Lyapunov-Optimized Two-Way Relay Networks With Stochastic Energy Harvesting

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

***Abstract*— Energy harvesting wireless cooperative communications have become more and more popular in recent years. In this paper, we consider a two-way relay cooperative network, where the relay is an energy harvesting node and uses decode-and-forward (DF) or amplify-and-forward (AF) cooperation protocol. We formulate the energy management problem in such a network as an optimization problem, which minimizes the long-term average outages with a long-term average battery constraint. We then apply Lyapunov optimization to transform the long-term optimization problem into the drift-plus-penalty. We also conduct theoretic analysis of the proposed energy management strategy. Specifically, we prove that the long-term average battery constraint can be guaranteed with the proposed strategy and derive an upper bound of the average outages with the proposed strategy. Furthermore, we analyze theoretically the diversity order and energy harvesting gain of the proposed strategy with the DF and AF protocols. Finally, simulation results using real-solar irradiance data measured by the solar site in Elizabeth City University show that the outage performance and the diversity order of our algorithms are better than the existing MDP-based method.**

***Index Terms*— Energy harvesting, cooperative communications, continuous energy, outage probability, Lyapunov optimization.**

I. Introduction

In the past decade, cooperative communications have gained much attention for its capability in improving the link quality of the wireless communications [1]. Among various cooperative communication models, two-way relay networks have the advantage of higher transmission efficiency since it can exchange information between two source nodes with the broadcasting nature of the relay's wireless transmission [2]. Two typical and widely adopted cooperation protocols are the decode-and-forward (DF) and amplify-and-forward (AF) protocols [3], [4], where the relay with DF protocol first decodes the received signals and then encodes and transmits the decoded signals to the destinations if the decoding is successful, while the relay with AF protocol directly amplifies the received signals and transmits to the destinations.

Many works on the cooperative communications have been conducted in the literature. Javan *et al.* in [5] considered the resource allocation problem in DF cooperative communications with limited rate feedback by maximizing the achievable rate with the total average transmit power constraint, while Zhao *et al.* in [6] proposed to minimize the total energy consumption of the users with the outage constraint. Chen *et al.* maximized the sum-throughput of multiple users in finite period and derived a water-filling based branch-and-bound solution [7]. In [8], Qiu *et al.* minimized the long-term average error rate in one-way relay cooperation communications with stochastic energy harvesting. As for two-way relay network, the authors in [9] developed an optimal relay transmission policy to maximize the long-term average throughput of the two-way relay network with energy harvesting. In addition, the optimal transmission strategy for wireless energy transfer in two-way relay networks was studied in [10].

Considering the limited size of the relay node, battery with large capacity is generally impractical at the relay, as the battery may need to be replaced or recharged frequently. For this reason, energy harvesting cooperative communications have been investigated [11]. There are three typical energy harvesting architectures [12], the harvest-use (HU) architecture which directly uses the instantaneous harvested energy without storage, the harvest-storage-use (HSU) architecture where the current harvested energy will be saved and only available in the following period, and the harvest-use-storage (HUS) which can immediately utilize the harvested energy and store the remaining energy for future use.

The energy management in energy harvesting wireless communications is very challenging due to the stochastic nature of harvested energy and wireless channel. To tackle this challenge, Markov decision process (MDP) has been widely used [13]–[18]. In point-to-point energy harvesting wireless communications, Ulukus *et al.* proposed to use MDP to maximize the throughput of the energy harvesting wireless communications [13], while Ma and Zhang maximized the throughput of the energy harvesting communications with one-bit channel feedback by adapting MQAM modulation [14]. The optimal power level allocation algorithm to maximize the long-term expected throughput was proposed in [15], while Rezaee *et al.* in [16] maximized the single-user throughput of energy harvesting communications with continuous energy

Manuscript received March 19, 2018; revised June 4, 2018; accepted July 17, 2018. Date of publication July 31, 2018; date of current version September 10, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61602090 and Grant 61672137, in part by the 111 Project No. B17008, and in part by the Thousand Youth Talents Program of China (to Yan Chen). *(Corresponding author: Yan Chen.)*

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

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2018.2858226

and data arrivals. The outage probability and the net bit rate for stochastic energy harvesting communications in fading channel were optimized with MDP in [17] and [18], respectively.

