functions of A_o. Nevertheless, the EH constraint, (10b), does not form a convex constraint since it bounds a convex function from below. Hence the resulting problem formulation is not convex.

Using the variable $K_{A_o} = A_o A_o^{\mathrm{H}}$, the objective function can be rewritten as follows:

$$\begin{aligned}\varepsilon_K(A_o, K_{A_o}, B) = \mathrm{tr}[BH_I\,(K_{A_o} + \alpha_v \,\mathrm{diag}(K_{A_o}))\, H_I^{\mathrm{H}} B^{\mathrm{H}}] \\ - 2\,\mathrm{Re}[\mathrm{tr}[BH_I A_0]] + n_s + \mathrm{tr}[BK_{w_I} B^{\mathrm{H}}].\end{aligned}$$

Hence the optimization problem in (25) can be written as

$$\min_{A_o,\, K_{A_o}} \quad \varepsilon_K(A_o, K_{A_o}, B) \tag{27a}$$

$$\text{s.t.} \quad (21\text{b}), (21\text{c}), \tag{27b}$$

$$K_{A_o} = A_o A_o^H. \tag{27c}$$

Due to (27c), this formulation is not a convex optimization problem. We relax (27c) as $K_{A_o} \succeq A_o A_o^{\mathrm{H}}$. Now the relaxed problem can be expressed as follows

$$\min_{A_o,\, K_{A_o}} \quad \varepsilon_K(A_o, K_{A_o}, B) \tag{28a}$$

$$\text{s.t.} \quad (21\text{b}), (21\text{c}), \tag{28b}$$

$$K_{A_o} \succeq A_o A_o^H. \tag{28c}$$

This is a convex optimization problem. The following result shows that this relaxation is tight:

Lemma 3.3: *Let* (25) *be feasible. Then the optimum error values for the relaxed problem in* (28) *and the problem in* (25) *are equal and can be attained. Using an optimal solution of* (28)*, an optimal solution for* (25) *can be constructed.*

The proof is given in Sec. VI. In Sec. III-C, this result is used as an intermediate step to propose solutions for (10). We note that feasible regions for (25) and (10) are the same. We discuss the conditions for the feasibility of (25)/(10) in Sec. III-C.

Algorithm 1 Algorithm for Problem P1

Initialize:
- **if** ((30) is infeasible) **then**
  - Quit Algorithm 1. *// Problem P1 is infeasible.*
- **end if**
- Solve (30) and find A_o^0.
- Using A_o^0 and (6), find B^0. Let i=1.

repeat
- Using B^{i-1}, solve (28) for $(A_o^i, K_{A_o}^i)$.
- **if** Rank constraint is not satisfied **then**
  - Generate new A_o^i using [13, Algorithm RED].
- **end if**
- Using A_o^i and (6), find B^i.
- Using A_o^i,B^i and (26), find the error e^i.

until $(e^{i-1} - e^i \leq \epsilon)$ or $(i > n_{max})$) *// The stopping criterion is met.*

Output: A_o^i.

C. Joint Precoder and Receiver Filter Design

In Sec. III-B, we have considered the case where the estimator B is fixed. In general, optimum A_o depends on B. We will now consider the joint optimization of A_o and B, i.e. Problem P1 in (10). We rewrite it equivalently as follows

$$(P1) \quad \min_{A_o, B} \quad \varepsilon(A_o, B) \tag{29}$$

subject to (10c), (10b). Since the optimization over A_o for fixed B, (Problem P2) does not form a convex formulation, in general (29) is not a convex formulation in (A_o, B), either.

To find a design for Problem P1, we propose a block coordinate descent approach, which is summarized in Algorithm I. Here we alternate between fixing A_o and B. For fixed B, by Lemma 3.3, an optimal solution for A_o is found using (28). For fixed A_o, an optimal B is found using (6). By monotone convergence theorem this block coordinate descent technique is guaranteed to converge since the objective function is bounded from below and it decreases during the steps with both fixed A_o and fixed B. We note that due to non-convexity of the formulation, the proposed method provides possibly sub-optimal solutions for Problem P1. The algorithm is initialized using the solution of

