

## Research Article

# The Precoder Design with Covariance Feedback for Simultaneous Information and Energy Transmission Systems

Wen Zhou <sup>1</sup>, Dan Deng <sup>2</sup>, Junjuan Xia <sup>3</sup>, and Ziyun Shao <sup>4</sup>

<sup>1</sup>The College of Information Science and Technology, Nanjing Forestry University, Nanjing, China

<sup>2</sup>The School of Information Engineering, Guangzhou Panyu Polytechnic, Guangzhou, China

<sup>3</sup>School of Computer Science and Educational Software, Guangzhou University, Guangzhou, China

<sup>4</sup>The School of Mechanic and Electrical Engineering, Guangzhou University, Guangzhou, China

Correspondence should be addressed to Junjuan Xia; [xiajunjuan@gzhu.edu.cn](mailto:xiajunjuan@gzhu.edu.cn)

Received 28 January 2018; Revised 10 April 2018; Accepted 16 April 2018; Published 14 June 2018

Academic Editor: Nathalie Mitton

Copyright © 2018 Wen Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the optimal precoder design with the assumption that the transmitter only has channel covariance information, for the multi-input multi-output (MIMO) information and energy transmission system. The objective of the system design is to maximize the average system information rate, meanwhile meeting the minimum energy requirement of the energy receiver. Following this objective, we formulate the problem as a semidefinite programming (SDP) and further transform it into a dual problem. Two methods are proposed to solve this problem: the first method decomposes the transmission covariance as a product of precoders so that the constrained optimization becomes an unconstrained one, whereas the second method derives the structure of the optimal transmission covariance analytically. Both methods are proved to be convergent and their overheads and complexity are also analyzed. The achievable rate-energy (R-E) regions for the proposed methods are presented in the simulation. Under various system settings, the superiority of the proposed methods is shown by comparing with a few existing transmission schemes.

## 1. Introduction

The rapidly development of wireless sensor networks has created many applications such as environmental monitoring, telemedicine system, and intelligent house system [1, 2]. However, the main bottleneck of these applications is the energy limitation of mobile equipment [3–5]. Energy harvesting is a promising technique to deal with this problem, prolonging the lifetime of the network, and hence has attracted a great deal of attention [6–9].

Nowadays, the techniques for wireless power transfer (WPT) can be divided into three categories [7]: (1) near-field power transfer; (2) far-field directive power beaming; (3) far-field power transfer based on radio frequency (RF). Note that low-power RF signals are ambient and can be harvested by receivers from remote transmitters such as base stations and free WiFi hotspots. Further, since RF signals carrying energy can also be used for information transmission, a novel research direction, simultaneous wireless information and power transfer (WIPT), appears. It is clear that, compared

with the former two techniques, RF-based WPT is more suitable for simultaneous wireless information and power transfer (WIPT). In fact, most researches on WIPT are based on the third kind. Therefore, this article also applies RF signals in energy scavenging.

In general, the researches on WIPT were gone from simpleness to complication and from single-input single-output (SISO) systems to multiple-antenna systems [10–34]. For the SISO WIPT system, the problems on the trade-off between the rate and power, the circuit design, the transmission power allocation, and so on, were thoroughly investigated [10–23].

The technical papers [10, 11] are two pioneer researches on how to transmit energy and information jointly. In [10], the concept of the achievable rate-energy (R-E) region was first proposed and a few R-E expressions were derived under some discrete channels. In [11], the authors considered the WIPT over a frequency-selective channel with Gaussian noise, providing the solutions of power allocation to the discrete and continuous versions, respectively. The authors in [12] studied

the designs of both the information and energy receiver over time-varying fading channels. They presented several practical circuits with different multiplexing methods. In [15], the authors studied an amplify-and-forward (AF) relaying system, in which the energy of the received RF signal is harvested by the relay and then used for the information exchange between the source and destination. Two relay protocols were proposed to facilitate the WIPT of the relay. There were also a few researches that considered the multiuser access problem, wireless resource management, or other issues [18–20].

For the WIPT system with multiple antennas, the issues on the design of transmission covariance, the achievable R-E region, the system performance, and others, were addressed in [21–34]. For instance, in [21], a robust beamforming scheme was designed for the multiple-input single-output (MISO) WIPT system. This article makes an assumption that channel state information (CSI) obtained at the transmitter is not perfect and the design is based on the criterion in which the worst-case harvested energy for the energy receiver is maximized while guaranteeing that the minimum information rate for the information receiver. Utilizing a similar system model, in [24], aiming at maximizing the harvesting energy, the authors proposed an adaptive energy beamforming method according to the instantaneous CSI. In addition, the relationship between transmit power, transfer duration, and limited feedback amount was also presented. In [26], a unified study on MIMO WIPT systems was conducted, in which two practical designs were proposed in the case of the colocated receiver and their respective achievable R-E regions were also characterized. In [27], the joint two sources and relay beamforming design problem was studied for orthogonal space-time block code (OSTBC) based AF relay systems. In [29], a low-complexity technique of antenna switching for WIPT MIMO relay systems was proposed to minimize the system outage probability.

Usually, in the above researches, channel state information (CSI) at the transmitter is indispensable for the design of transmission scheme, power control, and others issues. Some of them assumed that the transmitter is able to know complete instantaneous CSI. With such an assumption, the authors in [11, 26] presented the achievable R-E regions for SISO systems and MIMO systems, respectively. Whereas there were also a few studies that considered that, in practice, it is quite difficult to obtain perfect CSI, due to Doppler shift, channel noise, or other factors, and hence assumed that only incomplete instantaneous CSI can be accessed by the transmitter. For instance, the authors in [21] designed a robust beamformer for the MIMO WIPT broadcasting system, with the assumption that the transmitter only has the incomplete CSI. In the multiantenna WIPT system, the authors in [24] analyzed the performance of energy beamforming to maximize the energy efficiency with the limited CSI feedback. Besides, a robust beamformer was designed by the use of the hybrid CSI, including the statistical CSI and instantaneous CSI [31].

It is clear that most existing researches use the instantaneous CSI at the transmitter to improve the system performance, including the cases of perfect CSI and imperfect CSI. It is suitable for the feedback of instantaneous CSI in slow fading channels. However, in the case of fast fading channel,

frequent CSI exchange will inevitably bring much burden on the communication system. Under such a scenario, the scheme of feeding back the instantaneous CSI is no longer appropriate. On the other front, in general, the channel statistical information does not change or changes very slowly, feeding back which can reduce the amount of CSI exchange significantly. It is no doubt that studying the precoder design with channel statistical information feedback makes sense. In fact, the optimal transmission covariance matrices with only statistical CSI feedback were thoroughly investigated in point-to-point MIMO systems or interference systems [32–37]. However, these cannot be applied to MIMO WIPT systems directly. To the best of our knowledge, the precoder with only statistical CSI feedback in WIPT systems has not been studied yet.

Therefore, this article focuses on the scenario in which the transmitter can only have the statistical CSI. Instead of using the instantaneous information rate as the optimization objective, the objective function becomes the expected system information rate. Consequently, the optimal precoder is designed under two constraints which include maximizing the expected information rate and meeting the minimum energy requirement. The contributions of this article are briefly summarized as follows:

- (i) Unlike most existing works, we design the optimal precoder by only use of statistical CSI, channel covariance feedback. Two precoding methods are proposed: one is a numerical method by employing the classical gradient-descent algorithm; the other one is an analytical method by use of the analytical expression of the optimal precoder. In the analytical method, we present the close-form expression of the optimal transmission covariance matrix via a proposition, as a function of the number of transmitter antennas, channel covariance matrices, and so on.
- (ii) We analyze the complexity and overheads of the proposed methods. A few classical existing works are also analyzed and compared, including time-switching method [26], isotropic transmission, and the precoder with hybrid CSI [31]. To the best of our knowledge, such work has not been done in existing literature. Furthermore, under various system settings, we compare the system performance of the proposed methods with that of these existing methods.
- (iii) Besides, the boundary of R-E region is characterized, analytically resending two important points of the boundary.

The rest of this paper is organized as follows. Section 2 describes the WIPT MIMO system model. Section 3 formulates the problem and proposes two precoding methods. Section 4 analyzes the complexity and overheads of the proposed methods. Simulation results are presented in Section 5, with concluding remarks in Section 6.

*Notations.* Vectors are denoted by boldface lowercase letters and matrices are denoted by boldface uppercase letters.  $E(\cdot)$  stands for the statistical expectation;  $\mathcal{C}^{m \times n}$  stands for the sets

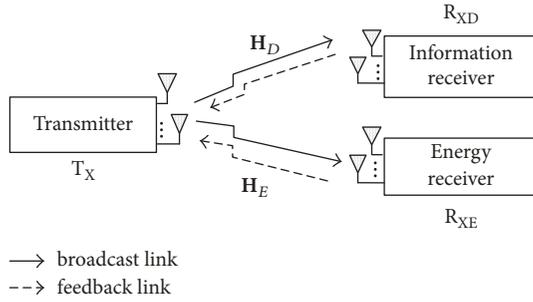


FIGURE 1: A three-node MIMO WIPT system model.

of  $m \times n$  complex matrices. For a vector  $\mathbf{x}$ ,  $\mathbf{x} \sim CN(\boldsymbol{\mu}, \mathbf{R})$  means that  $\mathbf{x}$  follows a complex Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{R}$ ; for a matrix  $\mathbf{X}$ , the notations  $\mathbf{X}^{1/2}$ ,  $\text{tr}(\mathbf{X})$ ,  $\mathbf{X}^H$ , and  $\mathbf{X}^*$  denote its square root, trace, Hermitian transpose, and conjugate, respectively; we write  $\mathbf{X} \geq \mathbf{0}$  to mean that  $\mathbf{X}$  is positive semidefinite; besides,  $\mathbf{I}_m$  is an  $m \times m$  identity matrix.

## 2. System Model

With reference to Figure 1, this paper considers a three-node MIMO WIPT system, which includes a transmitter  $T_X$ , an energy receiver  $R_{XE}$ , and an information receiver  $R_{XD}$ . The numbers of antennas for  $T_X$ ,  $R_{XE}$ , and  $R_{XD}$  are  $N_T$ ,  $N_E$ , and  $N_D$ , respectively. There are two kinds of links: one is the broadcast link for data or power transmission; the other is the feedback link for statistical CSI feedback [38–41]. The signal  $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$ , sent from the transmitter, has  $N_T$  independent data streams with its autocorrelation matrix  $E(\mathbf{x}\mathbf{x}^H) = \mathbf{I}_{N_T}$ . Before transmission, we multiply the signal  $\mathbf{x}$  by a precoder matrix  $\mathbf{W} \in \mathbb{C}^{N_T \times N_T}$ , and after passing through flat-fading channels, the received signals of the information receiver and energy receiver are given by

$$\mathbf{y}_D = \mathbf{H}_D \mathbf{W} \mathbf{x} + \mathbf{n}_D, \quad (1)$$

$$\mathbf{y}_E = \mathbf{H}_E \mathbf{W} \mathbf{x} + \mathbf{n}_E, \quad (2)$$

where  $\mathbf{H}_D \in \mathbb{C}^{N_D \times N_T}$  and  $\mathbf{H}_E \in \mathbb{C}^{N_E \times N_T}$  are flatting MIMO channels from the transmitter to the information receiver and energy receiver, respectively;  $\mathbf{n}_D \in \mathbb{C}^{N_D \times 1}$  and  $\mathbf{n}_E \in \mathbb{C}^{N_E \times 1}$  are *i.i.d* Gaussian noise vectors following  $CN(\mathbf{0}, \sigma_D^2 \mathbf{I}_{N_D})$  and  $CN(\mathbf{0}, \sigma_E^2 \mathbf{I}_{N_E})$ , respectively. In the followings, for the sake of convenience, we refer to  $\mathbf{H}_D$  and  $\mathbf{H}_E$  as the information channel and energy channel, respectively.