There are some works on energy harvesting cooperative communications based on MDP [19]–[22]. Minasian *et al.* maximized the average throughput of two-hop AF-based energy harvesting cooperative communications with the aid of MDP in [19]. In [20], the long-term average error rates of the DF-based cooperative communications with an energy harvesting relay was optimized. The average outage probabilities of the cooperative communications in AF and DF protocols with an energy harvesting two-way relay were minimized separately by means of MDP [21]. Li *et al.* in [22] developed strategies for more efficient energy usage with MDP. However, the computational complexity of an MDP can be very challenging when the state and action spaces are large.

To avoid the computational complexity challenge introduced by the MDP, an alternative way to the energy management problem in energy harvesting wireless communications is to formulate the problem with the long-term average objective function and/or constraints, and Lyapunov optimization [27] can be utilized for this type of formulation. Examples include the resource allocation in energy harvesting multi-hop networks [28], [29], [32], and resource allocation in cellular networks powered by both energy harvesting and the grid [30], [31], [33]. In [34], Cui *et al.* utilized Lyapunov optimization to study the delay-aware resource control problem, where the system throughput, the sum delay and the power consumption are jointly optimized. In [35], the problem to optimize the time average expected transmission cost for the mobile video streaming system was solved with Lyapunov optimization by balancing the data wastage cost and the freezing time of the mobile user. By considering the playback interruption and fluctuation, Wu *et al.* [36] maximized the received video quality of an energy harvesting video communication system with Lyapunov optimization. Amirnavaei *et al.* in [37] used Lyapunov optimization to maximize the throughput in energy harvesting wireless communications with the constraint where the average energy used for transmission equals the average harvested energy. In [38], Zhang *et al.* utilized Lyapunov optimization to manage the resource for energy harvesting cognitive radio sensor networks with the constraints of collision, data queue and the spare capacity of the battery.

In this paper, we consider a two-way relay cooperative communication network with two source nodes and one energy harvesting relay node. We propose an optimal continuous energy management for the relay to minimize the average outages based on Lyapunov optimization. The major contributions are summarized as follows:

- We formulate the energy management problem in a two-way relay energy harvesting cooperative communication network as an optimization problem which minimizes the long-term average outages with a long-term average battery constraint. We then apply Lyapunov optimization to transform the long-term optimization problem into the drift-plus-penalty. To the best of our knowledge, this is the first work considering Lyapunov optimization for the two-way relay energy harvesting cooperative communication network.
- We conduct theoretic analysis of the proposed energy management strategy. We first prove that the long-term average battery constraint can be guaranteed with the proposed strategy. Then we derive an upper bound of the average outages with the proposed strategy. Furthermore, we theoretically show that with the proposed strategy, the diversity order is 1 for both DF and AF protocols, and derive the bounds for the corresponding energy harvesting gains. Finally, we theoretically analyze the potential instantaneous energy used for transmission.
- We conduct simulations to evaluate the outage performance and the diversity order of our proposed algorithms. We use real solar irradiance data measured by the solar site in Elizabeth City University. It is shown that the outage performance and the diversity order of our algorithms are better than the state-of-the-art MDP based method [21].

Perhaps the most related works to this paper are [8] and [21]. The main differences between this paper and [8] are as follows. First, this paper considers a two-way relay network without the source-destination link and adopts the outage probability as the performance metric, while [8] considers a classic one-way relay network with the source-destination link and adopts the symbol error rate as the performance metric. Thus, this paper is different from [8] in both the network architecture and the objective function. Second, the major parts of the analytical results in this paper, which are the diversity order, energy harvesting gain and the energy distribution, are completely new contributions. Third, the detailed proof of the battery constraint is different since the objective function is different. Fourth, the simulation results and the insights are different. The main difference between this paper and [21] is the optimization technique. This paper utilizes the Lyapunov optimization while [21] uses MDP. In such a case, [21] needs to discretize the variable and uses the computational expensive MDP to solve the problem, while this paper directly uses the continuous variable and can find the solution in a computational efficient way. Moreover, as shown in the simulation results, by using the Lyapunov optimization and the continuous variable, the proposed method can achieve much better performance compared with [21].

