Lyapunov Optimized Cooperative Communications With Stochastic Energy Harvesting Relay

Chengrun Qiu, Yang Hu, and Yan Chen, *Senior Member, IEEE*

***Abstract*—Energy harvesting (EH) wireless communications have become more and more popular due to its capability of effectively reducing the battery replacement time. In this paper, we focus on decode-and-forward-based cooperative wireless communications with an EH relay. The stochastic characteristics of the harvestable energy and the wireless channel make it difficult to optimally manage the energy harvested at the relay for better communication performance. We formulate such a problem as an optimization by minimizing the long-term average symbol error rate (SER) subject to a battery constraint. Based on the Lyapunov optimization theory, we utilize the virtual queue technique and transform the optimization problem into a drift-plus-penalty minimization. Then, we conduct theoretic analysis of the optimal energy strategy derived by the proposed scheme. Specifically, we show the convexity of the drift-plus-penalty minimization problem, and prove that the virtual queue is bounded and the battery constraint is always satisfied. We also show that the proposed optimal energy management strategy is limited by an upper bound that is independent of the operation time index, derive the closed-form expression for the asymptotical average SER, and analyze the corresponding diversity order and EH gain. Finally, simulation results using the real solar irradiance data show that the proposed algorithm can achieve much better performance in terms of both average SER and diversity order, compared with the Markov decision process-based method.**

***Index Terms*—Cooperative communication, diversity order, energy harvesting (EH) gain, energy harvesting, Lyapunov optimization.**

I. Introduction

WIRELESS communications may suffer a temporal disturbance due to the stochastic nature of wireless channel. Cooperative communications have recently raised much attention as it can make up the shortness by applying the transmission relay and achieve spatial diversity gains [1]. Cooperative communications are especially useful in situations where the communication node's size is limited and is not able to accommodate multiple transmission terminals. Different cooperative communication patterns have been proposed, among which decode-and-forward (DF) and amplify-and-forward (AF) are considered as the most popular patterns to provide full diversity gains [2], [3]. In DF, the relay first decodes the signals from the source, and then encodes the decoded signals and transmits to the destination if the decoding is successful. On the other hand, in AF, the relay directly amplifies the received signals and sends to the destination without judging the signals. Comparing with AF, DF can achieve better symbol error rate (SER) performances [1].

Another important problem of the wireless communications is the energy supplement of the transmission terminals. Due to the limited size of the transmission nodes, long lifetime battery may not be suitable and battery with small capacity has to be replaced frequently, which is very inconvenient and may cause harm to the components and/or the communications. In the past few years, energy harvesting (EH) wireless communications have gained much attention. With the aid of chargeable battery and renewable energy, the time of battery replacement can be reduced effectively [4]. Hence, it is very attractive to apply the EH techniques in the cooperative communication systems. Three typical EH architectures have been proposed in the EH system [5]. The first one is harvest-use (HU), which directly uses the harvested energy without storage. The second one is HU-storage (HUS), where the harvested energy is saved and only available at the next time instance. The last one is harvest-storage-use (HSU), where the harvested energy can be immediately used and the remaining is stored in the energy battery for future use.

While the EH wireless communications have the unique advantages, the corresponding energy management is challenging due to the stochastic nature of harvested energy and wireless channel. To solve this challenge, Markov decision process (MDP) has been widely used [6]–[10]. Ulukus *et al.* [6] proposed to use MDP to maximize the throughput of the point-to-point EH wireless communications. With MDP, the throughput of the EH communications with one-bit channel feedback is maximized by finding a suitable MQAM modulation [7]. By utilizing the MDP technique, Blasco *et al.* [8] derived the optimal power level to maximize the long-term expected throughput, while Rezaee *et al.* [9] maximized the single-user throughput of an EH communication system with continuous energy and data arrivals and derived a three-step optimal energy scheduling algorithm with the knowledge of the harvested energy and data arrival. Ku *et al.* [10] considered the problem to maximize the net bit rate of an EH wireless communication system and utilized MDP to derive the

Manuscript received December 19, 2017; accepted January 10, 2018. Date of publication January 15, 2018; date of current version April 10, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61672137 and Grant 61602090, in part by the 111 Project under Grant B17008, and in part by the Thousand Youth Talents Program of China (to Y. Chen). *(Corresponding author: Yan Chen.)*

The authors are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China, and also with the Center of Future Media, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: cr_qiu@std.uestc.edu.cn; yanghu@uestc.edu.cn; eecyan@uestc.edu.cn).

Digital Object Identifier 10.1109/JIOT.2018.2793850

optimal energy-and-modulation management with the states of the solar, battery, and channel. There have been also some works on utilizing MDP for EH cooperative communications [11]–[16]. In [11], MDP is utilized to minimize the average SER of the DF-based cooperative communications with an EH relay. The long-term average outage probabilities of the cooperative communications in AF and DF protocols with an EH relay code is minimized by means of MDP [12]–[14]. The average throughput of two-hop AF-based EH cooperative communication is maximized by MDP [15]. Li *et al.* [16] developed strategies for more efficient energy usage with MDP. However, the computational complexity of an MDP is a very challenging problem when the state and action spaces are large.

An alternative method to solve the challenge is the water-filling technique where the KKT optimality conditions are utilized to derive the optimal power and/or energy [17]–[20]. In [17], a directional water-filling algorithm is proposed to maximize the number of bits sent in a finite duration and minimize the time taken to sent a given amount of data for an EH system with fading channels. In [18], the problem of the energy allocation over a finite horizon is studied by taking into account the time-varying channel conditions and energy sources. It was shown that when the battery capacity is unlimited and the information of the channel conditions and harvested energy is known, the water-filling energy allocation solution is optimal. An iterative dynamic water-filling algorithm is proposed to optimally schedule the energy for fading multiple access channels with EH [19]. Such an iterative algorithm was shown to be able to converge in a few iterations and achieved much better sum-rate performance than various suboptimal energy scheduling algorithms. While achieving promising results, the water filling algorithm is limited only to the optimization with finite time slots and cannot be used for the problem with long-term optimization.

To tackle the challenge of the optimization with the long-term objective function and/or constraints with continuous variables, Lyapunov optimization has been utilized [21]–[24]. Since Lyapunov optimization does not rely on the state transition, its computational complexity is generally much smaller than the MDP-based approaches. In [22], the delay-aware resource control problem is studied with the Lyapunov optimization, where the system throughput, the sum delay and the power consumption are jointly optimized. He *et al.* [23] proposed to utilize the Lyapunov optimization to optimize the time average expected transmission cost for the mobile video streaming system by balancing the data wastage cost and the freezing time of the mobile user. By considering the playback interruption and fluctuation, Wu *et al.* [24] maximized the received video quality of an EH video communication system with the Lyapunov optimization.

In this paper, we consider the continuous energy management problem of a DF-based cooperative communication system with an EH relay. We propose to utilize the Lyapunov optimization theory to derive the optimal energy that the relay should use at each time slot. The major contributions are summarized as follows.

1) We formulate the energy management problem as an optimization problem, where the objective function is to minimize the long-term average SER with the constraint that the long-term average battery level is above a certain threshold. To solve the optimization, we introduce a virtual queue and employ the Lyapunov optimization theory to transform the optimization with long-term average format into optimizing the drift-plus-penalty problem. We also upper bound the drift-plus-penalty with all the terms that are only related to current time slot, which greatly simplifies the optimization problem.
2) We conduct the theoretic analysis of the optimal strategy derived by the proposed algorithm. Specifically, we first show the convexity of the drift-plus-penalty in terms of the transmit power. Then, we prove that the virtual queue is bounded and the battery constraint is satisfied with the energy management strategy derived by the proposed algorithm. We also show that the proposed optimal energy management strategy is limited by an upper bound that is independent of the operation time index. Finally, we analyze the diversity order and the corresponding EH gain of the proposed energy management strategy.
3) We run the numerical simulations to evaluate the average SER and the diversity order that can be achieved by the proposed method. In the simulations, we deploy the data of real solar irradiance measured by the solar site in Elizabeth City University. We show with the simulation results that the proposed method, compared with the MDP-based method in [11], can achieve much better performance in terms of both average SER and diversity order. We also evaluate with simulations the sensitivities of the specific parameters in Lyapunov optimization theorem to the average SER and diversity order.