$$\max_{K_{A_o} \succeq 0} \mathcal{J}_K(K_{A_o}), \tag{30}$$

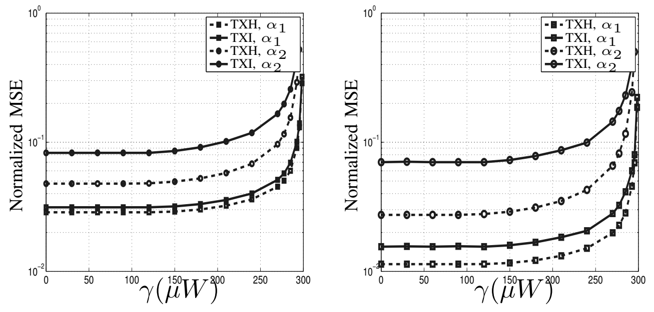
subject to (21c) and (21b). We note that $\mathcal{J}_K(K_{A_o})$ in (19) can be equivalently written as

$$\mathcal{J}_K(K_{A_o}) = \mathrm{tr}[A_o^{\mathrm{H}} \left(H_E^{\mathrm{H}} H_E + \alpha_v \,\mathrm{diag}(H_E^{\mathrm{H}} H_E)\right) A_o]$$

Whenever (30) is feasible, an optimal analytical solution of (30), $K_{A_o}^*$ in the form of beamforming with full power in the direction associated with the largest eigenvalue of $H_E^{\mathrm{H}} H_E + \alpha_v \,\mathrm{diag}(H_E^{\mathrm{H}} H_E)$ exists, see for instance [1]. Hence feasibility of (30) and equivalently (25)/(10) can be checked a priori by comparing $\mathcal{J}_K(K_{A_o}^*)$ and the EH constraint γ.

IV. Numerical Results

We now illustrate the performance of the hardware impairment aware designs. We consider $H_c = 10^{-3/2}\sqrt{n_t}\frac{\bar{H}_c}{||\bar{H}_c||}$ where $\bar{H}_c$ is given by the practical uniform linear array model $\bar{H}_c = \sum_{i=1}^{L_c} \kappa_{c,i} a_c(\theta_{c,i}) a_T^{\mathrm{T}}(\theta_{T,i})$; see, for instance, [14] for discussions on the validity and the applications of the model. The coefficient $10^{-3/2}$ is due to the path loss corresponding



(a) SNR= 35 dB.

(b) SNR= 40 dB

Fig. 1: Error versus energy harvesting requirements

to a path loss exponent of 3 and a distance of 10m between the transmitter and the receiver which introduces an average power loss of 30dB. Here $a_c(\theta) = [1\ e^{j2\pi d_c \cos(\theta)} \dots e^{j2\pi(n_c-1)d_c\cos(\theta)}]^T$ where c=T, IR, ER; $a_T(\theta_{T,i})$ is the array steering vector at the transmitter and $a_{IR}(\theta_{IR,i})/a_{ER}(\theta_{ER,i})$ is the array response vector at the IR/ER corresponding to the ith path in the IR/ER channel; $\kappa_{c,i}$ is the corresponding complex path amplitude [14]. Let $n_t = 3, n_r = 2,\ n_s = 2$, $L_{IR} = L_{ER} = 2$, $\kappa_1 = \kappa_2 = 1$, $d = 0.5$, $\theta_{IR,1} = \pi/6$, $\theta_{IR,2} = \pi/3$, $\theta_{ER,1} = \pi/2$, $\theta_{ER,2} = \pi/2$, $\theta_{T,1} = \pi/4$, $\theta_{T,2} = \pi/5$, $w_E = 0$, $K_{w_I} = \sigma^2_{w_I} I$, $\epsilon = 10^{-8} n_s$, $n_{max} = 1000$, SNR= $||H_{IR}||^2 P/(n_t \sigma^2_{w_I})$ (dB) where $P = 100$(mWs). The mean-square error (MSE) values are normalized by dividing with $n_s = \mathrm{tr}[K_s]$. TXH denotes the proposed hardware impairment aware designs obtained by Algorithm 1. TXI denotes the strategy that assumes ideal hardware. A practical quality measure for non-ideal hardware is the error vector magnitude (EVM) [4]. EVM and α_v are related as follows

$$\text{EVM} = \sqrt{\mathbb{E}[||v||^2]}/\sqrt{\mathbb{E}[||A_o s||^2]} = \sqrt{\alpha_v}. \tag{31}$$

We consider two values of α_v, $\alpha_1 = 0.05^2$, and $\alpha_2 = 0.15^2$ which correspond to an EVM of 0.05 and 0.15, respectively. We note that 3GPP LTE specifies EVM in the range $0.08 \leq$ EMV ≤ 0.175 [4].