The rest of this paper is organized as follows. Section II introduces the system model and problem formulation in detail. In section III we discuss how to utilize Lyapunov optimization to solve the problem. Theoretical analysis is conducted in Section IV, including the satisfaction of the battery constraint, the outage performance, the diversity order and energy harvesting gain, as well as the instantaneous energy used for transmission distribution. Finally, section V shows the simulation results and section VI draws the conclusions.

II. System Model and Problem Formulation

In this paper, we consider an energy harvesting two-way relay transmission network, similar to [21], as shown in Fig. 1, where there are two wireless source nodes sending

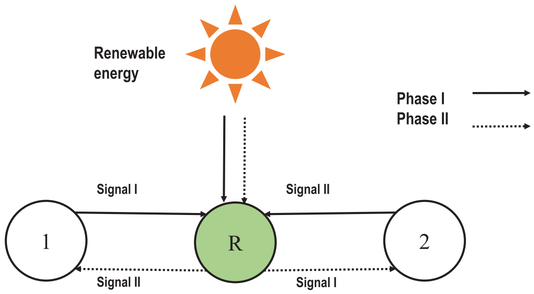


Fig. 1. An illustration of the system model.

and receiving information to/from each other with the aid of an energy harvesting relay node. Note that two-way relay cooperative communications have been practically used in wireless sensor networks (WSN) [39], D2D communications [40], cellular communications [41], etc. The relay first receives information from the two source nodes in the multiple access (MA) phase, and then broadcasts (BC) the modified (decoded or amplified) information to the two source nodes in broadcast phase. The relay is assumed to be able to harvest energy in both phases and adapt continuous energy for broadcasting. The transmission power of source node 1 and 2 is assumed to be the same, which is represented by P_S. The transmission power of the relay is denoted as P_R and the instantaneous energy consumed for transmission is thus $\omega = P_R \times T_M$ with T_M being the transmission duration of the broadcasting phase. It is also assumed that there is no direct transmission between source node 1 and source node 2, and the wireless channels are reciprocal, quasi-static and Rayleigh flat fading. The instantaneous power of the channels between the source nodes and the relay are denoted as γ_1 and γ_2. Specifically, we define the instantaneous power of the channels between the source nodes and the relay at the t^{th} management period as $\gamma_1(t)$ and $\gamma_2(t)$, and the average noise power is assumed to be the same at all time periods and is represented as N_0. We study the amplify-and-forward (AF) and decode-and-forward (DF) cooperation protocols in this paper.

A. Harvestable Energy

Several different kinds of renewable energy such as solar and wind can be used in the energy harvesting wireless sensor networks [11]. In this paper, we assume that the harvestable energy at every time slot is positive,[1] i.e., $E_{Hmin} > 0$. The power of the renewable energy source is assumed to be quasi-static in one policy management duration T_M. Note that the energy harvesting interval will influence the value of $E_H(t)$ and the resource allocation strategy is determined by $E_H(t)$. Thus, when the two intervals are different, the solution will be slightly different and the technique in [42] may be useful. To make the full use of the harvested energy, the rechargeable battery is employed at the transmission relay. While the proposed method is suitable to all three typical battery structures [12], in this paper, we utilize the harvest-use-store (HUS) model for fair comparison with the existing benchmark,

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

where b_{max} denotes the capacity of the battery and $\omega(t) \in [0, b(t)]$ is the energy used for transmission.

B. Transmission Outage

The outage will happen if the transmission data rate, either the source-to-relay link or the relay-to-source link, is out of the achievable rate region, which is determined by the cooperation protocols in the two-way relay network. According to [43] and [44], the outage events of DF and AF protocols are listed in this section. In the following, we will discuss the outage under the DF and AF cooperation protocols, respectively.