The rest of this paper is organized as follows. In Section II, we describe in detail the system model and problem formulation. Then, we discuss how to utilize the Lyapunov optimization theory to solve the optimization problem in Section III. In Section IV, we conduct the performance analysis of the proposed algorithm, including the guarantee of the long-term battery constraint, the SER performance, and the diversity order and EH gain. Finally, simulation results are shown in Section V and the conclusions are drawn in Section VI.

A summary of major notations used in this paper is shown in Table I.

II. System Model and Problem Formulation

In this paper, we consider a cooperative communication system with an energy harvester on the cooperative relay as shown in Fig. 1. In this system, there is a source node (S), an EH relay node (R), and a destination node (D). The channels between source to destination (SD), source to relay (SR), and relay to destination (RD) are denoted as $\zeta_{\rm sd}$, $\zeta_{\rm sr}$, and $\zeta_{\rm rd}$, respectively. These channels are assumed to be independent Rayleigh fading with the mean power $\eta_{\rm sd}$, $\eta_{\rm sr}$, and $\eta_{\rm rd}$, respectively. The M-ary phase-shift-keying (M-PSK) modulation is used in the SR, RD, and SD communications.

TABLE I
BRIEF SUMMARY OF MAJOR NOTATIONS

T_L: Time duration of one phase	T_M: Policy management duration
ζ : Instantaneous channel power	E_H: Harvested energy during T_M
η : Average channel power	$\tilde{x}$: Decoded symbol at destination
Ψ_s: Source transmission power	γ_T : Decoding threshold
ω : Relay transmission energy	M : M-PSK modulation scheme
p : Relay transmission power	c_M : M-PSK modulation specific parameter
B : Virtual queue for the additive constraint	V : Parameter of penalty weight
L : Lyapunov function	d : Diversity order
$\mathbf{s}$: Set of the random variables	g_E : Energy harvesting gain
m : Decoding state	δ : Long-term average battery constraint
ρ : Probability of the almost silent relay when $m = 1$	N_0 : Average noise power

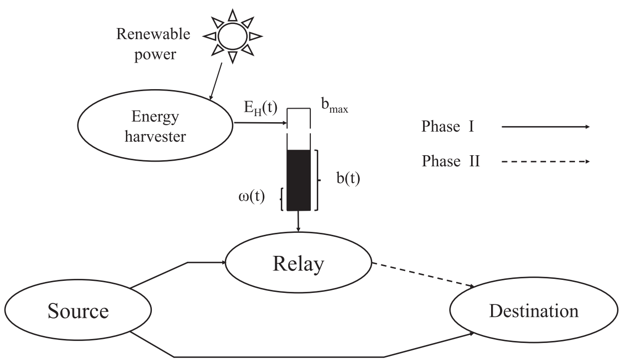


Fig. 1. Illustration of the system model.

The entire communication procedures are divided into two phases, each of which is with duration T_L. In phase I, the source node broadcasts the signals to the relay node and the destination node. The transmission power of the source node is fixed as Ψ_s, and thus the instantaneous received SNR of the SR link and SD link is $[(\Psi_s\zeta_{\text{sr}})/N_0]$ and $[(\Psi_s\zeta_{\text{sd}})/N_0]$, respectively. In phase II, the relay node may forward the signals received in phase I to the destination node. Here, we assume that the selective DF strategy is used in phase II, i.e., if the decoding is successful at the relay node in phase I, the relay node will adopt energy in the battery to forward the re-encoded modulated symbols to the destination node. Note that the harvester collects energy in both two phases but only consumes in phase II.

We assume that the EH system adopts an HUS strategy [4], which means that the relay node can only use the energy in the battery and the harvested energy in current time slot can be only available in the subsequent time slots. Thus, the energy in the battery can be updated as follows:

$$b(t+1) = \min\{b(t) - \omega(t) + E_H(t), b_{\max}\} \tag{1}$$

where $\omega(t) \in [0, b(t)]$ is the energy consumed at time slot t which cannot exceed the energy in the battery $b(t)$ at time slot t, $E_H(t)$ is the energy harvested at time slot t, and $b_{\max}$ is the battery capacity. We assume that $\forall t, E_H(t) > 0$ because our EH networks only work with active harvestable energy. The relay transmission power in each time slot t, $p(t)$, can be written as

$$p(t) = \frac{\omega(t)}{T_L}. \tag{2}$$

With the selective DF strategy, the relay will forward the signals only when the decoding is successful. Let $m = 1$ and $m = 0$ stand for the decoding is successful and failure, respectively. In general, the decoding status is determined by the received SNR of the SR link, $[(\Psi_s\zeta_{\text{sr}})/N_0]$. If the received SNR is larger than a predefined threshold γ_T, i.e., $[(\Psi_s\zeta_{\text{sr}})/N_0] > \gamma_T$, then the decoding is successful $m = 1$. Given Ψ_s, N_0 and γ_T, the probability of successful decoding at the relay node can be written as follows:

$$p(m) = \begin{cases} P\left(\zeta_{\text{sr}} \geq \frac{\gamma_T N_0}{\Psi_s}\right) = \exp\left(-\frac{\gamma_T N_0}{\Psi_s \eta \text{sr}}\right), & m = 1 \\ P\left(\zeta_{\text{sr}} < \frac{\gamma_T N_0}{\Psi_s}\right) = 1 - \exp\left(-\frac{\gamma_T N_0}{\Psi_s \eta \text{sr}}\right), & m = 0. \end{cases} \tag{3}$$

The performance of the cooperative communications is determined by multiple factors including the expected power of SD and RD channel, the battery energy, the harvested energy, and the decoding status at the relay node. We define a set $\mathbf{s}$ to capture these factors as

$$\mathbf{s} = (\zeta_{\text{sd}}, \zeta_{\text{rd}}, b, E_H, m). \tag{4}$$

According to [25], the expression of the expected SER can be written as

$$P\left(x \neq \widetilde{x} \mid \mathbf{s}\right) = \begin{cases} \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \exp\left(-\frac{c_M(\Psi_s\zeta_{\text{sd}} + \omega(t)\zeta_{\text{rd}}/T_L)}{N_0 \sin^2\theta}\right) d\theta \\ \qquad m = 1 \\ \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \exp\left(-\frac{c_M \Psi_s \zeta_{\text{sd}}}{N_0 \sin^2\theta}\right) d\theta \\ \qquad m = 0. \end{cases} \tag{5}$$

Our goal is to help the relay determine the energy to spend at every time slot when the decoding is successful. To do so, we use the long-term average SER as the objective function as follows:

$$\lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} P\left(x \neq \widetilde{x} \mid \mathbf{s}\right). \tag{6}$$

Due to the stochastic nature of the wireless channel, there exists a non-negligible probability that the power of the RD channel is small. Obviously, forwarding signals in such a condition will bring little payback, and storing the energy and waiting for a better chance may be a better choice. Therefore, we include a long-term average battery buffer constraint to help the battery store power for further use when the channel condition is poor, and the constraint can be written as follows:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1} b(t) \geq \delta \tag{7}$$

where δ is a predefined energy level threshold.

Therefore, the problem can be formulated as to choose the optimal transmission power $\omega(t)$ at each time slot such that the long-term average SER is minimized while the battery constraint is satisfied as follows:

$$\begin{aligned} \min_{\omega(t)} \quad & \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1} P\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big) \\ \text{s.t.} \quad & \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1} b(t) \geq \delta \\ & \omega(t) \in [0, b(t)]. \end{aligned} \tag{8}$$

III. Stochastic Relay Transmission Policy Based on Lyapunov Optimization

Both the objective function and constraint in the optimization problem in (8) involve the long-term averaging, due to which the solution is difficult to find. To solve the problem, in this section we propose to apply virtual queue and the penalty drift in the Lyapunov optimization theory.

We first transform the long-term averaging form of the constraint into a virtual queue. Specifically, we define a virtual queue B as follows:

$$B(t+1) = \max\{B(t) + \delta - b(t+1), 0\}. \tag{9}$$

From (9), we can see that if at a specific time slot the battery energy is less than the threshold, the virtual queue B increases, i.e., the virtual queue B reflects how well the constraint is satisfied. To utilize the penalty drift in Lyapunov optimization theory, we define $\Theta(t) = B(t)$, and the corresponding Lyapunov function can be written as

$$L(\Theta(t)) = \frac{1}{2}B^2(t). \tag{10}$$

Based on [21], the penalty drift can be written as follows:

$$\Delta\Theta(t) = \mathbb{E}[L(\Theta(t+1)) - L(\Theta(t))|\Theta(t)] \tag{11}$$

and the long-term averaging terms in (8) can be rewritten as "drift-plus-penalty" as

$$\Delta\Theta(t) + V\mathbb{E}\Big[P\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big)|\Theta(t)\Big] \tag{12}$$

where parameter $V \geq 0$ is the penalty weight, which reflects the importance of the objective function comparing to the constraint at each time slot.