The trade-off between the MSE and the EH requirements are presented in Fig. 1a and Fig. 1b, for SNR= 35dB and SNR= 40dB respectively. The plots illustrate that there is a significant average performance gap between the impairment aware solutions (TXH) and the solutions that assumes ideal hardware (TXI) when the SNR is high enough. In particular, for instance, with $\gamma \approx 100(\mu W)$, $\alpha_v = \alpha_1$ the aggregate MSE performance corresponds to bit error rates (BER) of 0.124 and 0.051 for Fig. 1a and 0.1 and 0.01 for Fig. 1b under 16QAM where we have related the MSE to the BER performance through [15, Ch.3]. The performance gap becomes smaller when the SNR decreases. This is consistent with the fact that for low SNR values, the performance is already affected by high levels of channel noise, and the TXI designs are also made with the awareness of this noise. Hence the relatively low levels of noise introduced by the hardware impairments do not affect the performance significantly. As the EH constraints become more demanding (such as $\gamma \gtrsim 280(\mu W)$ in Fig. 1b), the performance gap between TXH and TXI becomes smaller.

V. Conclusions

Linear precoder design for SWIPT systems is investigated under transmitter impairments. Our results illustrated that when the channel SNR is high, significant gains can be obtained by the proposed impairment-aware designs.

VI. Appendix

Using Schur complement, (28c) can be written as $Z_{A_o} = [I\ A_o^{\mathrm{H}}; A_o\ K_{A_o}] \succeq 0$. Considering (28) in terms of Z_{A_o} instead of A_o and K_{A_o} reveals that (28) is the SDP relaxation of the problem in (25) [13, 2.7]. By [13, Thm 2.2], (25) and (28) have the same optimal value if the relaxation is solvable and the number of constraints in (25) is equal to or smaller than $2n_s$. This last condition is satisfied $\forall n_s$, since (25) has 2 two constraints. Since the matrix associated with the power constraint, i.e. identity, is positive definite, the regularity condition in [13, 2.10] holds. Under feasibility of (25), this implies solvability of the SDP relaxation [13, Cor. 2.1]. This proves the first part of Lemma 3.3. An optimal solution for (25) is found as follows: Using [13, Algorithm RED] on an optimal solution $Z^*_{A_o}$ of (28), a rank-constrained optimal solution $\bar{Z}^*_{A_o}$ is obtained. Due to [13, Lemma 2.1], the lower left $n_t \times n_s$ matrix of $\bar{Z}^*_{A_o}$ gives an optimal A_o for (25).

References

[1] 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.

[2] K. Huang and E. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Trans. Signal Process.*, pp. 5972–5986, Dec. 2013.

[3] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a K-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 5781–5796, Oct. 2014.

[4] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, pp. 7112–7139, Nov 2014.

[5] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *2010 Int. ITG Workshop on Smart Antennas (WSA)*, pp. 189–196, Feb 2010.

[6] P. Zetterberg, "Experimental investigation of TDD reciprocity-based zero-forcing transmit precoding," *EURASIP Journal on Advances in Signal Process.*, p. 137541, 2011.

[7] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, pp. 91–94, Jan. 2013.

[8] X. Zhang, M. Matthaiou, E. Björnson, M. Coldrey, and M. Debbah, "On the MIMO capacity with residual transceiver hardware impairments," in *IEEE Inter. Conf. on Comm.*, pp. 5299–5305, June 2014.

[9] 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, 2001.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[12] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Research Logistics Quarterly*, vol. 9, no. 3-4, pp. 181–186, 1962.

[13] A. Beck, "Convexity properties associated with nonconvex quadratic matrix functions and applications to quadratic programming," *Journal of Opt. Theory and App.*, vol. 142, no. 1, pp. 1–29, 2009.

[14] H. Bölcskei, M. Borgmann, and A. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 427–439, Apr. 2003.

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