The channel matrices  $\mathbf{H}_D$  and  $\mathbf{H}_E$  can be further modeled as

$$\mathbf{H}_D = \boldsymbol{\theta}_{RD}^{1/2} \mathbf{H}_{\omega,D} \boldsymbol{\theta}_{TD}^{1/2}, \quad (3)$$

$$\mathbf{H}_E = \boldsymbol{\theta}_{RE}^{1/2} \mathbf{H}_{\omega,E} \boldsymbol{\theta}_{TE}^{1/2}, \quad (4)$$

in which  $\boldsymbol{\theta}_{TD}$  and  $\boldsymbol{\theta}_{RD}$  are channel covariance matrices for the information channel at the transmitter end and receiver end, respectively;  $\boldsymbol{\theta}_{TE}$  and  $\boldsymbol{\theta}_{RE}$  are channel covariance matrices

for the energy channel at the transmitter end and receiver end, respectively.  $\mathbf{H}_{\omega,D}$  and  $\mathbf{H}_{\omega,E}$  are random matrices having *i.i.d.* zero-mean and unit variance complex Gaussian random variables, representing the uncorrelated scattering.

After measuring channels for a period of time, the transmitter acquires channel statistical information, including  $\boldsymbol{\theta}_{RD}$ ,  $\boldsymbol{\theta}_{RE}$ ,  $\boldsymbol{\theta}_{TD}$ , and  $\boldsymbol{\theta}_{TE}$ . Using the information, the optimal precoder  $\mathbf{W}$  is designed to maximize the system information transmission rate, meanwhile satisfying the minimum energy requirement. Clearly, the scheme with statistical CSI feedback can reduce the overheads, compared to the scheme with instantaneous CSI feedback. However, it may bring the loss of information rate or energy.

Finally, for the sake of performance evaluation, we define the system signal-to-noise-ratio (SNR). Since the autocorrelation of the user signal  $\mathbf{x}$  is an identity matrix, the transmission power is expressed as

$$E(\|\mathbf{W}\mathbf{x}\|_F^2) = \text{tr}(\mathbf{W}\mathbf{W}^H) \leq P_T, \quad (5)$$

where  $P_T$  is the maximum transmission power. Obviously,  $\mathbf{W}\mathbf{W}^H$  is the transmission covariance matrix after precoding. With (5), the system SNR is defined as

$$\text{SNR} = \frac{P_T}{\sigma_D^2}, \quad (6)$$

where  $\sigma_D^2$  is the noise power of the information channel  $\mathbf{H}_D$ . Without loss of generality, we assume that  $\sigma_D^2 = 1$ .

## 3. The Proposed Precoding Methods with Covariance Feedback

In this section, we first formulate the precoding problem as a SDP and then transform it into a dual problem by introducing some auxiliary variables and constructing a dual function. Based on these, two methods are proposed to solve the dual problem. Their differences mainly lie on how to solve the dual function: one does it by the classical numerical searching, whereas the other does it by exploiting the structure of the covariance matrix.

**3.1. Problem Formulation.** First, with (1), the information transmission rate is given by

$$J = E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\}, \quad (7)$$

where the matrix  $\mathbf{Q} \in \mathbb{C}^{N_T \times N_T}$ , defined by  $\mathbf{W}\mathbf{W}^H$ , is the covariance matrix of the transmitted signal.

Then, with (2), the received energy per unit time can be written as

$$\varepsilon_{ng} = \eta E \left\{ (\mathbf{H}_E \mathbf{W} \mathbf{x})^H \mathbf{H}_E \mathbf{W} \mathbf{x} \right\} = \eta E \left[ \text{tr} \left( \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H \right) \right]. \quad (8)$$

With (4), (8) can be further expressed as [31]

$$\varepsilon_{ng} = \eta \text{tr}(\mathbf{Q} \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}), \quad (9)$$

where the notation  $\eta$  is the energy transfer efficiency. Without loss of generality, we set it to 1.

Finally, we try to find the optimal  $\mathbf{Q}$  that maximizes the expected information rate, meanwhile meeting the minimum energy requirement of the energy receiver,  $\varepsilon_{ng}$ . The problem, termed as P1, can be summarized as follows:

$$\begin{aligned} \text{P1 : } \max_{\mathbf{Q}} \quad & J = E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\} \\ \text{s.t. } \quad & \mathbf{Q} \succeq \mathbf{0} \\ & \text{tr}(\mathbf{Q}) \leq P_T \\ & \text{tr}(\mathbf{Q} \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) \geq \varepsilon_{ng}. \end{aligned} \quad (10)$$

Notice that unreasonable  $\varepsilon_{ng}$  may make the problem P1 not feasible, and hence we present the following lemma about the range of  $\varepsilon_{ng}$  (see Appendix A).

**Lemma 1.** *When  $\varepsilon_{ng} > P_T \lambda_{\max}(\boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE})$ , the problem P1 is not feasible, where  $\lambda_{\max}(\boldsymbol{\theta}_{TE})$  is the maximum eigenvalue of  $\boldsymbol{\theta}_{TE}$ ; when  $\varepsilon_{ng} < \text{tr}(\boldsymbol{\theta}_{RE})$ , the energy constraint of P1 is inactive, where  $\mathbf{Q}'_{OPT}$  is the solution of the following problem  $P'_1$ :*

$$\begin{aligned} P'_1 : \max_{\mathbf{Q}} \quad & E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\} \\ \text{s.t. } \quad & \mathbf{Q} \succeq \mathbf{0}; \text{tr}(\mathbf{Q}) \leq P_T. \end{aligned} \quad (11)$$

Note that  $P'_1$  is a transmission optimization problem for the transmitter only has channel covariance information. Similar problems were well investigated and solved in the past years [33–37]. For instance, in [33], the optimal transmission covariance matrices for two cases were presented, including channel covariance and mean feedback of rank one. In [34], the precoding methods for MIMO interference channels with covariance feedback were presented. In fact,  $P'_1$  can be solved by the proposed method in [33]. In addition, similar to [26], we also use the rate-energy (R-E) region to describe all achievable rate and energy pairs, which is defined as follows:

$$\begin{aligned} \Omega(P_T) = \{ (J, \varepsilon_{ng}) : & E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\} \\ & \geq J, \text{tr}(\mathbf{Q} \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) \geq \varepsilon_{ng}, \text{tr}(\mathbf{Q}) \leq P_T, \mathbf{Q} \succeq \mathbf{0} \} \end{aligned} \quad (12)$$

Clearly, given  $\varepsilon_{ng}$ , the rate solution of P1 is the vertical ordinate of some boundary point of the R-E region.

*Remark.* Recently, there are practical nonlinear models being proposed [42]. As pointed in [42], for low power, the harvested RF power increases with increasing input power; however, there are limitations on the maximum possible harvested energy. In fact, from Figure 2 of [42], we find that the nonlinearity mainly occurs in high input power region and the use of the linear energy harvesting (EH) model in low-power region is still suitable. If there is no light-of-sight (LOS) path or the distance is long enough between the transmitter and the receiver, the assumption that the received power is in low region is likely to be reasonable. Hence, assuming the received power is low, the conventional linear EH model is adopted in this article.

**3.2. The Proposed First Precoding Method.** In this section, a numerical algorithm is proposed to solve P1 and the optimal  $\mathbf{Q}$  can be obtained subsequently. Obviously, once acquiring the optimal  $\mathbf{Q}$ , the optimal precoder  $\mathbf{W}$  can be readily solved by taking the square root of  $\mathbf{Q}$ .

First, the Lagrangian of P1 is given by

$$\begin{aligned} L(\mathbf{Q}, \lambda, u) = & E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\} \\ & + \lambda \left[ \text{tr}(\mathbf{Q} \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) - \varepsilon_{ng} \right] \\ & - u \left[ \text{tr}(\mathbf{Q}) - P_T \right] \end{aligned} \quad (13)$$

where  $\lambda \geq 0$  and  $u \geq 0$  are two introduced auxiliary variables. Observe that the Lagrangian combines the objective function with the transmission power and energy constraints of P1. By doing so, the number of the constraints on the covariance matrix  $\mathbf{Q}$  is reduced. The aim of introducing the Lagrangian is to transform the original problem P1 into an easier problem. With (13), a dual function is defined as

$$g(\lambda, u) = \max_{\mathbf{Q} \succeq \mathbf{0}} L(\mathbf{Q}, \lambda, u). \quad (14)$$

Consequently, the variable  $\mathbf{Q}$  does not exist in this function.

With (14), the dual problem of P1 can be formulated as (P2)

$$\text{P2 : } J_{du} = \min_{\lambda, u \geq 0} g(\lambda, u) \quad (15)$$

Clearly, P2 only involves two introduced variables  $\lambda$  and  $u$  and it is much simpler than P1. It is easy to verify the Slater's condition is satisfied (see Appendix B). Therefore, there is no dual gap between P1 and P2. In other words, the two problems are equivalent.

**Next, we consider how to solve P2.** The problem-solving trait is composed of two steps: (1) given  $(\lambda, u)$ , obtain the optimal  $\mathbf{Q}$  and  $g(\lambda, u)$ ; (2) search over the two-dimension space of  $(\lambda, u)$  and find the minimum value of  $g(\lambda, u)$ . To do those, the objective function  $g(\lambda, u)$  needs further arrangement. With (13) and (14), we have [20]

$$\begin{aligned} g(\lambda, u) = \max_{\mathbf{Q} \succeq \mathbf{0}} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H \right] \right\} \\ - \text{tr} \{ \mathbf{A} \mathbf{Q} \} + T_{const}, \end{aligned} \quad (16)$$

where  $\mathbf{A} = u \mathbf{I}_{N_T} - \lambda \text{tr}(\boldsymbol{\theta}_{RE}) \boldsymbol{\theta}_{TE}$ ,  $T_{const} = u P_T - \lambda \varepsilon_{ng}$ . Note that the matrix  $\mathbf{A}$  must be semidefinite, otherwise  $g(\lambda, u)$  will go to infinity. This implies that  $u > \lambda \text{tr}(\boldsymbol{\theta}_{RE}) \lambda_{\max}(\boldsymbol{\theta}_{TE})$ , where  $\lambda_{\max}(\boldsymbol{\theta}_{TE})$  is the largest eigenvalue of  $\boldsymbol{\theta}_{TE}$ .

Since  $\mathbf{Q}$  is semidefinite and hence can be decomposed as  $\mathbf{Q} = \mathbf{W} \mathbf{W}^H$ , where  $\mathbf{W} \in \mathbb{C}^{N_T \times N_T}$  is the precoder. Replacing  $\mathbf{Q}$  with  $\mathbf{W}$ , the semidefinite constraint of  $\mathbf{Q}$  will disappear. Hence, (16) can be further written as

$$\begin{aligned} g(\lambda, u) = \max_{\mathbf{W}} L(\mathbf{W}, \lambda, u) \\ = \max_{\mathbf{W}} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{W} \mathbf{W}^H \mathbf{H}_D^H \right] \right\} \\ - \text{tr} \{ \mathbf{A} \mathbf{W} \mathbf{W}^H \} + T_{const}. \end{aligned} \quad (17)$$

Differentiating  $L(\mathbf{W}, \lambda, u)$  with respect to  $\mathbf{W}^*$ , we have

$$\begin{aligned} \mathbf{D}(\mathbf{W}) &\triangleq \frac{\partial L(\mathbf{W}, \lambda, u)}{\partial \mathbf{W}^*} \\ &= E \left\{ \left( \mathbf{I}_{N_T} + \mathbf{H}_D^H \mathbf{H}_D \mathbf{W} \mathbf{W}^H \right)^{-1} \mathbf{H}_D^H \mathbf{H}_D \mathbf{W} \right\} \\ &\quad - \mathbf{A} \mathbf{W}. \end{aligned} \quad (18)$$

Observe that solving  $\max_{\mathbf{W}} L(\mathbf{W}, \lambda, u)$  is an unconstrained optimization problem. Given  $(\lambda, u)$ , we propose the following subalgorithm to solve  $g(\lambda, u)$  and obtain  $\mathbf{W}$  and  $\mathbf{Q}$  afterwards.

Given  $(\lambda, u)$ , the proposed subalgorithm (Algorithm 1) will output the optimal  $\mathbf{W}$  and corresponding  $g(\lambda, u)$ . Based on Subalgorithm 1 (see Algorithm 1), a decent algorithm to solve the dual problem P2 is proposed. We term it as Main Algorithm 1 and summarize it in Algorithm 2.