Therefore, the optimization problem in (8) can be rewritten as

$$\begin{aligned} \min_{\omega(t)} \quad & \Delta\Theta(t) + V\mathbb{E}\Big[P\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big)|\Theta(t)\Big] \\ \text{s.t.} \quad & 0 \leq \omega(t) \leq b(t). \end{aligned} \tag{13}$$

According to (11), we can see that the objective function in (13) involves the variables at time slot $t+1$, which can be simplified by finding the upper bound of $B^2(t+1) - B^2(t)$ as follows:

$$\begin{aligned} & B^2(t+1) - B^2(t) \\ &\quad = \{\max\{B(t) + \delta - b(t+1), 0\}\}^2 - B^2(t) \\ &\quad \leq [B(t) + \delta - b(t+1)]^2 - B^2(t) \\ &\quad \leq \delta^2 + b^2(t+1) + 2B(t)[\delta - b(t+1)] \\ &\quad = \delta^2 + \{\min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}\}^2 \\ &\qquad + 2B(t)[\delta - b(t+1)] \\ &\quad \leq \delta^2 + \{\min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}\}^2 \\ &\qquad + 2B(t)[\delta - b(t) + \omega(t)] \\ &\quad \leq \delta^2 + \{\max\{b(t) - \omega(t), 0\} + E_H(t)\}^2 \\ &\qquad + 2B(t)[\delta - b(t) + \omega(t)] \\ &\quad = \delta^2 + \{\max\{b(t) - \omega(t), 0\}\}^2 + E_H(t)^2 + 2EH(t) \\ &\qquad \times \max\{b(t) - \omega(t), 0\} + 2B(t)[\theta - b(t) + \omega(t)] \\ &\quad \leq \delta^2 + [b(t) - \omega(t)]^2 + E_H(t)^2 + 2E_H(t) \\ &\qquad \times b(t) + 2B(t)[\delta - b(t) + \omega(t)] \\ &\quad \leq \delta^2 + E_{H\max}^2 + b_{\max}^2 + \omega(t)^2 - 2b(t)\omega(t) \\ &\qquad + 2b_{\max} \times E_{H\max} + 2B(t)[\delta - b(t) + \omega(t)] \end{aligned} \tag{14}$$

where the first inequality is due to $\{\max\{a, 0\}\}^2 \leq a^2$, the third inequality is due to $b(t+1) = \min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}$ and $\max\{b(t) - \omega(t), 0\} + E_H(t) \geq b(t) - \omega(t)$, the fourth inequality is due to $\{\min\{a, b\}\}^2 \leq a^2$.

With (10), (11), and (14), the objective function in (13) can be upper bounded as

$$\begin{aligned} & \Delta\Theta(t) + VP\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big) \\ &\quad \leq \frac{1}{2} \times \Big\{\delta^2 + E_{H\max}^2 + b_{\max}^2 + \omega^2(t) - 2b(t)\omega(t) \\ &\qquad\qquad + 2b_{\max}E_{H\max} + 2B(t) \times [\delta - b(t) + \omega(t)]\Big\} \\ &\qquad + VP\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big) \\ &\quad = \frac{1}{2} \times \Big\{B_1 + B_2 + \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t)\Big\} \\ &\qquad + VP\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big) \end{aligned} \tag{15}$$

where $B_1 = \delta^2 + E_{H\max}^2 + b_{\max}^2 + 2b_{\max}E_{H\max}$ is a constant during all time slots and $B_2 = 2B(t) \times [\delta - b(t)]$ is fixed at a specific time slot. Therefore, by ignoring the terms that are not related to $\omega(t)$, the optimization problem (13) can be approximated as

$$\begin{aligned} \min_{\omega(t)} \quad & \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t) + 2VP\Big(x \neq \widetilde{x} \mid \mathbf{s}\Big) \\ \text{s.t.} \quad & 0 \leq \omega(t) \leq b(t). \end{aligned} \tag{16}$$

IV. Performance Analysis

In this section, we will conduct performance analysis of the proposed stochastic relay transmission strategy.

A. Battery Constraint

We first show in Lemma 1 that the objective function in (16) is convex in terms of $\omega(t)$ when the decoding is successful, i.e., $m = 1$.

Lemma 1: The objective function in (16) is convex in terms of $\omega(t)$ when $m = 1$.

Proof: Please see Appendix A for details. ■

Then, we would like to show that the solution derived by (16) satisfies the long-term average battery buffer constraint in (7). To do so, we first derive the following Lemma 2.

Lemma 2: The constraint $\lim_{T\to+\infty}(1/T)\sum_{t=0}^{T-1} b(t) \geq \delta$ is satisfied if the virtual queue in (9) is upper bounded.

Proof: Please see Appendix B for details. ■

With Lemma 2, we are ready to show that the long-term average battery energy constraint is satisfied with the following Theorem.

Theorem 1: The long-term battery energy constraint is satisfied with the method based on Lyapunov optimization applied.

Proof: According to Lemma 1, we know that the objective function $J(\omega(t))$ in (29) is convex, which means that the first-order derivative $[(dJ(\omega(t)))/(d\omega(t))]$ in (30) is increasing. Then, we evaluate the value of the first-order derivative at zero, i.e., $[(dJ(\omega(t)))/(d\omega(t))]|_{\omega(t)=0}$.

Case 1: When $[(dJ(\omega(t)))/(d\omega(t))]|_{\omega(t)=0} < 0$, since $[(dJ(\omega(t)))/(d\omega(t))]$ is increasing, the optimal transmit power $\omega^*(t) > 0$ and $[(dJ(\omega(t)))/(d\omega(t))]|_{\omega(t)=\omega^*(t)} \leq 0$. In such a case, according to (30), we have

$$\begin{aligned} B(t) &\leq b(t) - \omega^*(t) + V\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\left(\frac{c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right) \\ &\quad\times \exp\left\{-\frac{c_M\left[\Psi_s\zeta_{\mathrm{sd}} + (\omega^*(t)/T_L)\zeta_{\mathrm{rd}}\right]}{N_0\sin^2\theta}\right\}d\theta \\ &\leq b(t) + V\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\left(\frac{c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right) \\ &\quad\times \exp\left(-\frac{c_M\Psi_s\zeta_{\mathrm{sd}}}{N_0\sin^2\theta}\right)d\theta. \end{aligned} \tag{17}$$

Case 2: When $[(dJ(\omega(t)))/(d\omega(t))]|_{\omega(t)=0} \geq 0$, since $[(dJ(\omega(t)))/(d\omega(t))]$ is increasing, the optimal transmit power $\omega^*(t) = 0$. According to (1), we have

$$b(t+1) = \min\{b(t) + E_H(t), b_{\max}\}. \tag{18}$$

Suppose that there exists a t_0 such that

$$\begin{aligned} B(t_0) &\leq b(t_0) + V\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\left(\frac{c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right) \\ &\quad\times \exp\left(-\frac{c_M\Psi_s\zeta_{\mathrm{sd}}}{N_0\sin^2\theta}\right)d\theta. \end{aligned} \tag{19}$$

Then, we would like to show that $\forall t > t_0$, we have

$$\begin{aligned} B(t) &\leq b(t) + V\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\left(\frac{c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right) \\ &\quad\times \exp\left(-\frac{c_M\Psi_s\zeta_{\mathrm{sd}}}{N_0\sin^2\theta}\right)d\theta + T_0\delta \end{aligned} \tag{20}$$

where $T_0 = [\delta/(E_{H\min})]$ is a fixed constant.

According to (9), we can see that if $b(t+1) < \delta$, we have

$$B(t) < B(t+1) < B(t) + \delta. \tag{21}$$

With (18), we know that with at most $T_0 = [\delta/(E_{H\min})]$ time slots, $b(t+1) > \delta$, in which case $\omega^*(t) = 0$ and $B(t)$ satisfies (17). Therefore, (20) holds $\forall t > t_0$.

In all, with (17) and (20), $B(t)$ is upper bounded. According to Lemma 2, the constraint of the battery energy is satisfied.