1) Decode-and-Forward: With the DF cooperation protocol, the outage happens when the data rate is larger than the minimum of the two mutual information of the two transmission phases, or the aggregated data rate is larger than the sum-rate constraint, which leads to the failure of decoding two received signals simultaneously in the MA phase. In other words, the outage events can be written as (2), shown at the top of the next page.

2) Amplify-and-Forward: With the AF cooperation protocol, the relay amplifies the received signals and forwards them to the nodes. Thus, the outage happens when the data rate is larger than the mutual information computed by the corresponding end-to-end SNR. The outage events can be written as (3), shown at the top of the next page.

From (2) and (3), we can see that the outage events are determined by the transmission power of the relay, i.e., $P_R = \omega/T_M$. Therefore, the outage probability is determined by the energy used for transmission of the relay ω according to the following Theorem.

Theorem 1: In a two-way relay network, for any given transmission rates, R_1 and R_2, and the energy used for transmission $\omega(t)$ of the relay, the outage probability is a step function shown as below

$$P_{out}(\omega(t)) = \begin{cases} 1, & \text{if } \omega(t) < \hat{\omega}_P(t); \\ 0, & \text{if } \omega(t) \geq \hat{\omega}_P(t); \end{cases} \tag{4}$$

where the threshold $\hat{\omega}_P(t)$ for the DF protocol and AF protocol can be written in (5) and (6), as shown at the top of the next page, respectively.

Proof: Please see the proof in APPENDIX A. ■

C. Problem Formulation

Our goal is to help the relay determine the energy used for transmission at every time slot. To do so, we use the long-term average outage probability as the objective function as follows

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=0}^{T-1} P_{out}(\omega(t)). \tag{7}$$

There exists a non-negligible probability that the channel power is small according to the Rayleigh fading characters.

[1]In this paper, we only consider the energy management for the period where the harvestable energy is active. In such a case, it is reasonable to assume that the harvestable energy at each time slot is positive.

$$\xi_{DF}=\left\{R_1>\min\left\{\frac{1}{2}\log\left(1+\frac{\gamma_1P_S}{N_0}\right),\frac{1}{2}\log\left(1+\frac{\gamma_2P_R}{N_0}\right)\right\}\right\}\cup\left\{R_2>\min\left\{\frac{1}{2}\log\left(1+\frac{\gamma_2P_S}{N_0}\right),\frac{1}{2}\log\left(1+\frac{\gamma_1P_R}{N_0}\right)\right\}\right\}$$
$$\cup\left\{R_1+R_2>\frac{1}{2}\log\left(1+\frac{\gamma_1P_S}{N_0}+\frac{\gamma_2P_S}{N_0}\right)\right\}. \quad (2)$$

$$\xi_{AF}=\left\{R_1>\frac{1}{2}\log\left(1+\frac{\gamma_1\gamma_2P_SP_R}{N_0(\gamma_1P_S+\gamma_2P_S+\gamma_2P_R+N_0)}\right)\right\}\cup\left\{R_2>\frac{1}{2}\log\left(1+\frac{\gamma_1\gamma_2P_SP_R}{N_0(\gamma_1P_S+\gamma_2P_S+\gamma_1P_R+N_0)}\right)\right\}. \quad (3)$$

$$\hat{\omega}_{DF}(t)=\begin{cases}+\infty, & \text{if } [\frac{\gamma_1(t)P_S}{N_0}\le 2^{2R_1}-1]\cup[\frac{\gamma_2(t)P_S}{N_0}\le 2^{2R_2}-1]\cup\{\frac{[\gamma_1(t)+\gamma_2(t)]P_S}{N_0}\le 2^{(2R_1+2R_2)}-1\};\\ T_M\times\max\left\{\frac{(2^{2R_1}-1)N_0}{\gamma_2(t)},\frac{(2^{2R_2}-1)N_0}{\gamma_1(t)}\right\}, & \text{else.}\end{cases} \quad (5)$$

$$\hat{\omega}_{AF}(t)=\begin{cases}+\infty, & \text{if } [\frac{\gamma_1(t)P_S}{N_0}\le 2^{2R_1}-1]\cup[\frac{\gamma_2(t)P_S}{N_0}\le 2^{2R_2}-1];\\ T_M\times\max\left\{\frac{(2^{2R_1}-1)N_0[\gamma_1(t)P_S+\gamma_2(t)P_S+N_0]}{\gamma_1(t)\gamma_2(t)P_S-(2^{2R_1}-1)N_0\gamma_2(t)},\frac{(2^{2R_2}-1)N_0[\gamma_1(t)P_S+\gamma_2(t)P_S+N_0]}{\gamma_1(t)\gamma_2(t)P_S-(2^{2R_2}-1)N_0\gamma_1(t)}\right\}, & \text{else,}\end{cases} \quad (6)$$