It is worth noting that, in Main Algorithm 1 (see Algorithm 2),  $\Delta\lambda^{(i)}$  and  $\Delta u^{(i)}$  are derived as

$$\begin{aligned} \Delta\lambda^{(i)} &= \left. \frac{\partial g(\lambda, u)}{\partial \lambda} \right|_{\lambda=\lambda^{(i)}} \\ &= \text{tr} \left( \mathbf{W}^{(i)} \left( \mathbf{W}^{(i)} \right)^H \boldsymbol{\theta}_{TE} \right) \text{tr} \left( \boldsymbol{\theta}_{RE} \right) - \varepsilon_{ng} \end{aligned} \quad (19a)$$

$$\Delta u^{(i)} = \left. \frac{\partial g(\lambda, u)}{\partial u} \right|_{u=u^{(i)}} = -\text{tr} \left( \mathbf{W}^{(i)} \left( \mathbf{W}^{(i)} \right)^H \right) + P_T \quad (19b)$$

The flow chart of Main Algorithm 1 (see Algorithm 2) is depicted in Figure 2. With the proposed Main Algorithm 1 (see Algorithm 2), we can obtain the optimal  $\mathbf{W}$ , denoted as  $\mathbf{W}^{OPT}$ , and the minimum value of  $g(\lambda, u)$ . According to the duality principle, the optimal precoder of the original problem P1 is  $\mathbf{W}^{OPT}$ , and the maximum transmission rate of P1 is  $g(\lambda, u)$ 's minimum value.

*Remark 2.* In Subalgorithm 1 (see Algorithm 1), the initial  $\mathbf{W}$  can be set to an arbitrary matrix except a zero one. Since  $L(\mathbf{W}, \lambda, u)$  is not a convex function of  $\mathbf{W}$ , we had better run this algorithm several times and try different initial points to find a best value.

*Remark 3.* In the simulation, we compute the expected information rate in (7) by the following two steps. (1) With (3), generate some information channel samples, denoted as  $\mathbf{H}_D(n)$ , where  $n$  is the index of the samples. (2) Compute the following formula to approximate the expected rate in (7):

$$\sum_{n=0}^{N_{sam}-1} \log \det \left[ \mathbf{I}_{N_D} + \frac{\mathbf{H}_D(n) \mathbf{Q} \mathbf{H}_D^H(n)}{N_{sam}} \right], \quad (20)$$

where  $N_{sam}$  is the number of samples.

**3.3. The Proposed Second Precoding Method.** This subsection presents an indirect method for solving P2. To find the optimal covariance matrix, we show its structure by a proposition and then propose an iterative algorithm to solve an unknown diagonal matrix, which is similar to the so-called *power-allocation matrix*.

**First**, with (16), we have the following proposition.

**Proposition 4.** *The optimal covariance matrix  $\mathbf{Q}^{OPT}$  has the following form:*

$$\mathbf{Q}^{OPT} = \mathbf{A}^{-1/2} \mathbf{V}_B \boldsymbol{\Lambda} \mathbf{V}_B^H \mathbf{A}^{-1/2}, \quad (21)$$

where  $\mathbf{A} = u \mathbf{I}_{N_T} - \lambda \text{tr}(\boldsymbol{\theta}_{RE}) \boldsymbol{\theta}_{TE}$ , the columns of  $\mathbf{V}$  are taken from the right-singular vectors of  $\mathbf{B} \triangleq \boldsymbol{\theta}_{TD}^{1/2} \mathbf{A}^{-1/2}$ , and  $\boldsymbol{\Lambda} = \text{diag}(p_1 \ p_2 \ \dots \ p_{N_T})$  with  $p_1 \geq \dots \geq p_{N_T} \geq 0$ .

*Proof.* See Appendix C.  $\square$

From the above proposition, given  $(\lambda, u)$ , the unknown parameter to be determined is the diagonal matrix  $\boldsymbol{\Lambda}$ .

**Then**, we propose an iterative algorithm, like the so-called *power-allocation algorithm*, to solve  $\boldsymbol{\Lambda}$ . With a few manipulations (Appendix C),  $g(\lambda, u)$  in (14) can be expressed as

$$\begin{aligned} g(\lambda, u) &= \max_{\mathbf{Q} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{U}_{\theta_{RD}} \widetilde{\mathbf{H}}_D \mathbf{V}_B^H \widetilde{\mathbf{Q}} \mathbf{V}_B \widetilde{\mathbf{H}}_D^H \mathbf{U}_{\theta_{RD}}^H \right] \right\} \\ &\quad - \text{tr} \left\{ \widetilde{\mathbf{Q}} \right\} + T_{const}, \end{aligned} \quad (22)$$

where  $\widetilde{\mathbf{H}}_D = \boldsymbol{\Lambda}_{\theta_{RD}}^{1/2} \mathbf{U}_{\theta_{RD}}^H \mathbf{H}_{\omega, D} \mathbf{U}_B \boldsymbol{\Lambda}_B$ , the definitions of  $\mathbf{U}_B$  and  $\boldsymbol{\Lambda}_B$  are shown in (C.3), and  $T_{const} \triangleq u P_T - \lambda \varepsilon_{ng}$ .

With (21) and noticing  $\widetilde{\mathbf{Q}} = \mathbf{A}^{1/2} \mathbf{Q} \mathbf{A}^{1/2}$ ,  $g(\lambda, u)$  can be further expressed as

$$\begin{aligned} g(\lambda, u) &= \max_{\boldsymbol{\Lambda} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{U}_{\theta_{RD}} \widetilde{\mathbf{H}}_D \boldsymbol{\Lambda} \widetilde{\mathbf{H}}_D^H \mathbf{U}_{\theta_{RD}}^H \right] \right\} \\ &\quad - \text{tr} \{ \boldsymbol{\Lambda} \} + T_{const} \\ &= \max_{\boldsymbol{\Lambda} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \widetilde{\mathbf{H}}_D \boldsymbol{\Lambda} \widetilde{\mathbf{H}}_D^H \mathbf{U}_{\theta_{RD}}^H \mathbf{U}_{\theta_{RD}} \right] \right\} \\ &\quad - \text{tr} \{ \boldsymbol{\Lambda} \} + T_{const} \\ &= \max_{\boldsymbol{\Lambda} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \widetilde{\mathbf{H}}_D \boldsymbol{\Lambda} \widetilde{\mathbf{H}}_D^H \right] \right\} - \text{tr} \{ \boldsymbol{\Lambda} \} \\ &\quad + T_{const}. \end{aligned} \quad (23)$$

To solve  $\boldsymbol{\Lambda}$ , introducing auxiliary variables  $u'_i \geq 0$ ,  $i = 1, \dots, N_T$ , we construct the following Lagrangian function:

$$\begin{aligned} L_g(p_1, \dots, p_{N_T}, u'_1, \dots, u'_{N_T}) &= E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \widetilde{\mathbf{H}}_D \boldsymbol{\Lambda} \widetilde{\mathbf{H}}_D^H \right] \right\} - \sum_{i=1}^{N_T} p_i + \sum_{i=1}^{N_T} p_i u'_i \\ &\quad + T_{const} \\ &= E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \sum_{i=1}^{N_T} p_i \widetilde{\mathbf{h}}_{D,i} \widetilde{\mathbf{h}}_{D,i}^H \right] \right\} - \sum_{i=1}^{N_T} p_i \\ &\quad + \sum_{i=1}^{N_T} p_i u'_i + T_{const}, \end{aligned} \quad (24)$$

- (1) **Initialization:**  $\mathbf{W}^{(0)}$ , the iteration index  $i := 0$ , and the searching step  $t > 0$ .  
(2) **Repeat**  
At the  $i$ -th iteration, compute  $\mathbf{D}(\mathbf{W}^{(i)})$  according to (18);  
update  $\mathbf{W}$  according to  $\mathbf{W}^{(i+1)} = \mathbf{W}^{(i)} + t\mathbf{D}(\mathbf{W}^{(i)})$ ;  $i := i + 1$ .  
(3) **Until**  $|L(\mathbf{W}^{(i)}, \lambda, u) - L(\mathbf{W}^{(i-1)}, \lambda, u)| < \xi$ , where  $\xi$  is a pre-scribed threshold.

ALGORITHM 1: Subalgorithm 1: solve  $g(\lambda, u)$ .

- (1) **Initialization:** the iteration index  $i := 0$ ,  $\lambda^{(0)} \geq 0$ ,  $u^{(0)} \geq 0$ ,  
 $u^{(0)} > \lambda^{(0)} \text{tr}(\boldsymbol{\theta}_{RE}) \lambda_{\max}(\boldsymbol{\theta}_{TE})$ , the searching step  $s > 0$ .  
Compute  $g(\lambda^{(0)}, u^{(0)})$  and  $\mathbf{W}^{(0)}$  according to Sub-algorithm 1 (see Algorithm 1).  
(2) **Repeat**  
Update  $\lambda^{(i+1)}, u^{(i+1)}$  according to  
 $\lambda^{(i+1)} = \max(0, \lambda^{(i)} + s\Delta\lambda^{(i)})$ , (\*a)  
 $u^{(i+1)} = \max(0, u^{(i)} + s\Delta u^{(i)})$ , (\*b)  
where  $\Delta\lambda^{(i)}$  and  $\Delta u^{(i)}$  are the partial derivatives of  $g(\lambda, u)$   
with respect to  $\lambda$  and  $u$  at  $\lambda^{(i)}$  and  $u^{(i)}$ , respectively.  
Compute  $g(\lambda^{(i+1)}, u^{(i+1)})$  and  $\mathbf{W}^{(i+1)}$  via Sub-algorithm 1 (see Algorithm 1);  $i := i + 1$ .  
(3) **Until**  $g(\lambda, u)$  converges to the predetermined accuracy.

ALGORITHM 2: Main Algorithm 1: solve the dual problem P2.

in which the vector  $\tilde{\mathbf{h}}_{D,i}$  is the  $i$ -th column of  $\tilde{\mathbf{H}}_D$ . Consequently, the KKT (Karush-Kuhn-Tucker) optimality condition is given by

$$\frac{\partial L_g(p_1, \dots, p_{N_T}, u'_1, \dots, u'_{N_T})}{\partial p_l} = E \left\{ \text{tr} \left[ \left( \mathbf{I}_{N_D} + \sum_{i=1}^{N_T} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H \right)^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} - 1 + u'_l = 0. \quad (25)$$

$$u'_l p_l = 0, \quad u'_l \geq 0, \quad p_l \geq 0, \quad l = 1, \dots, N_T. \quad (26)$$

After a few manipulations, the optimality condition above can be expressed as

$$E \left\{ \text{tr} \left[ \left( \mathbf{I}_{N_D} + \sum_{i=1}^{N_T} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H \right)^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} = 1, \quad \text{for } p_l > 0; \quad (27)$$

$$E \left\{ \text{tr} \left[ \left( \mathbf{I}_{N_D} + \sum_{i=1}^{N_T} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H \right)^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} \leq 1, \quad \text{for } p_l = 0.$$

Notice that (27) is similar, but not identical to [32, Eq. (5)], because the problem of solving  $g(\lambda, u)$  and that in [32] are

somewhat different. Furthermore, as in Appendix D, (27) can be written in a simpler form,

$$E \left( \frac{\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) = 1, \quad \text{for } p_l > 0; \quad (28)$$

$$E \left( \frac{\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) \leq 1, \quad \text{for } p_l = 0,$$

where  $\mathbf{C}_l \triangleq \mathbf{I}_{N_D} + \sum_{i \neq l} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H$ . With (28), it is derived that (also see Appendix D)

$$p_l = 0, \quad \text{for } E \left( \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} \right) \leq 1; \quad (29)$$

$$p_l > 0, \quad p_l = 1 - E \left( \frac{1}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right),$$

otherwise.

Similar to the processing technique introduced in [32], we present the iterative algorithm (see Algorithm 3) to solve  $\{p_l\}$ , resulting from (29).

Consequently, given  $(\lambda, u)$ , the algorithm of solving  $g(\lambda, u)$  based on Proposition 4 is proposed and summarized in Algorithm 4.