This completes the proof. ■

From Theorem 1, we can see that the long-term average battery energy constraint is guaranteed with the proposed stochastic strategy.

B. SER Bound

In this section, we derive an upper bound of the expected SER when applying the proposed stochastic strategy, as shown in the following theorem.

Theorem 2: The average SER of the optimal strategy obtained by (16) is limited by an upper bound that is independent of the operation time, i.e.,

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[P\left(x \neq \tilde{x} \mid \mathbf{s}\right)\right] \leq P^{\mathrm{opt}} + \frac{B_1 + 2b_{\max}B_{\max}}{2V} \tag{22}$$

where $B_{\max}$ is the upper bound of the virtual queue B.

Proof: Please see Appendix C for details. ■

C. Diversity Order and Energy Harvesting Gain

With the DF protocol, the relay node works only when the decoding is successful, i.e., $m = 1$. In other words, if the decoding fails, i.e., $m = 0$, the SER is purely determined by the source-to-destination link. Therefore, the SER performance has a lower bound when the SNRs of the SD link and SR link are given.

Theorem 3: When the η_{sr} and η_{sd} are given, the expected SER is lower bounded as

$$\mathbb{E}\left[P\left(x \neq \tilde{x} \mid \mathbf{s}\right)\right] \geq \frac{\left(1 - \exp\left(-\frac{\gamma_T N_0}{\Psi_s\eta_{\mathrm{sr}}}\right)\right)}{\left(1 + \frac{c_M\Psi_s\eta_{\mathrm{sd}}}{N_0\sin^2\theta}\right)} \tag{23}$$

and the lower bound is achieved when $\delta \leq E_{H\min}$ and η_{rd} goes to infinite.

Proof: Please see Appendix D for details. ■

From Theorem 3, we can see that when η_{rd} goes to infinite, the expected SER is determined by the source-to-relay link and the source-to-destination link, i.e., η_{sr} and η_{sd}.

Then, we will discuss the diversity order of this EH cooperative wireless network. To do so, we first derive the expression of the asymptotical average SER in the following theorem.

Theorem 4: The asymptotical average SER can be expressed as

$$\begin{aligned} P_{M,\mathrm{asym}} &= \rho \times \frac{K_0}{c_M\eta_{\mathrm{sd}}}\left(\frac{\Psi_s}{N_0}\right)^{-1} \\ &\quad + \left(\frac{T_L K_1 c_{R,0}}{c_M^2\eta_{\mathrm{sd}}\eta_{\mathrm{rd}}} + \frac{\gamma_T K_0}{c_M\eta_{\mathrm{sr}}\eta_{\mathrm{sd}}}\right)\left(\frac{\Psi_s}{N_0}\right)^{-2} \\ &\quad - \frac{\gamma_T T_L K_1 c_{R,0}}{c_M^2\eta_{\mathrm{sd}}\eta_{\mathrm{rd}}\eta_{\mathrm{sr}}}\left(\frac{\Psi_s}{N_0}\right)^{-3} \end{aligned} \tag{24}$$

where $\rho = P\{\omega = 0|m = 1\}$, $c_{R,0}$, K_0 and K_1 are defined as follows:

$$c_{R,0} = \mathbb{E}\left(\frac{\Psi_s}{\omega}|\omega > 0\right)$$
$$= \Psi_s\left[\int_{0^+}^{b_{\max}} \omega_0^{-1} P(\omega = \omega_0|m = 1)d\omega_0\right] \tag{25}$$
$$K_0 = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \sin^2\theta d\theta = \frac{M-1}{2M} + \frac{\sin(2\pi/M)}{4\pi} \tag{26}$$

and

$$K_1 = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \sin^4\theta d\theta$$
$$= \frac{3(M-1)}{8M} + \frac{\sin(2\pi/M)}{4\pi} - \frac{\sin(4\pi/M)}{32\pi}. \tag{27}$$

Proof: Please see Appendix E for details. ■

Theorem 4 presents the asymptotical average SER when the average channel-to-noise power ratio of the three channels (SD, SR, and RD links) are all sufficiently high. From Theorem 4, we can tell that when $\rho \neq 0$ the diversity order $d = 1$ and the EH gain $g_E = [(c_M\eta_{\rm sd})/(K_0 P\{\omega = 0|m = 1\})]$, while when $\rho = 0$ the diversity order $d = 2$ and the EH gain $g_E = [(c_M^2\eta_{\rm sd}\eta_{\rm sr}\eta_{\rm rd})/(T_L K_1 c_{R,0}\eta_{\rm sr} + \gamma_T K_0 c_M \eta_{\rm rd})]$. That is to say the diversity order is determined by the status of the relay. Once there exists a positive probability that the relay does not work, the diversity order will decrease to 1.

Then we discuss the diversity order and the corresponding EH gain under the special cases when $\delta \leq E_{H\min}$ and $\delta > b_{\max}$ in the following Corollaries 1 and 2, respectively.

Corollary 1: When $\delta \leq E_{H\min}$, the diversity order $d = 2$ and the EH gain is bounded by $[(c_M^2\eta_{\rm sd}\eta_{\rm rd}\eta_{\rm sr}c_{R,2})/(T_L K_1\eta_{\rm sr} + \gamma_T K_0 c_M c_{R,2}\eta_{\rm rd})] \leq g_E \leq [(c_M^2\eta_{\rm sd}\eta_{\rm rd}\eta_{\rm sr}c_{R,1})/(T_L K_1\eta_{\rm sr} + \gamma_T K_0 c_M c_{R,1}\eta_{\rm rd})]$ with $c_{R,1} = [(b_{\max})/(\Psi_s)]$ and $c_{R,2} = [(E_{H\min})/(\Psi_s)]$.

Proof: Please see Appendix F for details. ■

On the contrary, the case where $\delta > b_{\max}$ is discussed as follows.

Corollary 2: When $\delta > b_{\max}$, the diversity order $d = 1$ and the EH gain $g_E = [(c_M\eta_{\rm sd})/K_0]$.

Proof: When $\delta > b_{\max}$, we have

$$B(t+1) = \max\{B(t) + \delta - b(t+1), 0\}$$
$$\geq \max\{B(t) + \delta - b_{\max}, 0\}$$
$$> B(t). \tag{28}$$

In such a case, there exists a t_0 such that $\forall t > t_0$, $[(dJ(\omega(t)))/(d\omega(t))]|_{\omega(t)=0} > 0$, which leads to $\omega^*(t) = 0$ and thus $\rho = 1$. According to Theorem 4, the diversity order $d = 1$ and the EH gain $g_E = [(c_M\eta_{\rm sd})/K_0]$.

This completes the proof. ■

V. Simulation Results

In this section, we conduct multiple simulations to evaluate the performance of the proposed scheme. The simulation settings are shown as follows. The policy management duration of the cooperative communications T_M is set as 300 s, which contains two transmission phases T_L. In other words, in one management period an entire procedure of cooperative communication is completed. The power of the transmission source is given as $\Psi_s = 4 \times 10^4$ μW. In the simulation, we deploy the real data records of the solar irradiance measured in solar site in Elizabeth City State University in June from 2010 to 2012 [26]. The solar panel size is set as 4 cm^2 and the harvesting efficiency is 20%. Considering the common location profile among source, relay and destination, the channel-to-noise ratio of the SR and RD channel are set to be 5 dB larger than that of the SD channel. By assuming that the sensor nodes are located in a rich scattering environment, we apply Jakes' model to generate channel information under a deterministic relative mobility among the source, the relay and the destination [27]. The Doppler frequency is set to be $f_D = 5 \times 10^{-2}$ to guarantee a slow variation of the wireless channel. The decoding threshold γ_T of the relay is set as 15 dB and the modulation scheme is set to be $M = 4$ (QPSK).

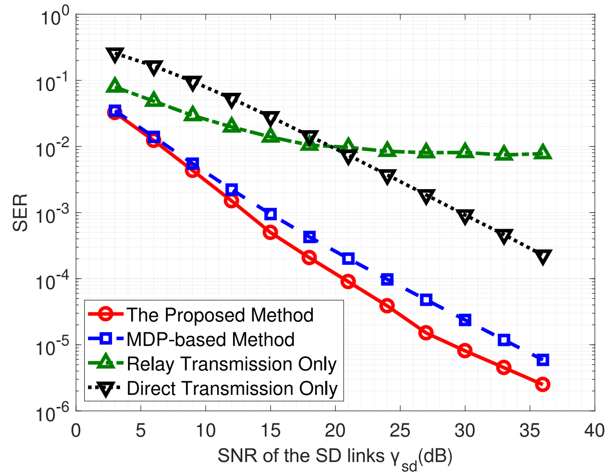


Fig. 2. Performance comparison of the average SER.