Consuming large amount of energy for avoiding outage when the channel is under very poor conditions may not be effective. Saving energy and waiting for better channel power will get more payback. 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 follow

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}b(t)\ge A, \quad (8)$$

where A is a pre-defined energy level threshold.

Therefore, the problem can be formulated as to adapt the energy used for transmission $\omega(t)$ at each time slot for optimizing the long-term average outage probability while satisfying the battery constraint as follows

$$\begin{aligned}&\min_{\omega(t)}\ \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}P_{out}(\omega(t)),\\ &s.t.\ \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}b(t)\ge A,\\ &\qquad \omega(t)\in[0,b(t)].\end{aligned} \quad (9)$$

III. Dynamic Strategy on Optimizing Outages

Both the objective function and constraint in the optimization problem in (9) 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

$$B(t+1)=\max\{B(t)+A-b(t+1),0\}. \quad (10)$$

From (10), 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 below

$$L(\Theta(t))=\frac{1}{2}B^2(t). \quad (11)$$

Based on [27], the penalty drift can be written as follows

$$\Delta\Theta(t)=\mathbb{E}[L(\Theta(t+1))-L(\Theta(t))|\Theta(t)], \quad (12)$$

and the long-term averaging terms can be re-written as "drift-plus-penalty"

$$\Delta\Theta(t)+V\mathbb{E}[P_{out}(\omega(t))|\Theta(t)], \quad (13)$$

where parameter $V\ge 0$ 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 (9) can be re-written as

$$\begin{aligned}&\min_{\omega(t)}\ \Delta\Theta(t)+V\mathbb{E}[P_{out}(\omega(t))|\Theta(t)],\\ &s.t.\ 0\le\omega(t)\le b(t).\end{aligned} \quad (14)$$

We can make some simplification of the target function by finding the terms related to the transmission power and ignoring other terms similar to that in [8]. Then, the objective function in (14) can be upper bounded as

$$\begin{aligned}&\Delta\Theta(t)+VP_{out}(\omega(t))\\ &\le\frac{1}{2}\times\{A^2+E_{Hmax}^2+b_{max}^2+\omega^2(t)-2b(t)\omega(t)\\ &\quad+2b_{max}E_{Hmax}+2B(t)\times[A-b(t)+\omega(t)]\}\\ &\quad+VP_{out}(\omega(t)),\\ &=\frac{1}{2}\times\{B_1+B_2+\omega(t)^2-2b(t)\omega(t)+2B(t)\omega(t)\}\\ &\quad+VP_{out}(\omega(t)),\end{aligned} \quad (15)$$

where $B_1=A^2+E_{Hmax}^2+b_{max}^2+2b_{max}E_{Hmax}$ is a constant during all time slots and $B_2=2B(t)\times[A-b(t)]$ is fixed at a specific time slot. Therefore, by ignoring the terms that are not related to $\omega(t)$, the optimization problem (9) can be approximated as

$$\begin{aligned}\min_{\omega(t)}\ & J(\omega(t)) = \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t)\\ &\qquad + 2VP_{out}(\omega(t)),\\ s.t.\ & 0 \le \omega(t) \le b(t). \end{aligned} \tag{16}$$

Remark: From (14) and (16), we can see that the penalty drift is upper bounded with a quadratic function of $\omega(t)$. By doing so, the optimization problem can be greatly simplified, which makes the analysis possible.

IV. Performance Analysis

A. Battery Constraint

In this subsection we would like to show that the solution derived in (16) satisfies the long-term average battery buffer constraint in (8) with any reasonable threshold $A < b_{max}$.