**Finally**, with Subalgorithms 2.2 (see Algorithm 4), the second algorithm of solving P2 is proposed as Algorithm 5.

Note that, for (\*a) and (\*b) (see Algorithm 2) mentioned in Main Algorithm 2 (see Algorithm 5), the partial derivatives of  $g(\lambda, u)$  should be modified as

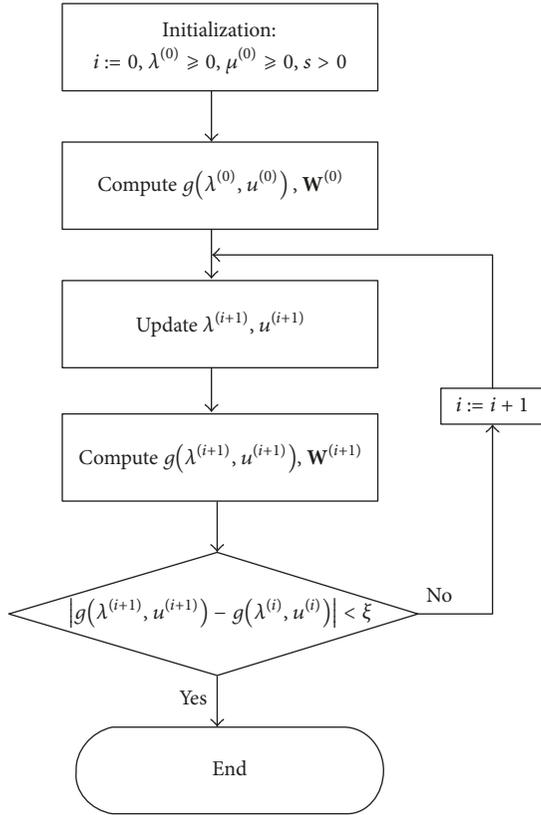


FIGURE 2: The flow chart of Main Algorithm 1 (see Algorithm 2).

$$\begin{aligned} \Delta \lambda^{(i)} &= \left. \frac{\partial g(\lambda, u)}{\partial \lambda} \right|_{\lambda=\lambda^{(i)}} = \text{tr}(\mathbf{Q}^{(i)} \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) - \varepsilon_{ng}, \\ \Delta u^{(i)} &= \left. \frac{\partial g(\lambda, u)}{\partial u} \right|_{u=u^{(i)}} = -\text{tr}(\mathbf{Q}^{(i)}) + P_T. \end{aligned} \quad (30)$$

Similar to the conclusion of the previous subsection, the proposed Main Algorithm 2 (see Algorithm 5) will output the optimal covariance  $\mathbf{Q}^{OPT}$  and the minimum  $g(\lambda, u)$ . According to the duality principle, the maximum transmission rate for P1 is the minimum  $g(\lambda, u)$  and the optimal precoder  $\mathbf{W}^{OPT}$  can be readily acquired by taking the square root of  $\mathbf{Q}^{OPT}$ .

*Remark 1.* In Subalgorithm 2.1 (see Algorithm 3), the initial power stream  $p_i^{(0)}$  should not be set to zero; otherwise it will be unchanged afterwards. To prove the convergence of this algorithm analytically is quite difficult; however, we show that it does converge rapidly by extensive simulations and similar conclusions have also been drawn in [32].

*Remark 2.* Since the dual function  $g(\lambda, u)$  is convex [43] and the two constraints  $\lambda \geq 0$  and  $u \geq 0$  are linear, the problem P2 is convex and its solution is unique. Searching from an arbitrary feasible initial point is able to reach its optimal value after a few iterations. On the other hand, it is easy to find that given  $\lambda$  and  $u$ ,  $g(\lambda, u) > -\infty$ , that is,  $g(\lambda, u)$  has a lower bound. Furthermore, the objective function  $g(\lambda, u)$  in either

Main Algorithm 1 or 2 (see Algorithm 2 or Algorithm 5) is nonincreasing because of the gradient-decent property of the proposed algorithm. Therefore, the proposed two algorithms are convergent.

*3.4. Discussion.* This subsection discusses two issues, including the difference between the proposed first method and second method and the difference between the proposed methods and the method in [32].

**First**, the two proposed methods solve the same optimization problem, i.e., solving  $g(\lambda, u)$  in two different ways. The first one is a *numerical method*, which transforms a constrained optimization problem into an unconstrained one, and hence can adopt the classical gradient-decent method to solve the problem, whereas the second one is an *analytical method*, which presents the structure of the optimal transmission covariance matrix in Proposition 4. Therefore, the second algorithm exhibits a deeper understanding of the original problem, by revealing the structure of the optimal precoder.

**Second**, we study the precoder design or covariance matrix in three-node WIPT systems, as compared with that in point-to-point systems [32]. Clearly, the structure of the optimal covariance matrix (21) is quite different from that in [32, Theorem 1]. The power-allocation algorithm is only a part of the structure of the covariance matrix. We draw on the trait of solving the power-allocation problem in [32] and the two problems do look similar to some degree. However, there are also a few differences between them. In addition, solving the power-allocation problem also yields new contributions.

- (i) The two problems are different in form: one is given by (23) and the other is shown in [32, Eq. (16)]. The derivation methods of solving the two problems are somewhat different. The authors in [32] derives the optimal matrix from the first principle [44, Section 15.2], a derivative version of the KKT conditions. However, the principle cannot be applied to the problem in this paper. Hence, we perform the derivations via the classical KKT conditions.
- (ii) When solving the power-allocation problem, we believe that the derivation from [32, Eq. (5)] to [32, Eqs. (10, 11)] is not such straightforward. The two formulas (27) and (29) in this paper are the counterparts of [32, Eq. (5)] and [32, Eqs. (10, 11)], respectively. We give the details on how (29) is derived from (27) in Appendix D. Using the techniques in Appendix D, one can derive [32, Eqs. (10, 11)] from [32, Eq. (5)].

## 4. Overhead and Complexity Issues

This section analyzes the overheads and computational complexity of the proposed precoding methods and a few other existing methods.

*4.1. The Overhead.* Table 1 lists the overheads of the proposed methods with other methods. We mainly consider the overhead that is required to compute the precoder.

Then, we make a brief review of the other three existing methods listed in Table 1. The isotropic transmission refers

- (1) **Initialization:** the iteration index  $k := 0$ ,  $p_l^{(0)} > 0$ , where  $l \in \Phi = \{1, \dots, N_T\}$ .
- (2) **Repeat**  
Update  $p_l^{(k+1)}$  according to  $p_l^{(k+1)} = 1 - E[1/(1 + p_l^{(k)} \tilde{\mathbf{h}}_{D,l}^H (\mathbf{C}_l^{-1})^{(k)} \tilde{\mathbf{h}}_{D,l})]$ ,  $l \in \Phi$ ;  $k := k + 1$ .
- (3) **Until** all  $p_l$ 's converge to the predetermined accuracy.
- (4) If there exists some  $l$  such that  $E[\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}] \leq 1$  but  $p_l$  does not converge to 0, find  $l_{\min} = \arg \min_{l \in \Phi} E[\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}]$ , set  $k := 0$  and  $p_{l_{\min}} = 0$ , update  $\Phi := \Phi - \{l_{\min}\}$ , and go to Step (2); otherwise end this procedure.

ALGORITHM 3: Subalgorithm 2.1: solve  $\{p_l\}$ .

TABLE 1: Overhead of computing the precoders.

Precoding method	Overhead
the proposed two methods	Tx obtains the channel covariance matrices $\{\boldsymbol{\theta}_{TE}, \boldsymbol{\theta}_{TD}, \boldsymbol{\theta}_{RE}, \boldsymbol{\theta}_{RD}\}$ and noise power $\sigma_D^2$ , by measurement over a period of time.
isotropic transmission	No overhead is needed.
the precoder with hybrid CSI feedback [31]	Tx obtains the energy channel covariance matrices $\{\boldsymbol{\theta}_{TE}, \boldsymbol{\theta}_{RE}\}$ and noise power $\sigma_D^2$ , by measurement over a period of time; Tx also requires the instant CSI of the information channel.
time switching method	It is the same with the proposed methods.

- (1) Obtain the matrix  $\boldsymbol{\Lambda} = \text{diag}(p_1 \ p_2 \ \dots \ p_{N_T})$ , according to Sub-algorithm 2.1 (see Algorithm 3).
- (2) Calculate the matrix  $\mathbf{Q}$  according to (21).
- (3) Calculate  $g(\lambda, u)$  by substituting  $\mathbf{Q}$  into (16).

ALGORITHM 4: Subalgorithm 2.2: solve  $g(\lambda, u)$ .

to the case in which the transmitter sends  $N_T$  data streams with equal power, i.e., the precoder matrix  $\mathbf{W} = \sqrt{P_T} \mathbf{I}_{N_T}$ . The time-switching scheme refers to the case in which each transmission period is divided into two time slots, one for data and the other one for energy. Let  $\beta$  with  $0 \leq \beta \leq 1$  be the percentage of transmitting time for the data time slot. The achievable R-E region of time-switching scheme is given by

$$\begin{aligned}
& \Omega'(P_T) \\
&= \bigcup_{0 \leq \beta \leq 1} \{(J, \varepsilon_{ng}) : \beta E \left\{ \log \det [\mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q}_1 \mathbf{H}_D^H] \right\} \\
&\geq J, \text{tr}(\mathbf{Q}_1) \leq P_T, (1 - \beta) \text{tr}(\mathbf{Q}_2 \boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) \\
&\geq \varepsilon_{ng}, \text{tr}(\mathbf{Q}_2) \leq P_T, \mathbf{Q}_1 \geq \mathbf{0}, \mathbf{Q}_2 \geq \mathbf{0}\}
\end{aligned} \tag{31}$$

in which  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  are the transmission covariance matrices for data and energy transmission phases, respectively. For the precoder with hybrid CSI feedback [31], only statistical CSI is available between the transmitter and energy receiver, while instant CSI is available between the transmitter and data receiver. What is slightly different from the precoder [31], herein we do not consider the channel estimation error brought by limited-length pilots and noises, since it is not the focus of this article.

From Table 1, observe that the isotropic transmission does not need the overhead; however, it has the worst performance, which will be shown in the simulations. Compared with the precoder with hybrid CSI feedback, the proposed methods only require the statistical channel information and do not need the instant CSI of the information channel, which reduces the overall overhead significantly in the long run. In addition, the proposed methods have the same overhead with the time-switching method.

**4.2. The Computational Complexity.** The number of floating-point flops is often used to measure of the complexity of an algorithm. Here, a flop is defined as one multiplication or division of two floating-point numbers, and one flop has the computational complexity  $O(1)$ . The addition and subtraction operations are neglected, since they are much quicker. Note that the matrix product  $\mathbf{X}_1 \mathbf{X}_2$  requires  $O(mnk)$  flops, where  $\mathbf{X}_1 \in \mathbb{C}^{m \times n}$ ,  $\mathbf{X}_2 \in \mathbb{C}^{n \times k}$ ; for an  $n \times n$  matrix, both its inverse and eigendecomposition operations require  $O(n^3)$  flops [45]; for an  $m \times n$  matrix, its singular value decomposition (SVD) requires  $O(mn^2)$  flops [45].

For the precoder with hybrid CSI feedback, the problem of solving it has the similar form with P3 [26] and hence can be solved by the proposed method in [26]. Note that the complexity for searching  $u$  and  $\lambda$  in this method is  $O(2^2) = O(1)$  [46]. For each iteration (searching  $u$  and  $\lambda$ ), the main computations lie in [26, Eq. (5) of Theorem 3.1] and have the complexity of  $O(N_T^2 N_D + N_D^2 N_T)$ . Therefore, if the precoder with hybrid CSI feedback is solved by the method [26], the complexity is  $O(N_T^2 N_D + N_D^2 N_T)$ .

For the time-switching method, it needs to compute two precoders. The first precoder is for energy transmission and the second is for information transmission. For the first one, it mainly involves the calculation of the eigenvectors of  $\boldsymbol{\theta}_{TE}$

- (1) **Initialization:** the iteration index  $i := 0$ ,  $\lambda^{(0)} \geq 0$ ,  $u^{(0)} \geq 0$ ,  $u^{(0)} > \lambda^{(0)} \text{tr}(\boldsymbol{\theta}_{RE})\lambda_{\max}(\boldsymbol{\theta}_{TE})$ , and  $s > 0$ . Compute  $g(\lambda^{(0)}, u^{(0)})$  and  $\mathbf{Q}^{(0)}$  according to Sub-algorithm 2.2 (see Algorithm 4).
- (2) **Repeat**  
Update  $(\lambda^{(i+1)}, u^{(i+1)})$  according to (\*a) and (\*b) in Algorithm 2, respectively; compute  $g(\lambda^{(i+1)}, u^{(i+1)})$  and  $\mathbf{Q}^{(i+1)}$ , according to Sub-algorithm 2.2 (see Algorithm 4); increase  $i := i + 1$ .
- (3) **Until**  $g(\lambda, u)$  converges to the prescribed accuracy.