We compare the proposed scheme with the MDP-based method in [11]. Besides the parameters set above, the additional parameters used in the MDP-based method are set as below. The channel quantization thresholds for the RD and SD channel are $\{0, 2.0, 3.0, +\infty\}$. Since the MDP-based method cannot deal with the continuous energy, the number of the battery energy state is 8, which is the same level as the source transmit energy $4 \times 10^4 \times T_L$ μJ, and the number of solar state is 4. The discount factor in the MDP-based method is set as 0.99.

A. Average SER and Diversity Order Comparison

In Fig. 2, we show the average SER performance comparison between the proposed strategy and the MDP-based method under various SNR value $\Upsilon_{\rm sd} = [(\Psi_s\eta_{\rm sd})/N_0]$, where we also include the performance comparison with the direct transmission without relay link (denoted as "direct transmission only") and relay transmission without direct link (denoted as "relay transmission only"). The $\Upsilon_{\rm sr} = [(\Psi_s\eta_{\rm sr})/N_0]$ is set to be 35 dB. From the figure, we can see that with cooperative communication, both the proposed strategy and the MDP-based

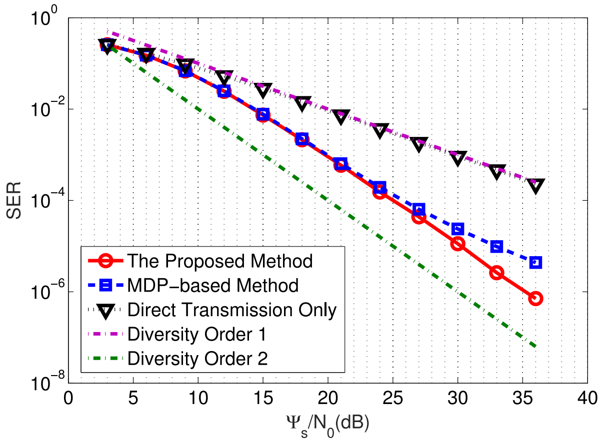


Fig. 3. Performance comparison of the diversity order.

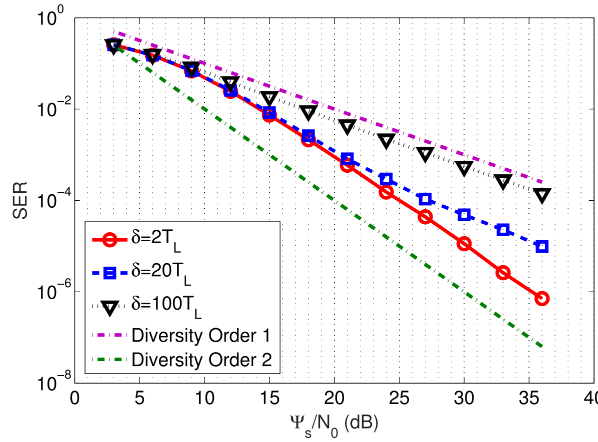


Fig. 4. Sensitivity analysis of the parameter δ with $V = 100$.

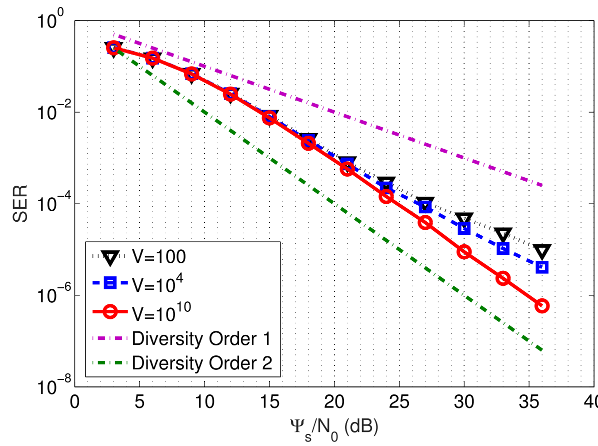


Fig. 5. Sensitivity analysis of the parameter V with $\delta = 20T_L$.

strategy achieve better performance than the direct transmission only and the relay transmission only. The proposed strategy based on the Lyapunov optimization can achieve much better performance compared with the MDP-based strategy, especially when $\Upsilon_{\rm sd}$ is larger than 15 dB. This is mainly because at the high SNR region, the continuous energy management have more advantage than the discrete one. We also observe from the figure that when the $\Upsilon_{\rm sd}$ increases, the SER performance of the relay transmission only saturates. This is because the $\Upsilon_{\rm sr}$ is fixed, leading to that the decoding successfully probability at the relay is smaller than 1, and according to Theorem 3, the SER performance is lower bounded.

Then, in Fig. 3, we evaluate the diversity order achieved by the proposed method and MDP-based method, where we vary the SNR $\Upsilon_{\rm sd}$ from 3 dB to 36 dB. From the figure, we can see that the blue line becomes parallel to the purple line as the SNR increases, i.e., the MDP-based method can only achieve diversity order 1, which is the same as the direct transmission only. Note that the EH gain of MDP-based method is much better than that of direct transmission only since for any fixed SER, the SNR of the blue line is much smaller than that of the purple line. With the Lyapunov optimization, the proposed method handles the continuous energy management and can achieve the diversity order 2 since the red line becomes parallel to the green line as the SNR increases. The gap between the MDP-based method and the proposed method becomes larger when (Ψ_s/N_0) is larger than 25 dB.

B. Parameter Sensitivity Analysis

From (16), we can see that the performance of the proposed method may be affected by the parameters δ and V. In this section, we will analyze the sensitivity of the performance of the proposed method with δ and V.

In Fig. 4, we evaluate the influence of the δ on the diversity order and the average SER under different values of (Ψ_s/N_0). The channel condition is the same as that in Fig. 3 and the parameter V is set as 100. From the figure, we can see that the diversity order is 1 when $\delta = 20T_L$ or $\delta = 100T_L$ since the blue line and black line are both parallel to the purple line in the high SNR region. However, the EH gain with $\delta = 20T_L$ is much better than that with $\delta = 100T_L$ since the SNR of the blue line to achieve any fixed SER is much smaller than that of the black line. When δ further reduces to $\delta = 2T_L$, the diversity order becomes 2 as the red line is parallel to the green line in the high SNR region. Therefore, the diversity order can be improved with a smaller δ using the proposed algorithm. Such a phenomenon is partially because with a smaller δ, more energy will be used to transmit the data, which reduces the probability ρ. When δ is small enough, the probability ρ can be zero, which leads to the diversity order 2.

In the last simulation, we shows in Fig. 5 the average SER and the diversity order versus (Ψ_s/N_0) under different values of V. The channel condition is the same as that in Fig. 3 and the δ is set as $20T_L$. From the figure, we can see that with the performance is improved as V becomes larger, where the diversity order is 1 when V is 100 or 10^4, and is 2 when V is 10^{10}.

VI. Conclusion

In this paper, we studied the continuous energy management problem in an EH cooperative wireless communication system with DF relaying, where the objective is to minimize the long-term average SER subject to a battery constraint. We employed the Lyapunov optimization theory to transform the optimization problem into a drift-plus-penalty minimization problem, which was proved to be convex. We also showed that the battery constraint is guaranteed, and the proposed optimal energy management strategy is limited by an upper bound. The closed-form expression of the asymptotical average SER is derived, according to which the diversity order and the EH gain of the proposed strategy were analyzed. Finally, simulation results using the real solar irradiance data showed that, compared with the MDP-based method, the proposed scheme achieved much better performance in terms of both average SER and diversity order.