ALGORITHM 5: Mainalgorithm 2: solve the dual problem P2.

TABLE 2: Computational complexity to compute the precoder.

Precoding method	Complexity
The proposed first method	$T_{11}T_{12}O(N_T^3 + N_D^2N_D)$ .
The proposed second method	$T_{21} \{T_{23} [N_T T_{22} - T_{22}(T_{22} - 1)/2] \times O(N_D^3) + O(N_T^3)\}$
isotropic transmission	No need to compute
the precoder with hybrid CSI feedback [31]	$O(N_T^2N_D + N_D^2N_T)$
time switching method	$O(N_T^3 + N_D^3)$

and hence has the complexity of  $O(N_T^3)$ . For the second one, it can be solved by the proposed method in [27] and this method has the complexity of  $O(N_T^3 + N_D^3)$ . For the above, it is concluded that the time-switching method has the complexity of  $O(N_T^3 + N_D^3)$ .

For the proposed first method, suppose that, for Main Algorithm 1 and Subalgorithm 1 (see Algorithms 2 and 1), the numbers of iterations required to reach the given accuracy are  $T_{11}$  and  $T_{12}$ , respectively. Observe that the main computation of this method lies in (18) and has the complexity of  $O(N_T^3 + N_T^2N_D)$ . Therefore, the complexity of the proposed first method is  $T_{11}T_{12}O(N_T^3 + N_T^2N_D)$ .

For the proposed second method, suppose that the number of iterations required to reach the given accuracy for Main Algorithm 2 (see Algorithm 5) is  $T_{21}$ . Observe that there are two loops in Subalgorithm 2.1 (see Algorithm 3), termed as the outer loop and inner loop. Suppose that the number of iterations for the outer loop is  $T_{22}$  and clearly  $1 \leq T_{22} \leq N_T$ . For the inner loop, assume that the average number of iterations required to calculate the power allocated to each eigenvector is  $T_{23}$ . Observe that the main computations lie in Step (2) of Subalgorithm 2.1 (see Algorithm 3) and (21). With the above, it is easy to find that the overall complexity of this method is  $T_{21}\{T_{23}[N_T T_{22} - T_{22}(T_{22} - 1)/2]O(N_D^3) + O(N_T^3)\}$ . Note that, we find  $T_{22} = 1 \sim 2$  and  $T_{23} = 5 \sim 30$  usually in the simulation. The complexity of the proposed methods and a few existing methods is listed in Table 2.

## 5. Simulation Results

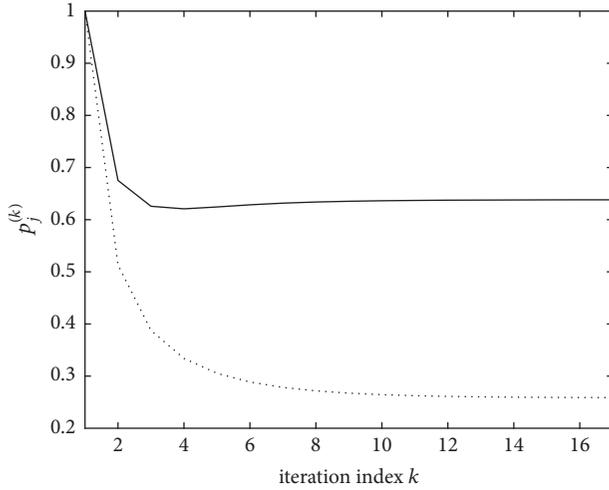
Computer simulation is employed to evaluate the performance of the proposed precoders of solving the optimal precoder and compare a few other existing methods. A three-node MIMO WIPT system is considered, where all nodes have the same number of antennas. We adopt the exponential

correlation model as the channel correlation matrix, with its  $(i, j)$ -th entry being  $\rho^{|i-j|}$ , in which the constant  $\rho$  is the correlation coefficient. Unless specified otherwise, for  $\boldsymbol{\theta}_{TD}$ ,  $\boldsymbol{\theta}_{TE}$ ,  $\boldsymbol{\theta}_{RD}$ , and  $\boldsymbol{\theta}_{RE}$ , their correlation coefficients are set to 0.1, 0.7, 0.2, and 0.3, respectively.

*5.1. The Convergence of the Proposed Methods.* Figure 3 shows the convergence of the proposed Subalgorithm 2.1 (see Algorithm 3). Two cases with  $N_T = N_D = N_E = 2$  and  $N_T = N_D = N_E = 3$  are considered. For both cases, the system SNR is 5 dB, the energy threshold  $\varepsilon_{ng}$  is set to 10, and the initial power matrix  $\boldsymbol{\Lambda}^{(0)} = \mathbf{I}$ . As in Figure 3(a), observe that, after 10 iterations, the power matrix converges to about  $\text{diag}\{0.64, 0.26\}$ . Similar behavior can be found in Figure 3(b), with the corresponding power matrix converging to  $\text{diag}\{0.92, 0.49, 0.33\}$ . These results have well demonstrated the convergence of Subalgorithm 2.2 (see Algorithm 4).

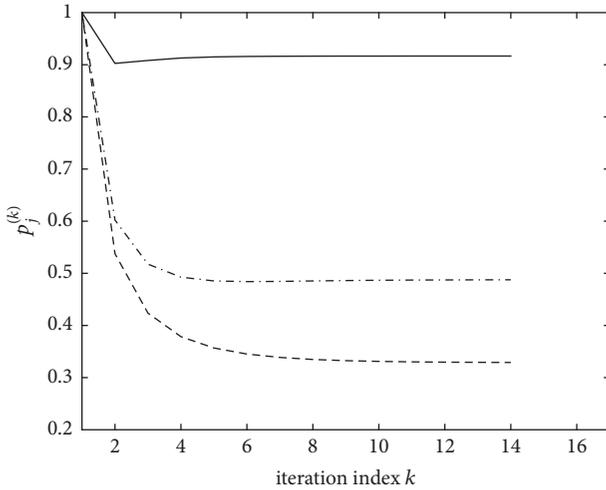
Figure 4 plots the convergence of the proposed two precoding methods which correspond to Main Algorithms 1 and 2 (see Algorithms 2 and 5). In this figure,  $N_T = N_D = N_E = 2$ ,  $\varepsilon_{ng} = 10$ , and the initial  $(\lambda, u)$  pairs are  $(0.1, 0.8)$  and  $(0, 0.5)$  for the first method and the second method, respectively. Two cases with SNR = 5 dB and SNR = 6 dB are considered. Observe that at SNR = 5 dB, the dual function  $g(\lambda, u)$  with the proposed first method decreases from 4.02 bit to 3.66 bit after 4 iterations. Further increasing iterations only marginally decrease  $g(\lambda, u)$ . Furthermore, the values of the dual function  $g(\lambda, u)$  using the proposed two methods converge to the same minimum, i.e., about 3.66 bit. Similar behavior can be found at SNR = 6 dB.

To sum up, the above results indicate that the proposed two methods are convergent; with the same system configuration, the solutions to problem P1 using two precoding methods are identical, which also demonstrates the validity of the proposed methods.



—  $j = 1$   
 .....  $j = 2$

(a) Case 1



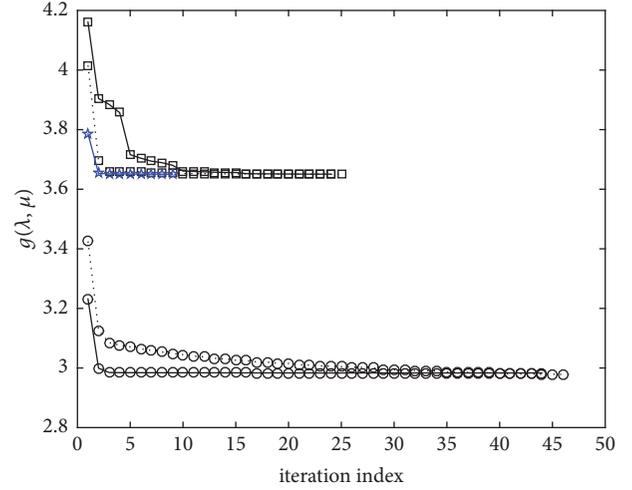
—  $j = 1$   
 - - -  $j = 2$   
 - · -  $j = 3$

(b) Case 2

FIGURE 3: The convergence of the Subalgorithm 2.1 (see Algorithm 3).

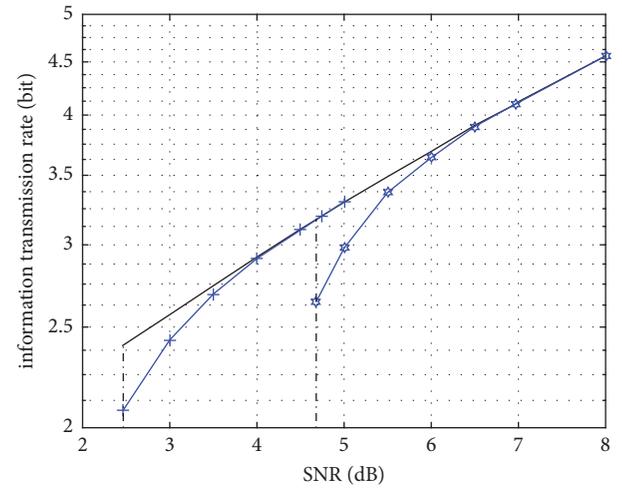
### 5.2. The R-E Trade-Offs with Different Transmission Schemes.

Figure 5 is a plot of the information transmission rate versus SNR with different energy thresholds. Three cases with  $\epsilon_{ng} = 6$ ,  $\epsilon_{ng} = 10$ , and no energy constraint are considered. For all cases,  $N_T = N_D = N_E = 2$ . The information rate is solved by the proposed first method. In fact, both methods yield the same result, as will be illustrated later. Obviously, it is seen that the case with no energy constraint is the upper bound of the other two cases. When SNR is less than a threshold, the problem P1 has no solution. This is because that even allocating the total power to  $\theta_{TE}$ 's eigenvector having maximum eigenvalue cannot meet the energy requirement. For instance,



·····  $\circ$  the proposed first method, SNR = 5 dB  
 ·····  $\square$  the proposed first method, SNR = 6 dB  
 —  $\circ$  the proposed second method, SNR = 5 dB  
 —  $\square$  the proposed second method, SNR = 6 dB

FIGURE 4: The convergence of the proposed two precoding methods.



— no energy constraint  
 — + energy threshold  $\epsilon_{ng} = 6$   
 — x energy threshold  $\epsilon_{ng} = 10$

FIGURE 5: The information transmission rate versus SNR, with different energy thresholds.

the SNR threshold for  $\epsilon_{ng} = 6$  is about 2.47 dB. This value can also be calculated by the first argument of Lemma 1. On the other front, at relatively low feasible SNR, there is a certain gap between the case with energy constraint and that with no energy constraint. This is due to the fact that the energy constraint decreases the achievable information rate. However, with increasing SNR, the gap becomes smaller and smaller until zero. This is due to the fact that the energy constraint becomes inactive gradually.

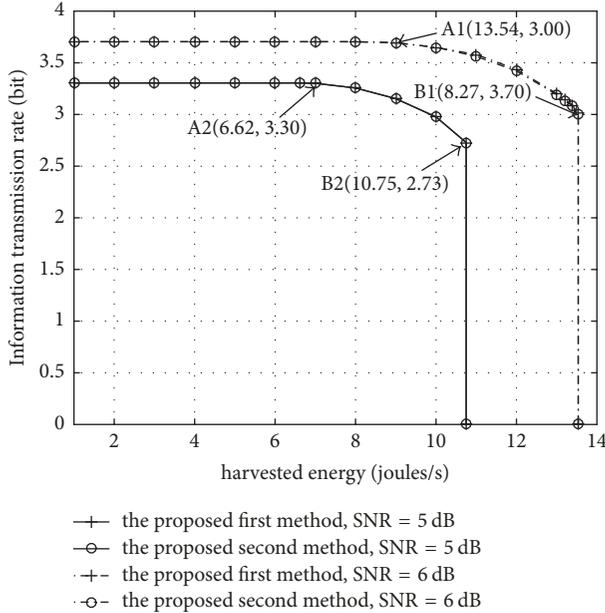


FIGURE 6: The R-E trade-off comparison with the proposed two precoding methods.

Figure 6 presents the R-E trade-offs with the proposed two precoding methods. Two cases with SNR = 5 dB and SNR = 6 dB are considered. For both cases, we set  $N_T = N_D = N_E = 2$  and  $\epsilon_{ng} = 10$ . Observe that, for each case, the trade-off curve with either the first or second method is horizontal at small energy threshold, and at a flexion point (A1 or A2), further increase of  $\epsilon_{ng}$  will result in the decrease of transmission rate. This is because that the information rate is nearly not affected by the energy constraint when  $\epsilon_{ng}$  is small and increasing  $\epsilon_{ng}$  will result in that the energy constraint becomes more and more active. There is also a cut-off point (B1 or B2), which corresponds to the maximum power, and exceeding the maximum power will result in no solution to  $P_1$ . This is because the harvested energy is limited by the maximum transmission power. The flexion point and cut-off point can also be calculated by Lemma 1. In fact, the coordinates of two points obtained by Lemma 1 are accordance with those obtained by numerical method. Moreover, with the same system configuration, the trade-off curve with the first method nearly overlaps that with the second method and hence, the proposed two methods are cross-validated.

Figure 7 compares several transmission schemes, including the proposed precoder, time-switching scheme, the isotropic transmission, and the precoder with hybrid CSI feedback [31]. Since the proposed two methods yield the same precoder, we may as well use one of them, say, the second one. Two scenarios including a  $2 \times 2$  MIMO WIPT system and a  $3 \times 3$  MIMO WIPT system are investigated, and the corresponding results are plotted in Figures 7(a) and 7(b), respectively. Note that, the correlation coefficients for  $\theta_{TD}$ ,  $\theta_{TE}$ ,  $\theta_{RD}$ , and  $\theta_{RE}$  are 0.3, 0.5, 0.2, and 0.3, respectively. In Figure 7(a), observe that the achievable R-E region with the proposed precoder contains that with the isotropic transmission or time-switching scheme. The R-E region of isotropic transmission

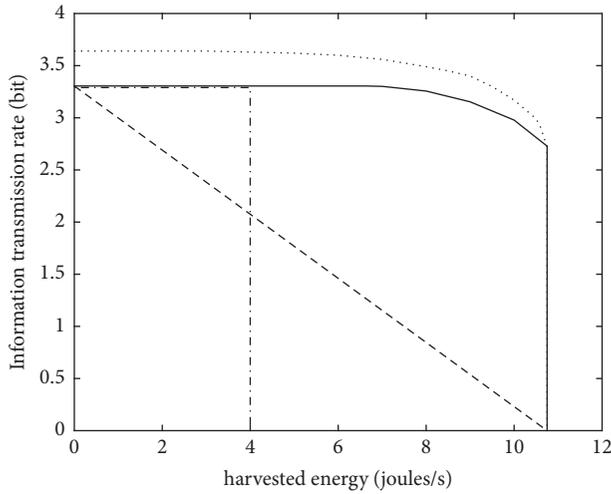
TABLE 3: The optimal covariance matrices and precoders with different cases.

Case	Optimal covariance matrix	Optimal Precoder
$2 \times 2$ system, SNR = 5 dB, $\epsilon_{ng} = 10$	$\begin{bmatrix} 1.58 & 1.31 \\ 1.31 & 1.58 \end{bmatrix}$	$\begin{bmatrix} 1.26 & 1.14 \\ 1.14 & 1.26 \end{bmatrix}$
	$\begin{bmatrix} 1.99 & 0.72 \\ 0.72 & 1.99 \end{bmatrix}$	$\begin{bmatrix} 1.41 & 0.85 \\ 0.84 & 1.41 \end{bmatrix}$
$3 \times 3$ system, SNR = 5 dB, $\epsilon_{ng} = 15$	$\begin{bmatrix} 1.05 & 0.53 & 0.36 \\ 0.53 & 1.07 & 0.53 \\ 0.36 & 0.53 & 1.05 \end{bmatrix}$	$\begin{bmatrix} 1.02 & 0.73 & 0.60 \\ 0.73 & 1.03 & 0.73 \\ 0.60 & 0.73 & 1.02 \end{bmatrix}$

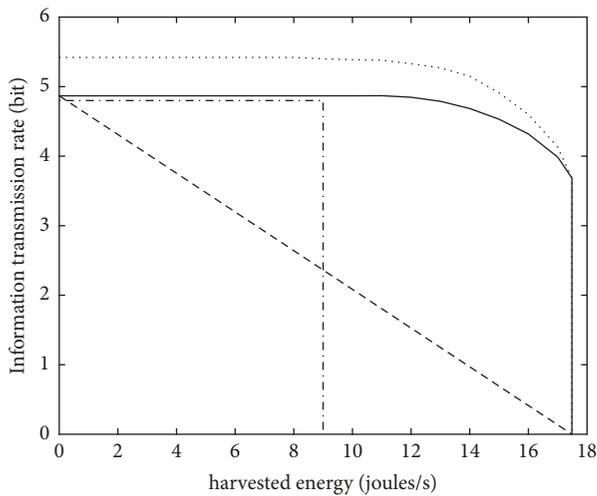
is a rectangular, whose upper boundary is very close to that of  $\Omega(P_T)$ . This is because that the coefficient for  $\theta_{TD}$  is small and hence the isotropic transmission is nearly optimal for data transmission. The boundary of  $\Omega'(P_T)$  is simply a line by connecting the two points  $(J_{\max}, 0)$  and  $(\epsilon_{ng, \max}, 0)$ . This can be derived by ranging  $\beta$  from 0 to 1. Also observe that the R-E region of the scheme with hybrid CSI feedback is much broader than the proposed scheme. This is because that the former utilizes the instant CSI of the information channel, whereas the latter only has the statistical CSI of the information channel. That is, the increase of average rate is an exchange of additional overhead. Similar phenomena can be found in Figure 7(b). In addition, the optimal precoders with a few cases are listed in Table 3.

To sum up, the results in Figure 5 imply that the energy constraint decreases the achievable information rate at low feasible SNR evidently and it becomes gradually inactive with increasing SNR. The results in Figure 6 indicate that the proposed two methods can be cross-validated and the precalculation of the flexion point and cut-off point by Lemma 1 will help to find a rough R-E boundary curve. The results in Figure 7 demonstrate that the proposed precoder is superior to both the isotropic transmission and time-switching scheme in terms of the R-E region; however, it is inferior to the precoder with hybrid CSI feedback, since the latter utilizes more CSI.

**5.3. The Effects of the Number of Transmitter Antennas on the R-E Trade-Off.** In Figures 8 and 9,  $N_E = N_D = 2$  and SNR = 5 dB. The figures investigate the behavior of the proposed algorithms when the number of transmit antennas is large. Observe that, in Figure 8, we find that increasing  $N_T$  results in much broader R-E regions. The right boundary of the R-E region, whose horizontal coordinate is the maximum harvested energy  $\epsilon_{ng, \max}$ , has a limit corresponding to the limit in Figure 9(a). In Figure 9(a), increasing  $N_T$  results in an  $\epsilon_{ng, \max}$  limit around 35.7 joules/s. On the other hand, the maximum achievable rate of the R-E region is characterized by Figure 9(b). In Figure 9(b), for small  $N_T$ , increasing  $N_T$  makes the maximum rate increase rapidly. For large  $N_T$ ,



— the proposed method  
 --- time division multiplexing  
 -.- isotropic transmission  
 ..... the precoder with hybrid CSI

(a) A  $2 \times 2$  MIMO WIPT system

— the proposed method  
 --- time division multiplexing  
 -.- isotropic transmission  
 ..... the precoder with hybrid CSI

(b) A  $3 \times 3$  MIMO WIPT system

FIGURE 7: The R-E trade-offs with several transmission schemes.

e.g., 64, further increase of  $N_T$  can only marginal increase the maximum rate. However, we cannot conclude that the maximum rate has a limit because the transmitter utilizes the CSI of information channel.

Recall that the capacity or the maximum rate for an open-loop point-to-point MIMO system has a limit when the number of transmitter antennas is large [47]. However, it does not hold for a closed-loop MIMO system because the transmitter has the CSI.

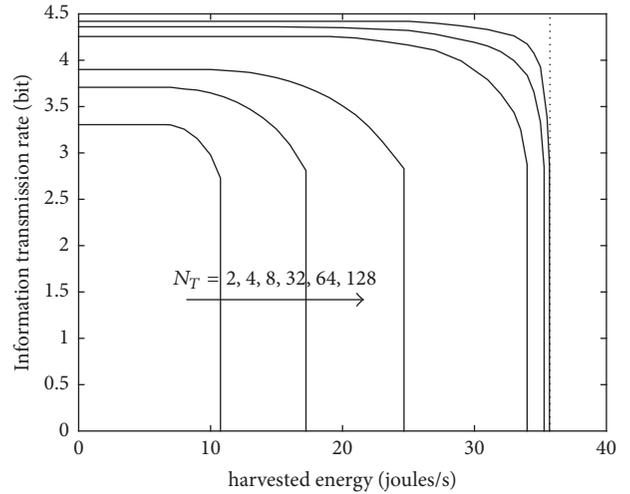
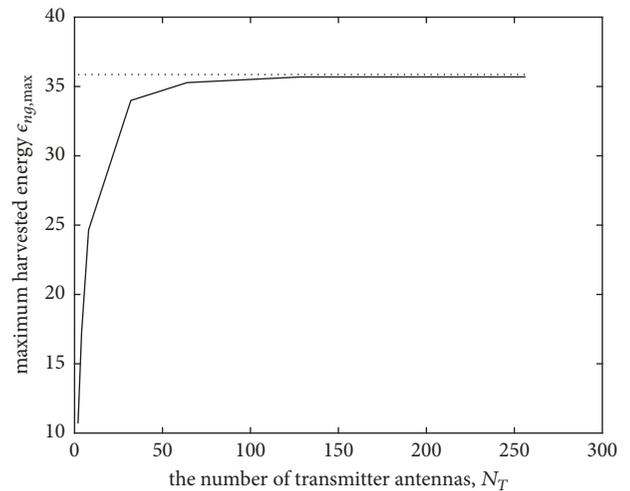
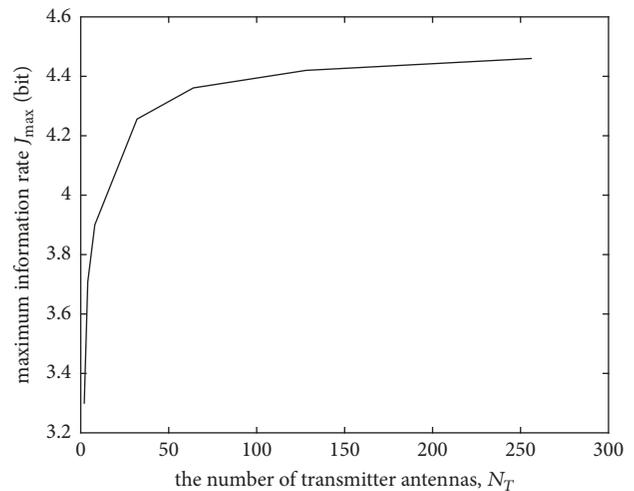


FIGURE 8: The effect of transmitter antenna on the rate-energy regions.

(a)  $\varepsilon_{ng,max}$  versus  $N_T$ (b)  $J_{max}$  versus  $N_T$ FIGURE 9: The effect of transmitter antenna number,  $N_T$ , on the maximum information rate  $J_{max}$  and harvested energy  $\varepsilon_{ng,max}$ .

## 6. Conclusions

For a typical three-node wireless MIMO information and energy transmission system, the precoder design with only covariance feedback is studied. By using the convex optimization technique, and after a few transformations, we transform the precoder design problem into a dual one. The dual problem is proved to have no duality gap by *Slater's condition*. Then, two algorithms have been proposed to solve this problem. The first one decomposes the transmission covariance and emphasizes numerically searching the optimal precoder, while the second one emphasizes exploiting the structure of the covariance, which is presented by Proposition 4. Their convergence, complexity, etc. have also been addressed by theoretical analysis or computer simulation.