APPENDIX A
PROOF OF LEMMA 1

When $m = 1$, substituting (5) into (16), the objective function in (16) can be written as

$$J(\omega(t)) = \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t) + \frac{2V}{\pi}\int_0^{\frac{M-1}{M}\pi} \exp\left\{-\frac{c_M\left[\Psi_s\zeta_{\mathrm{sd}} + (\omega/T_L)\zeta_{\mathrm{rd}}\right]}{N_0\sin^2\theta}\right\}d\theta. \quad (29)$$

By taking the first derivative of $J(\omega(t))$ with respect to $\omega(t)$, we have

$$\frac{dJ(\omega(t))}{d\omega(t)} = 2\omega(t) - 2b(t) + 2B(t) + \frac{2V}{\pi}\int_0^{\frac{M-1}{M}\pi} \times \left(\frac{-c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right)\exp\left\{-\frac{c_M\left[\Psi_s\zeta_{\mathrm{sd}} + (\omega/T_L)\zeta_{\mathrm{rd}}\right]}{N_0\sin^2\theta}\right\}d\theta. \quad (30)$$

Similarly, the second derivative of $J(\omega(t))$ with respect to $\omega(t)$ can be derived as

$$\frac{d^2J(\omega(t))}{d\omega(t)^2} = 2 + 2V\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\left(\frac{c_M\zeta_{\mathrm{rd}}}{T_L N_0\sin^2\theta}\right)^2 \times \exp\left\{-\frac{c_M\left[\Psi_s\zeta_{\mathrm{sd}} + (\omega(t)/T_L)\zeta_{\mathrm{rd}}\right]}{N_0\sin^2\theta}\right\}d\theta \quad (31)$$

which is positive for all $\omega(t)$. Therefore, the objective function is convex with respect to the variable $\omega(t)$. This completes the proof.

APPENDIX B
PROOF OF LEMMA 2

Suppose that the virtual queue B is upper bounded by $B_{\max} > 0$. Then, we have

$$\lim_{t\to+\infty}\frac{B(t)}{t} \le \lim_{t\to+\infty}\frac{B_{\max}}{t} = 0. \quad (32)$$

The (9) can be rewritten as follows:

$$B(t+1) = \begin{cases} B(t) + \delta - b(t+1), & \text{if } B(t) \ge b(t+1) - \delta \\ 0, & \text{if } B(t) < b(t+1) - \delta. \end{cases} \quad (33)$$

With (33), we have

$$\begin{aligned} B(t+1) - B(t) &= \begin{cases} \delta - b(t+1), & \text{if } B(t) \ge b(t+1) - \delta \\ -B(t), & \text{if } B(t) < b(t+1) - \delta \end{cases} \\ &= \max\{\delta - b(t+1), -B(t)\} \\ &\ge \delta - b(t+1). \end{aligned} \quad (34)$$

By taking the average from time index 0 to time index $T-1$ on both sides of (34) and letting T goes to infinite, we have

$$\begin{aligned} \lim_{T\to+\infty}\frac{B(T)}{T} &\ge \delta - \lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T} b(t) \\ &= \delta - \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1} b(t). \end{aligned} \quad (35)$$

With (32) and (35), we have

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1} b(t) \le \delta. \quad (36)$$

This completes the proof.

APPENDIX C
PROOF OF THEOREM 2

By substituting the expression of B_2, (15) can be rewritten as

$$\begin{aligned} &\Delta\Theta(t) + V\mathbb{E}\left[P\left(x \ne \tilde{x} \mid \mathbf{s}\right)\right] \\ &\le \frac{1}{2}\Big\{B_1 + 2B_{\max}\mathbb{E}\{[\delta - b(t)]|\Theta(t)\} + \omega(t)^2 - 2b(t)\omega(t) + 2B_{\max}\times\omega(t)\Big\} \\ &\quad + V\mathbb{E}\left[P\left(x \ne \tilde{x} \mid \mathbf{s}\right)\right]. \end{aligned} \quad (37)$$

Since "$\delta - b(t)$" is a factor of the inequality constraint in (8), with ω-only policy [21], (37) can be simplified as

$$\begin{aligned} &\Delta\Theta(t) + V\mathbb{E}\left[P\left(x \ne \tilde{x} \mid \mathbf{s}\right)\right] \\ &\le \frac{1}{2}\Big\{B_1 + \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t)\Big\} + VP^{\mathrm{opt}}. \end{aligned} \quad (38)$$

With some manipulations, (38) can be rearranged as follows:

$$\begin{aligned} &\mathbb{E}\left[P\left(x \ne \tilde{x} \mid \mathbf{s}\right)\right] \\ &\le P^{\mathrm{opt}} + \frac{-2\Delta\Theta(t) + B_1 + \omega(t)[+\omega(t) - 2b(t) + 2B(t)]}{2V} \\ &\le \begin{cases} P^{\mathrm{opt}} + \frac{-2\Delta\Theta(t) + B_1 - b^2(t) + 2b(t)B(t)}{2V}, & B(t) \ge \frac{1}{2}b(t) \\ P^{\mathrm{opt}} + \frac{-2\Delta\Theta(t) + B_1}{2V}, & \text{others} \end{cases} \\ &\le P^{\mathrm{opt}} + \frac{-2\Delta\Theta(t) + B_1 + 2b_{\max}B_{\max}}{2V}. \end{aligned} \quad (39)$$

By taking the average from time index 0 to time index $T-1$ on both sides of (39) and letting T goes to infinite, we have

$$\begin{aligned}
&\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid\mathbf{s}\Big)\Big]\\
&\leq P^{\text{opt}}+\frac{B_1+2b_{\max}B_{\max}}{2V}-\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\frac{\Delta\Theta(t)}{V}\\
&=P^{\text{opt}}+\frac{B_1+2b_{\max}B_{\max}}{2V}-\lim_{T\to+\infty}\frac{L(\Theta(T))-L(\Theta(0))}{VT}\\
&=P^{\text{opt}}+\frac{B_1+2b_{\max}B_{\max}}{2V}.
\end{aligned}\tag{40}$$

Appendix D
Proof of Theorem 3

We first prove that when $\delta\leq E_{H\min}$ and η_{rd} goes to infinite, the expected SER with the proposed algorithm goes to zero if the decoding at the relay node is successful.

According to (1) and (9), when $\delta\leq E_{H\min}$, we have $B(t+1)\leq\max\{0,B(t)\}$. Therefore, we have $B(t)=0,\forall t$. In such a case, according to (30), we have $[(dJ(\omega(t)))/(d\omega(t))]<0,\forall t$. Due to the convexity of $J(\omega(t))$, the optimal power is $\omega^*(t)=b(t)=E_H(t-1),\forall t$.

With (5), the expected SER when $m=1$ can be derived as follows:

$$\begin{aligned}
&\lim_{\eta_{\text{rd}}\to+\infty}\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=1\Big)\Big]\\
&=\lim_{\eta_{\text{rd}}\to+\infty}\int_{\sigma}^{b_{\max}}\int_0^{+\infty}\int_0^{\frac{M-1}{M}\pi}\frac{1}{\pi}\\
&\quad\times\exp\left\{-\frac{c_M\big[\Psi_s\zeta_{\text{sd}}+(E_H/T_L)\zeta_{\text{rd}}\big]}{N_0\sin^2\theta}\right\}\\
&\quad p(\zeta_{\text{rd}})p(E_H)d\zeta_{\text{rd}}dE_{\text{Hd}}\theta\\
&\quad+\int_0^{+\infty}\int_0^{\frac{M-1}{M}\pi}\frac{1}{\pi}\exp\left\{-\frac{c_M\big[\Psi_s\zeta_{\text{sd}}+(b_{\max}/T_L)\zeta_{\text{rd}}\big]}{N_0\sin^2\theta}\right\}\\
&\quad P(E_H>b_{\max})p(\zeta_{\text{rd}})d\zeta_{\text{rd}}d\theta\\
&=\lim_{\eta_{\text{rd}}\to+\infty}\int_{\sigma}^{b_{\max}}\int_0^{+\infty}\int_0^{\frac{M-1}{M}\pi}\frac{1}{\pi}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)\\
&\quad\times\exp\left(-\frac{c_ME_H\zeta_{\text{rd}}}{T_LN_0\sin^2\theta}\right)\frac{1}{\eta_{\text{rd}}}\exp\left(-\frac{\zeta_{\text{rd}}}{\eta_{\text{rd}}}\right)p(E_H)d\zeta_{\text{rd}}dE_{\text{Hd}}\theta\\
&\quad+\int_0^{+\infty}\int_0^{\frac{M-1}{M}\pi}\frac{1}{\pi}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)P(E_H>b_{\max})\\
&\quad\times\exp\left(-\frac{c_Mb_{\max}\zeta_{\text{rd}}}{T_LN_0\sin^2\theta}\right)\frac{1}{\eta_{\text{rd}}}\exp\left(-\frac{\zeta_{\text{rd}}}{\eta_{\text{rd}}}\right)d\zeta_{\text{rd}}d\theta\\
&=\lim_{\eta_{\text{rd}}\to+\infty}\int_{\sigma}^{b_{\max}}\int_0^{\frac{M-1}{M}\pi}\frac{1}{\left(1+\frac{\eta_{\text{rd}}E_H}{T_LN_0\sin^2\theta}\right)}p(E_H)dE_H\\
&\quad\times\frac{1}{\pi}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)d\theta+\int_0^{\frac{M-1}{M}\pi}\frac{1}{\left(1+\frac{\eta_{\text{rd}}b_{\max}}{T_LN_0\sin^2\theta}\right)}\\
&\quad\times P(E_H>b_{\max})\frac{1}{\pi}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)d\theta.
\end{aligned}\tag{41}$$