On the other hand, the R-E region is often used to measure the WIPT system performance. We have exploited the key features of the R-E boundary curve by Lemma 1. In fact, the precalculation of the flexion point and cut-off point by this lemma will help to find a rough R-E boundary curve. Furthermore, we compare the proposed methods with a few existing transmission schemes. Simulation results show that the proposed precoder is superior to the isotropic transmission and time-switching scheme and it has the broadest R-E region compared with the other two regions. However, the proposed precoder is inferior to the precoder with hybrid CSI feedback, since the latter one utilizes more CSI.

The proposed precoding methods and above results can be a reference for practical system implementation in the future. We will further improve the system performance by incorporating the other wireless transmission techniques [48–52].

## Appendix

### A. Proof of Lemma 1

Since

$$\begin{aligned} \text{tr}(\mathbf{Q}\boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) &\leq \text{tr}(\mathbf{Q}) \lambda_{\max}(\boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}) \\ &\leq P_T \lambda_{\max}(\boldsymbol{\theta}_{TE}) \text{tr}(\boldsymbol{\theta}_{RE}), \end{aligned} \quad (\text{A.1})$$

it is clear that the first conclusion of Lemma 1 holds. Then, when  $\varepsilon_{ng} < \text{tr}(\mathbf{Q}'_{OPT}\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE})$ , where  $\mathbf{Q}'_{OPT}$  is the solution of  $P'_1$ , it is deduced that  $\mathbf{Q}'_{OPT}$  is also the solution of P1 because

of the following: (1) the objective function of P1 is concave and all constraints are linear; therefore, P1 is convex and has a global optimal solution; (2)  $\mathbf{Q}'_{OPT}$  maximizes the objective function of P1 and satisfies all constraints. Obviously, the energy constraint of P1 is not active because the equality sign does not hold.

### B. Proof of Slater's Condition

For one thing, if  $\varepsilon_{ng} = P_T \lambda_{\max}(\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE})$ , the structure of the optimal covariance is given by

$$\mathbf{Q}^{OPT} = \mathbf{U}_{\boldsymbol{\theta}_{TE}} \text{diag}[P_T \ 0 \ \cdots \ 0] \mathbf{U}_{\boldsymbol{\theta}_{TE}}^H, \quad (\text{B.1})$$

where  $\mathbf{U}_{\boldsymbol{\theta}_{TE}}$  corresponds with the eigenvectors of  $\boldsymbol{\theta}_{TE}$ . Without loss of generality, we assume that the eigenvalues of  $\boldsymbol{\theta}_{TE}$  are sorted in descending order, so that the total power  $P_T$  is allocated to the eigenvector with the maximum eigenvalue. In this scenario, there is no need transforming P1 into a dual problem.

For another, if  $\varepsilon_{ng} < P_T \lambda_{\max}(\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE})$ , let  $\mathbf{Q} = \mathbf{U}_{\boldsymbol{\theta}_{TE}} \times \text{diag}[P_T - \delta \ 0 \ \cdots \ 0] \mathbf{U}_{\boldsymbol{\theta}_{TE}}^H$ , where  $P_T - \varepsilon_{ng} / \lambda_{\max}(\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE}) > \delta > 0$ . We have  $\text{tr}(\mathbf{Q}) = (P_T - \delta) < P_T$  and  $\text{tr}(\mathbf{Q}\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE}) = (P_T - \delta) \times \lambda_{\max}(\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE}) > \varepsilon_{ng}$ . Therefore, there exists a strict feasible point for P1, and the Slater's condition is satisfied for  $\varepsilon_{ng} < P_T \times \lambda_{\max}(\boldsymbol{\theta}_{TE})\text{tr}(\boldsymbol{\theta}_{RE})$ .

### C. Proof of Proposition 4

First, we reexpress  $g(\lambda, u)$  by transforming  $\mathbf{Q}$ . With (16), letting  $\tilde{\mathbf{Q}} = \mathbf{A}^{1/2}\mathbf{Q}\mathbf{A}^{1/2}$ , one has

$$\begin{aligned} g(\lambda, u) &= \max_{\mathbf{Q} \geq \mathbf{0}} E \left\{ \log \det [\mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{Q} \mathbf{H}_D^H] \right\} - \text{tr} \{ \mathbf{A} \mathbf{Q} \} + T_2 \\ &= \max_{\mathbf{A}^{-1/2} \tilde{\mathbf{Q}} \mathbf{A}^{-1/2} \geq \mathbf{0}} E \left\{ \log \det [\mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{A}^{-1/2} \tilde{\mathbf{Q}} \mathbf{A}^{-1/2} \mathbf{H}_D^H] \right\} \\ &\quad - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_2, \end{aligned} \quad (\text{C.1})$$

where  $\mathbf{A} = u\mathbf{I}_{N_T} - \lambda \text{tr}(\boldsymbol{\theta}_{RE})\boldsymbol{\theta}_{TE}$ . Since  $\mathbf{A} > \mathbf{0}$  and  $\mathbf{Q} \geq \mathbf{0}$ , it is deduced that the condition  $\mathbf{Q} \geq \mathbf{0}$  is equivalent to  $\tilde{\mathbf{Q}} \geq \mathbf{0}$ . Hence, (C.1) can be written as

$$\begin{aligned} g(\lambda, u) &= \max_{\tilde{\mathbf{Q}} \geq \mathbf{0}} E \left\{ \log \det [\mathbf{I}_{N_D} + \mathbf{H}_D \mathbf{A}^{-1/2} \tilde{\mathbf{Q}} \mathbf{A}^{-1/2} \mathbf{H}_D^H] \right\} - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_2 \\ &= \max_{\tilde{\mathbf{Q}} \geq \mathbf{0}} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \boldsymbol{\theta}_{RD}^{1/2} \mathbf{H}_{\omega, D} \boldsymbol{\theta}_{TD}^{1/2} \mathbf{A}^{-1/2} \tilde{\mathbf{Q}} \mathbf{A}^{-1/2} \left( \boldsymbol{\theta}_{TD}^{1/2} \right)^H \times \mathbf{H}_{\omega, D}^H \left( \boldsymbol{\theta}_{RD}^{1/2} \right)^H \right] \right\} - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_2 \\ &= \max_{\tilde{\mathbf{Q}} \geq \mathbf{0}} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \boldsymbol{\theta}_{RD}^{1/2} \mathbf{H}_{\omega, D} \boldsymbol{\theta}_{TD}^{1/2} \mathbf{A}^{-1/2} \tilde{\mathbf{Q}} \mathbf{A}^{-1/2} \boldsymbol{\theta}_{TD}^{1/2} \mathbf{H}_{\omega, D}^H \boldsymbol{\theta}_{RD}^{1/2} \right] \right\} - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_2 \end{aligned} \quad (\text{C.2})$$

Notice that although  $\boldsymbol{\theta}_{TD}^{1/2}$  is positive semidefinite and  $\mathbf{A}^{1/2}$  is positive definite, the product of them,  $\boldsymbol{\theta}_{TD}^{1/2}\mathbf{A}^{-1/2}$ , may not be positive semidefinite.

**Then**, we further arrange the expectation term of  $g(\lambda, u)$ . Since  $\boldsymbol{\theta}_{TD}^{1/2}\mathbf{A}^{-1/2}$  may not be positive semidefinite, the eigendecomposition is no longer suitable. Let  $\mathbf{B} = \boldsymbol{\theta}_{TD}^{1/2}\mathbf{A}^{-1/2}$  and its SVD is expressed as

$$\mathbf{B} = \boldsymbol{\theta}_{TD}^{1/2}\mathbf{A}^{-1/2} = \mathbf{U}_B\boldsymbol{\Lambda}_B\mathbf{V}_B^H. \quad (\text{C.3})$$

Meanwhile, the eigendecomposition of  $\boldsymbol{\theta}_{RD}^{1/2}$  is expressed as

$$\boldsymbol{\theta}_{RD}^{1/2} = \mathbf{U}_{\theta_{RD}}\boldsymbol{\Lambda}_{\theta_{RD}}^{1/2}\mathbf{U}_{\theta_{RD}}^H. \quad (\text{C.4})$$

Substituting (C.3) and (C.4) into (C.2), one obtains

$$g(\lambda, u) = \max_{\tilde{\mathbf{Q}} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{U}_{\theta_{RD}}\boldsymbol{\Lambda}_{\theta_{RD}}^{1/2}\mathbf{U}_{\theta_{RD}}^H \mathbf{H}_{\omega,D} \mathbf{U}_B\boldsymbol{\Lambda}_B\mathbf{V}_B^H \times \tilde{\mathbf{Q}}\mathbf{V}_B\boldsymbol{\Lambda}_B\mathbf{U}_B^H \mathbf{H}_{\omega,D}^H \mathbf{U}_{\theta_{RD}}\boldsymbol{\Lambda}_{\theta_{RD}}^{1/2}\mathbf{U}_{\theta_{RD}}^H \right] \right\} - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_{const}. \quad (\text{C.5})$$

Letting  $\tilde{\mathbf{H}}_D = \boldsymbol{\Lambda}_{\theta_{RD}}^{1/2}\mathbf{U}_{\theta_{RD}}^H \mathbf{H}_{\omega,D} \mathbf{U}_B\boldsymbol{\Lambda}_B$  and noticing that pre- or postmultiplying a unitary matrix does not change the distribution of  $\mathbf{H}_{\omega,D}$ , it is easily found that  $\tilde{\mathbf{H}}_D$  is a random matrix whose entries are independent, with zero-mean and arbitrary variance. Then, (C.5) is further written as

$$\begin{aligned} g(\lambda, u) &= \max_{\tilde{\mathbf{Q}} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \mathbf{U}_{\theta_{RD}}\tilde{\mathbf{H}}_D \times \mathbf{V}_B^H \tilde{\mathbf{Q}}\mathbf{V}_B \tilde{\mathbf{H}}_D^H \mathbf{U}_{\theta_{RD}}^H \right] \right\} \\ &\quad - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_{const}. \end{aligned} \quad (\text{C.6})$$

Letting  $\bar{\mathbf{H}}_D = \mathbf{U}_{\theta_{RD}}\tilde{\mathbf{H}}_D\mathbf{V}_B^H$  and substituting it into (C.6), one has

$$\begin{aligned} g(\lambda, u) &= \max_{\tilde{\mathbf{Q}} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \bar{\mathbf{H}}_D \tilde{\mathbf{Q}} \bar{\mathbf{H}}_D^H \right] \right\} \\ &\quad - \text{tr} \{ \tilde{\mathbf{Q}} \} + T_{const}. \end{aligned} \quad (\text{C.7})$$

**Finally**, we transform  $g(\lambda, u)$  into an equivalent problem and present the eigenvectors of the optimal  $\tilde{\mathbf{Q}}^{OPT}$ . Assume that the total transmission power, corresponding to the trace of the optimal  $\tilde{\mathbf{Q}}^{OPT}$ , is  $P_T^*$ . Now, consider what are the eigenvectors of  $\tilde{\mathbf{Q}}^{OPT}$ . It is easy to find that this task is equivalent to the following problem:

$$\begin{aligned} \max_{\tilde{\mathbf{Q}} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \bar{\mathbf{H}}_D \tilde{\mathbf{Q}} \bar{\mathbf{H}}_D^H \right] \right\} \\ \text{s.t. } \text{tr} \{ \tilde{\mathbf{Q}} \} = P_T^*. \end{aligned} \quad (\text{C.8})$$

Obviously, it is also equivalent to

$$\begin{aligned} \max_{\tilde{\mathbf{Q}} \geq 0} E \left\{ \log \det \left[ \mathbf{I}_{N_D} + \bar{\mathbf{H}}_D \tilde{\mathbf{Q}} \bar{\mathbf{H}}_D^H \right] \right\} \\ \text{s.t. } \text{tr} \{ \tilde{\mathbf{Q}} \} \leq P_T^*, \end{aligned} \quad (\text{C.9})$$

because the objective function achieves its maximum value at the maximum transmission power. Since  $\bar{\mathbf{H}}_D$  satisfies the definition in [32, Definition 2], according to [32, Theorem 1], the eigenvectors of  $\tilde{\mathbf{Q}}^{OPT}$  are given by the columns of  $\mathbf{V}_B$ . Since  $\tilde{\mathbf{Q}}^{OPT} = \mathbf{A}^{1/2}\mathbf{Q}^{OPT}\mathbf{A}^{1/2} = \mathbf{V}_B\boldsymbol{\Lambda}\mathbf{V}_B^H$ , we have the optimal covariance  $\mathbf{Q}^{OPT} = \mathbf{A}^{-1/2}\mathbf{V}_B\boldsymbol{\Lambda}\mathbf{V}_B^H\mathbf{A}^{-1/2}$  and hence, this proposition is proved.

## D. Derivation of (29)

To start with, we have the following lemma.

**Lemma D.1.** For the matrix  $\mathbf{D} = \mathbf{C} + \mathbf{x}\mathbf{y}^H$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are two column vectors, the inverse of  $\mathbf{D}$  is  $\mathbf{D}^{-1} = \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{x}\mathbf{y}^H\mathbf{C}^{-1}/(1 + \mathbf{y}^H\mathbf{C}^{-1}\mathbf{x})$ .

With the above lemma, we will reexpress the optimality condition (27).

$$\begin{aligned} E \left\{ \text{tr} \left[ \left( \mathbf{I}_{N_D} + \sum_{i=1}^{N_T} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H \right)^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} \\ = E \left\{ \text{tr} \left[ \left( \mathbf{I}_{N_D} + \sum_{i \neq l} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H + p_l \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right)^{-1} \right. \right. \\ \left. \left. \cdot \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} = E \left\{ \text{tr} \left[ \left( \mathbf{C}_l^{-1} \right. \right. \right. \\ \left. \left. - \frac{\mathbf{C}_l^{-1} p_l \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\}, \end{aligned} \quad (\text{D.1})$$

where  $\mathbf{C}_l \triangleq \mathbf{I}_{N_D} + \sum_{i \neq l} p_i \tilde{\mathbf{h}}_{D,i} \tilde{\mathbf{h}}_{D,i}^H$ . Noticing  $\text{tr}[\mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} \times \tilde{\mathbf{h}}_{D,l}^H] = \text{tr}[\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}] = \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}$ , (D.1) is further arranged as

$$\begin{aligned} E \left\{ \text{tr} \left[ \left( \mathbf{C}_l^{-1} - \frac{\mathbf{C}_l^{-1} p_l \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \right] \right\} \\ = E \left\{ \text{tr} \left[ \left( \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} \right. \right. \right. \\ \left. \left. - \frac{p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) \right] \right\} \\ = E \left( \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} - \frac{p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l} \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right) \\ = E \left( \frac{\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}} \right). \end{aligned} \quad (\text{D.2})$$

Therefore, the optimality condition (27) is reexpressed as

$$E\left(\frac{\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}\right) = 1, \quad \text{for } p_l > 0;$$

$$E\left(\frac{\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}{1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}}\right) \leq 1, \quad \text{for } p_l = 0. \quad (\text{D.3})$$

With (D.3), we can prove that (29) holds. For one thing, consider the case of  $E(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) \leq 1$ . If  $p_l > 0$ , one has  $E[(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) / (1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l})] < 1$ , because the denominator is larger than 1. However, this contradicts the optimality condition (D.3). Hence,  $p_l = 0$  for  $E(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) \leq 1$ . For another, consider another case of  $E(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) > 1$ . If  $p_l = 0$ , one obtains  $E(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) = E[(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) / (1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l})] > 1$ , which also contradicts (D.3). Therefore, in this case we have  $p_l > 0$  and  $E[(\tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l}) / (1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l})] = 1$ , or equivalently,  $p_l = 1 - E[1 / (1 + p_l \tilde{\mathbf{h}}_{D,l}^H \mathbf{C}_l^{-1} \tilde{\mathbf{h}}_{D,l})]$ . In a word, (29) has been proved.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Disclosure

This paper has been presented in part at the IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Hong Kong, China, 2016.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Science Foundation of China (no. 61601275); Guangdong Province Natural Science Fund (no. 2016A030313068); Guangdong Natural Science Funds for Distinguished Young Scholar (no. 2014A030306027); the Innovation Team Project of Guangdong Province University (no. 2016KCXTD017); the Science and Technology Program of Guangzhou (no. 201807010103); highly educated talent research start-up fund of Nanjing Forestry University (no. GXL2016030); the Scientific Research Projects of Guangzhou Municipal Universities (no. 1201430722); the Scientific Research Project of Education Department of Guangdong, China (no. 2017GKTSCX045); Science and Technology Program of Guangzhou, China (no. 201707010389); the Scientific Research Project of Guangzhou Municipal University (no. 1201620439); and Qingshanhu Young Scholar Program in GZPYP (no. 2016Q001).

## References

- [1] X. Lai, W. Zou, D. Xie, X. Li, and L. Fan, "DF relaying networks with randomly distributed interferers," *IEEE Access*, vol. 5, pp. 18909–18917, 2017.
- [2] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secure multiple amplify-and-forward relaying with cochannel interference," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1494–1505, 2016.
- [3] L. Fan, S. Zhang, T. Q. Duong, and G. K. Karagiannidis, "Secure switch-and-stay combining (SSSC) for cognitive relay networks," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 70–82, 2016.
- [4] L. Fan, R. Zhao, F.-K. Gong, N. Yang, and G. K. Karagiannidis, "Secure multiple amplify-and-forward relaying over correlated fading channels," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2811–2820, 2017.
- [5] G. Pan, H. Lei, Y. Deng et al., "On secrecy performance of MISO SWIPT systems with TAS and imperfect CSI," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3831–3843, 2016.
- [6] G. Huang, Q. Zhang, and J. Qin, "Joint time switching and power allocation for multicarrier decode-and-forward relay networks with SWIPT," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2284–2288, 2015.
- [7] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 104–110, 2014.
- [8] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secrecy cooperative networks with outdated relay selection over correlated fading channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7599–7603, 2017.
- [9] R. Zhao, Y. Yuan, L. Fan, and Y.-C. He, "Secrecy performance analysis of cognitive decode-and-forward relay networks in nakagami-m fading channels," *IEEE Transactions on Communications*, vol. 65, no. 2, pp. 549–563, 2017.
- [10] L. R. Varshney, "Transporting information and energy simultaneously," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '08)*, pp. 1612–1616, IEEE, Toronto, Canada, July 2008.
- [11] P. Grover and A. Sahai, "Shannon meets tesla: wireless information and power transfer," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '10)*, pp. 2363–2367, Austin, Tex, USA, June 2010.
- [12] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: architecture design and rate-energy tradeoff," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4754–4761, 2013.
- [13] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multiuser OFDM systems with wireless information and power transfer," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '13)*, pp. 3823–3828, Shanghai, China, April 2013.
- [14] H. Xing, L. Liu, and R. Zhang, "Secrecy wireless information and power transfer in fading wiretap channel," in *Proceedings of the 2014 1st IEEE International Conference on Communications, ICC 2014*, pp. 5402–5407, Australia, June 2014.
- [15] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3622–3636, 2013.

- [16] L. Liu, R. Zhang, and K.-C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 288–300, 2013.
- [17] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4808–4818, 2012.
- [18] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer in multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2282–2294, 2014.
- [19] A. M. Fouladgar and O. Simeone, "On the transfer of information and energy in multi-user systems," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1733–1736, 2012.
- [20] M. Zhao and Y. Yang, "Bounded relay hop mobile data gathering in wireless sensor networks," *IEEE Transactions on Computers*, vol. 61, no. 2, pp. 265–277, 2012.
- [21] Z. Xiang and M. Tao, "Robust beamforming for wireless information and power transmission," *IEEE Wireless Communications Letters*, vol. 1, no. 4, pp. 372–375, 2012.
- [22] K. Huang and E. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5972–5986, 2013.
- [23] W. Wu and B. Wang, "Robust secrecy beamforming for wireless information and power transfer in multiuser MISO communication system," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–11, 2015.
- [24] X. Chen, C. Yuen, and Z. Zhang, "Wireless energy and information transfer tradeoff for limited-feedback multiantenna systems with energy beamforming," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 407–412, 2014.
- [25] C.-F. Liu, M. Maso, S. Lakshminarayana, C.-H. Lee, and T. Q. S. Quek, "Simultaneous wireless information and power transfer under different CSI acquisition schemes," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 1911–1926, 2015.
- [26] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989–2001, 2013.
- [27] Q. Zhang, Q. Li, and J. Qin, "Beamforming design for OSTBC-based AF-MIMO two-way relay networks with simultaneous wireless information and power transfer," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7285–7296, 2016.
- [28] Y. Huang and B. Clerckx, "Joint wireless information and power transfer in a three-node autonomous MIMO relay network," *IEEE Communications Letters*, vol. 19, no. 7, pp. 1113–1116, 2015.
- [29] I. Krikidis, S. Sasaki, S. Timotheou, and Z. Ding, "A low complexity antenna switching for joint wireless information and energy transfer in MIMO relay channels," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1577–1587, 2014.
- [30] W. Zhou, D. Deng, and Z. Shao, "The precoder design with covariance feedback for information and energy transmission systems," *IEEE ICSPCC*, pp. 1–5, 2016.
- [31] C. Xing, N. Wang, J. Ni, Z. Fei, and J. Kuang, "MIMO beamforming designs with partial CSI under energy harvesting constraints," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 363–366, 2013.
- [32] A. M. Tulino and A. Lozano, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 662–671, 2006.
- [33] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, 2004.
- [34] W. Zhou, "Precoding methods for multi-input multi-output interference channels with channel covariance feedback," *IEEE Communications*, vol. 9, no. 4, pp. 517–525, 2015.
- [35] E. A. Jorswieck, R. Mochaourab, and M. Mittelbach, "Effective capacity maximization in multi-antenna channels with covariance feedback," *IEEE Transactions on Wireless Communications*, vol. 9, no. 10, pp. 2988–2993, 2010.
- [36] R. T. Krishnamachari, M. K. Varanasi, and K. Mohanty, "MIMO systems with quantized covariance feedback," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 485–495, 2014.
- [37] X. Li, S. Jin, X. Gao, and K.-K. Wong, "Near-optimal power allocation for MIMO channels with mean or covariance feedback," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 289–300, 2010.
- [38] X. Wang, H. Zhang, L. Fan, and Y. Li, "Performance of distributed switch-and-stay combining for cognitive relay networks with primary transceiver," *Wireless Personal Communications*, vol. 97, no. 2, pp. 3031–3042, 2017.
- [39] W. Tan, M. Matthaiou, S. Jin, and X. Li, "Spectral efficiency of DFT-based processing hybrid architectures in massive MIMO," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 586–589, 2017.
- [40] F. Shi, L. Fan, X. Liu, Z. Na, and Y. Liu, "Probabilistic caching placement in the presence of multiple eavesdroppers," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 2104162, 10 pages, 2018.
- [41] Y. Liang, H. Wu, G. Huang, J. Yang, and H. Wang, "Thermal performance and service life of vacuum insulation panels with aerogel composite cores," *Energy and Buildings*, vol. 154, pp. 606–617, 2017.
- [42] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2082–2085, 2015.
- [43] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.
- [44] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [45] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, Pa, USA, 1997.
- [46] R. G. Bland, D. Goldfarb, and M. J. Todd, "The ellipsoid method: a survey," *Operations Research. The Journal of the Operations Research Society of America*, vol. 29, no. 6, pp. 1039–1091, 1981.
- [47] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [48] X. Lai, J. Xia, M. Tang, H. Zhang, and J. Zhao, "Cache-aided multiuser cognitive relay networks with outdated channel state information," *IEEE Access*, pp. 1–10, 2018.
- [49] J. Xia, F. Zhou, X. Lai et al., "Cache aided decode-and-forward relaying networks: from the spatial view," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 5963584, 9 pages, 2018.
- [50] J. Yang, H. Wu, M. Wang, S. He, and H. Huang, "Prediction and optimization of radiative thermal properties of ultrafine fibrous insulations," *Applied Thermal Engineering*, vol. 104, pp. 394–402, 2016.

- [51] J. Yuan, S. Jin, W. Xu, W. Tan, M. Matthaiou, and K.-K. Wong, "User-centric networking for dense C-RANs: high-SNR capacity analysis and antenna selection," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 5067–5080, 2017.
- [52] G. Huang and W. Tu, "Optimal resource allocation in wireless-powered OFDM relay networks," *Computer Networks*, vol. 104, pp. 94–107, 2016.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