Since $E_H>0$, using $(1/x+1)\approx(1/x)$ when $x\gg1$, (41) can be approximated as

$$\begin{aligned}
&\lim_{\eta_{\text{rd}}\to+\infty}\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=1\Big)\Big]\\
&\approx\lim_{\eta_{\text{rd}}\to+\infty}\frac{T_LN_0\sin^2\theta}{\eta_{\text{rd}}}\\
&\quad\times\Bigg[\int_{\sigma}^{b_{\max}}\frac{1}{E_H}p(E_H)dE_H\int_0^{\frac{M-1}{M}\pi}\frac{1}{\pi}\\
&\quad\times\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)d\theta+\int_0^{\frac{M-1}{M}\pi}\frac{1}{b_{\max}}\\
&\quad\times P(E_H>b_{\max})\frac{1}{\pi}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)d\theta\Bigg]\\
&=0.
\end{aligned}\tag{42}$$

Then, the expected SER can be derived as follows:

$$\begin{aligned}
&\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid\mathbf{s}\Big)\Big]\\
&=\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=0\Big)\Big]P(m=0)\\
&\quad+\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=1\Big)\Big]P(m=1)\\
&\geq\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=0\Big)\Big]P(m=0)\\
&=\int_0^{\frac{M-1}{M}\pi}\int_0^{+\infty}\exp\left(-\frac{c_M\Psi_s\zeta_{\text{sd}}}{N_0\sin^2\theta}\right)\frac{1}{\eta_{\text{sd}}}\exp\left(-\frac{\zeta_{\text{sd}}}{\eta_{\text{sd}}}\right)d\theta\\
&\quad\times\left(1-\exp\left(-\frac{\gamma_TN_0}{\Psi_s\eta_{\text{sr}}}\right)\right)\\
&=\frac{\left(1-\exp\left(-\frac{\gamma_TN_0}{\Psi_s\eta_{\text{sr}}}\right)\right)}{\left(1+\frac{c_M\Psi_s\eta_{\text{sd}}}{N_0\sin^2\theta}\right)}
\end{aligned}\tag{43}$$

and the equality holds when $\mathbb{E}[P(x\neq x^{\sim}\mid m=1)]=0$, i.e., $\delta<E_{H\min}$ and η_{rd} goes to infinite according to (42).

Appendix E
Proof of Theorem 4

By applying that $e^x\approx1+x$ for $0\leq x\ll1$, the successful decoding probability can be expressed as $P(m=1)=P([(\zeta_{\text{sr}}\Psi_s)/N_0]\geq\gamma_T)=\exp(-[(\gamma_TN_0)/(\Psi_s\eta_{\text{sr}})])\approx1-[(\gamma_TN_0)/(\Psi_s\eta_{\text{sr}})]$, and the asymptotical average SER can be written as

$$\begin{aligned}
P_{M,\text{asym}}&=P(m=1)\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=1\Big)\Big]\\
&\quad+P(m=0)\mathbb{E}\Big[P\Big(x\neq\tilde{x}\mid m=0\Big)\Big]\\
&\approx\left(1-\frac{\gamma_TN_0}{\Psi_s\eta_{\text{sr}}}\right)\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\int_0^{+\infty}\int_0^{+\infty}\int_0^{b_{\max}}\\
&\quad\times\exp\left(-\frac{c_M(\Psi_s\zeta_{\text{sd}}+(\omega_0/T_L)\zeta_{\text{rd}})}{N_0\sin^2\theta}\right)\\
&\quad\times p(\zeta_{\text{rd}})p(\zeta_{\text{sd}})P\{\omega=\omega_0|m=1\}\\
&\quad\times d(\zeta_{\text{sd}})d\theta d(\zeta_{\text{rd}})d\omega_0\\
&\quad+\frac{\gamma_TN_0}{\Psi_s\eta_{\text{sr}}}\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\int_0^{+\infty}
\end{aligned}$$

$$
\times \ \exp\left(-\frac{c_M \Psi_s \zeta_{\rm sd}}{N_0 \sin^2\theta}\right) p(\zeta_{\rm sd}) d(\zeta_{\rm sd}) d\theta
$$
$$
= \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{1}{\pi} \int_0^{\frac{(M-1)\pi}{M}} \int_0^{b_{\max}}
$$
$$
\times \ \frac{1}{\left(\frac{c_M \Psi_s \eta_{\rm sd}}{N_0 \sin^2\theta}\right) + 1} \frac{P\{\omega = \omega_0 | m = 1\}}{\left[\frac{c_M(\omega_0/T_L)\eta_{\rm rd}}{N_0 \sin^2\theta}\right] + 1} d\omega_0 d\theta
$$
$$
+ \ \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} \frac{1}{\pi} \int_0^{\frac{(M-1)\pi}{M}} \frac{1}{\left(\frac{c_M \Psi_s \eta_{\rm sd}}{N_0 \sin^2\theta}\right) + 1} d\theta. \quad (44)
$$

Since $(1/x + 1) \approx (1/x)$ when $x \gg 1$, (44) can be further approximated as

$$
P_{M,\rm asym} \approx \lim_{\varepsilon \to 0} \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{1}{\pi} \int_0^{\frac{(M-1)\pi}{M}} \int_\varepsilon^{b_{\max}} \frac{N_0 \sin^2\theta}{c_M \Psi_s \eta_{\rm sd}}
$$
$$
\times \ \frac{N_0 \sin^2\theta}{c_M(\omega_0/T_L)\eta_{\rm rd}} P\{\omega = \omega_0 | m = 1\} d\omega_0 d\theta
$$
$$
+ \left(\frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} + P\big\{\omega = 0 \big| m = 1\big\}\right)
$$
$$
\times \ \frac{1}{\pi} \int_0^{\frac{(M-1)\pi}{M}} \frac{N_0 \sin^2\theta}{c_M \Psi_s \eta_{\rm sd}} d\theta
$$
$$
= \rho \times \frac{K_0}{c_M \eta_{\rm sd}} \left(\frac{\Psi_s}{N_0}\right)^{-1}
$$
$$
+ \left(\frac{T_L K_1 c_{R,0}}{c_M^2 \eta_{\rm sd} \eta_{\rm rd}} + \frac{\gamma_T K_0}{c_M \eta_{\rm sr} \eta_{\rm sd}}\right) \left(\frac{\Psi_s}{N_0}\right)^{-2}
$$
$$
- \ \frac{\gamma_T T_L K_1 c_{R,0}}{c_M^2 \eta_{\rm sd} \eta_{\rm rd} \eta_{\rm sr}} \left(\frac{\Psi_s}{N_0}\right)^{-3} \quad (45)
$$

where $\rho = P\{\omega = 0 | m = 1\}$, $c_{R,0}$, K_0 and K_1 are defined in (25)–(27), respectively.

This completes the proof.

Appendix F
Proof of Corollary 1

According to the proof of Theorem 3, when $\delta < E_{Hmin}$, $\omega(t) = E_H(t-1)$ and $\rho = P\{\omega = 0 | m = 1\} = 0$. In such a case, the average SER can be written as

$$
P_{M,\rm asym} \approx \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{1}{\pi} \times \frac{N_0^2 T_L}{c_M^2 \Psi_s \eta_{\rm sd} \eta_{\rm rd}} \int_0^{\frac{M-1}{M}\pi}
$$
$$
\times \ \sin^4\theta \left[\int_{E_{Hmin}}^{b_{\max}} E_H^{-1} p(E_H) dE_H \right.
$$
$$
\left. + \ b_{\max}^{-1} \times P\{E_H > b_{\max}\}\right] d\theta
$$
$$
+ \ \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} \frac{1}{\pi} \times \frac{N_0}{c_M \Psi_s \eta_{\rm sd}} \int_0^{\frac{M-1}{M}\pi} \sin^2\theta d\theta
$$
$$
= \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{N_0^2 T_L}{c_M^2 \Psi_s \eta_{\rm sd} \eta_{\rm rd}} K_1
$$
$$
\times \left[\int_{E_{Hmin}}^{b_{\max}} E_H^{-1} p(E_H) dE_H + b_{\max}^{-1} \times P\{E_H > b_{\max}\}\right]
$$
$$
+ \ \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} \times \frac{N_0}{c_M \Psi_s \eta_{\rm sd}} K_0. \quad (46)
$$

The (46) can be further bounded as follows:

$$
P_{M,\rm asym} \geq \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{N_0^2 T_L}{c_M^2 \Psi_s \eta_{\rm sd} \eta_{\rm rd} b_{\max}} K_1
$$
$$
+ \ \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} \times \frac{N_0}{c_M \Psi_s \eta_{\rm sd}} K_0
$$
$$
= \left(\frac{T_L K_1}{c_M^2 \eta_{\rm sd} \eta_{\rm rd} c_{R,1}} + \frac{\gamma_T K_0}{c_M \eta_{\rm sr} \eta_{\rm sd}}\right) \left(\frac{\Psi_s}{N_0}\right)^{-2}
$$
$$
- \ \frac{\gamma_T T_L K_1}{c_M^2 \eta_{\rm sd} \eta_{\rm rd} \eta_{\rm sr} c_{R,1}} \left(\frac{\Psi_s}{N_0}\right)^{-3} \quad (47)
$$

and

$$
P_{M,\rm asym} \leq \left(1 - \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}}\right) \frac{N_0^2 T_L}{c_M^2 \Psi_s \eta_{\rm sd} \eta_{\rm rd} E_{Hmin}} K_1
$$
$$
+ \ \frac{\gamma_T N_0}{\Psi_s \eta_{\rm sr}} \times \frac{N_0}{c_M \Psi_s \eta_{\rm sd}} K_0
$$
$$
= \left(\frac{T_L K_1}{c_M^2 \eta_{\rm sd} \eta_{\rm rd} c_{R,2}} + \frac{\gamma_T K_0}{c_M \eta_{\rm sr} \eta_{\rm sd}}\right) \left(\frac{\Psi_s}{N_0}\right)^{-2}
$$
$$
- \ \frac{\gamma_T T_L K_1}{c_M^2 \eta_{\rm sd} \eta_{\rm rd} \eta_{\rm sr} c_{R,2}} \left(\frac{\Psi_s}{N_0}\right)^{-3}. \quad (48)
$$

According to (47) and (48), the diversity order $d = 2$, and the EH gain is lower and upper bounded by $[(c_M^2 \eta_{\rm sd} \eta_{\rm rd} \eta_{\rm sr} c_{R,2})/(T_L K_1 \eta_{\rm sr} + \gamma_T K_0 c_M c_{R,2} \eta_{\rm rd})] \leq g_E \leq [(c_M^2 \eta_{\rm sd} \eta_{\rm rd} \eta_{\rm sr} c_{R,1})/(T_L K_1 \eta_{\rm sr} + \gamma_T K_0 c_M c_{R,1} \eta_{\rm rd})]$.

This completes the proof.

References

[1] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[3] W. Su, A. K. Sadek, and K. J. R. Liu, "Cooperative communication protocols in wireless networks: Performance analysis and optimum power allocation," *Wireless Pers. Commun.*, vol. 44, no. 2, pp. 181–217, Jan. 2008.

[4] M.-L. Ku, W. Li, Y. Chen, and K. J. R. Liu, "Advances in energy harvesting communications: Past, present, and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1384–1412, 2nd Quart., 2016.

[5] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, 3rd Quart., 2011.

[6] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[7] R. Ma and W. Zhang, "Adaptive MQAM for energy harvesting wireless communications with 1-bit channel feedback," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6459–6470, Nov. 2015.

[8] P. Blasco, D. Gunduz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.

[9] M. Rezaee, M. Mirmohseni, and M. R. Aref, "Energy harvesting systems with continuous energy and data arrivals: The optimal offline and heuristic online algorithms," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3739–3753, Dec. 2016.

[10] M.-L. Ku, Y. Chen, and K. J. R. Liu, "Data-driven stochastic models and policies for energy harvesting sensor communications," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1505–1520, Aug. 2015.

[11] M.-L. Ku, W. Li, Y. Chen, and K. J. R. Liu, "On energy harvesting gain and diversity analysis in cooperative communications," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2641–2657, Dec. 2015.

[12] W. Li, M.-L. Ku, Y. Chen, and K. J. R. Liu, "On outage probability for stochastic energy harvesting communications in fading channels," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1893–1897, Nov. 2015.

[13] W. Li, M.-L. Ku, Y. Chen, and K. J. R. Liu, "On outage probability for two-way relay networks with stochastic energy harvesting," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1901–1915, May 2016.

[14] W. Li, M.-L. Ku, Y. Chen, K. J. R. Liu, and S. H. Zhu, "Performance analysis for two-way network-coded dual-relay networks with stochastic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5747–5761, Sep. 2017.

[15] A. Manasian, S. ShahbazPanahi, and R. S. Adve, "Energy harvesting cooperative communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6118–6131, Nov. 2014.

[16] H. Li, N. Jaggi, and B. Sikdar, "Relay scheduling for cooperative communications in sensor networks with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 2918–2928, Sep. 2011.

[17] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

[18] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[19] Z. Wang, V. Aggarwal, and X. D. Wang, "Iterative dynamic water-filling for fading multiple-access channels with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 382–395, Mar. 2015.

[20] Z. Ding, S. M. Perlaza, I. Esnaola, and H. V. Poor, "Power allocation strategies in energy harvesting wireless cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 846–860, Feb. 2014.

[21] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," in *Synthesis Lectures on Communication Networks*, vol. 3. San Rafael, CA, USA: Morgan & Claypool, 2010, pp. 1–211.

[22] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. Q. Zhang, "A survey on delay-aware resource control for wireless systems—Large deviation theory, stochastic Lyapunov drift, and distributed stochastic learning," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1677–1701, Mar. 2012.

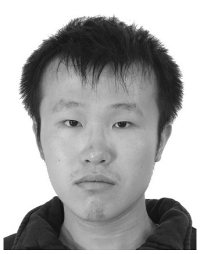
[23] J. He, Z. Xue, D. Wu, D. P. O. Wu, and Y. G. Wen, "CBM: Online strategies on cost-aware buffer management for mobile video streaming," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 242–252, Jan. 2014.

[24] X. M. Wu, J. Yang, Y. Y. Ran, and H. S. Xi, "Adaptive scalable video transmission strategy in energy harvesting communication system," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2345–2353, Dec. 2015.

[25] M. K. Simon and M. S. Alouini, *Digital Communication Over Fading Channels*. Hoboken, NJ, USA: Wiley, 2000.

[26] *Cooperative Networks for Renewable Resource Measurements (CONFRRM) Solar Energy Resource Data*, NREL, Elizabeth City, NC, USA, Sep. 2012. [Online]. Available: http://www.nrel.gov/rredc

[27] W. C. Jakes, *Microwave Mobile Communications*. New York, NY, USA: Wiley, 1974.



Chengrun Qiu is currently pursuing the B.S. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include reinforcement learning, stochastic optimization, and green communications.

Yang Hu received the B.S. and Ph.D. degrees in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2004 and 2009, respectively.

She was with the Institute for Advanced Computer Studies, University of Maryland at College Park, College Park, MD, USA, as a Research Associate from 2010 to 2015. She is currently an Associate Researcher with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include computer vision, machine learning, and multimedia signal processing.

Yan Chen (SM'14) received the bachelor's degree from the University of Science and Technology of China, Hefei, China, in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, in 2007, and the Ph.D. degree from the University of Maryland at College Park, College Park, MD, USA, in 2011.

He was with Origin Wireless Inc., Greenbelt, MD, USA, as a Founding Principal Technologist. Since 2015, he has been a Full Professor with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. His current research interests include multimedia, signal processing, game theory, and wireless communications.

Dr. Chen was a recipient of multiple honors and awards including the Best Student Paper Award of the PCM in 2017, the Best Student Paper Award of IEEE ICASSP in 2016, the Best Paper Award of IEEE GLOBECOM in 2013, the Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, University of Maryland at College Park in 2011, and the Chinese Government Award for Outstanding Students Abroad in 2010.