Theorem 2: The long-term battery constraint is guaranteed under the Lyapunov optimization for fading with bounded channel gain.

Proof: If $B(t) < b(t)$, $B(t)$ is bounded by $B(t) \le b_{max}$. On the other hand, if $B(t) \ge b(t)$, then we take the derivative of the objective function in (16) and get

$$\frac{dJ(\omega(t))}{d\omega(t)} = 2\{\omega(t) - b(t) + B(t) - V\delta[\omega(t) - \hat{\omega}(t)]\}, \tag{17}$$

where $\delta(x)$ is the impulse function.

According to (17), when $B(t) \ge b(t)$, the derivative $\frac{dJ(\omega(t))}{d\omega(t)}$ is increasing except at the point $\omega(t) = \hat{\omega}(t)$, thus the optimal $J(\omega(t))$ is achieved at either 0 or $\hat{\omega}(t)$. In the following, we show that $B(t)$ is upper bounded in both situations where $\omega^*(t) = 0$ or $\omega^*(t) = \hat{\omega}(t)$.

Case 1. The optimum energy used for transmission $\omega^*(t) = \hat{\omega}(t)$. In such a case, we have

$$J(\omega(t))\Big|_{\omega(t)=0} > J(\omega(t))\Big|_{\omega(t)=\hat{\omega}(t)}, \tag{18}$$

which means that

$$V > \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t). \tag{19}$$

With (19), $B(t)$ is bounded by

$$\begin{aligned} B(t) &< \frac{V}{\hat{\omega}(t)} - \frac{1}{2}\hat{\omega}(t) + b(t)\\ &\le \frac{V}{\hat{\omega}_{min}} - \frac{1}{2}\hat{\omega}_{min} + b_{max}. \end{aligned} \tag{20}$$

Case 2. The optimum energy used for transmission $\omega^*(t) = 0$. In such a case, we have

$$J(\omega(t))\Big|_{\omega(t)=0} \le J(\omega(t))\Big|_{\omega(t)=\hat{\omega}(t)}, \tag{21}$$

i.e.

$$V \le \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t). \tag{22}$$

We can tell the condition of $B(t)$ to enter this case is

$$B(t) \ge \max\left[\frac{V}{\hat{\omega}(t)} - \frac{1}{2}\hat{\omega}(t) + b(t), b(t)\right]. \tag{23}$$

Suppose that there exists a t_0 such that

$$B(t_0) \ge \max\left[\frac{V}{\hat{\omega}(t_0)} - \frac{1}{2}\hat{\omega}(t_0) + b(t_0), b(t_0)\right], \tag{24}$$

but

$$B(t_0) \le \max\left[\frac{V}{\hat{\omega}_{min}} - \frac{1}{2}\hat{\omega}_{min} + b_{max}, b_{max}\right]. \tag{25}$$

According to (10), if $b(t+1) < A$, we have $B(t) < B(t+1) < B(t)+A$, and if $b(t+1) \ge A$, we have $B(t+1) \le B(t)$. From (1), we can see that for $\omega(t) = 0$, after at most $T_0 = \frac{A}{E_{Hmin}}$ time slots, $b(t+T_0) > A$. Therefore, $B(t)$ will increase at most T_0 time slots. After that, the $B(t)$ starts to decrease until $B(t) < \frac{V}{\hat{\omega}(t)} - \frac{1}{2}\hat{\omega}(t) + b(t)$, which is the condition to enter Case 1. Therefore, for Case 2, we have

$$B(t) \le \max\left[\frac{V}{\hat{\omega}_{min}} - \frac{1}{2}\hat{\omega}_{min} + b_{max}, b_{max}\right] + T_0 A, \tag{26}$$

where $T_0 = \frac{A}{E_{Hmin}}$.

With (20) and (26), we can see that $B(t)$ is upper bounded. According to the *Lemma 1* in [8], the constraint of the average battery energy is guaranteed.

This completes the proof. ■

B. Bounded Performance

In this subsection, we analyze the performance of the proposed algorithm. Let P_{out}^* denote the minimum objective value of the expected outage probability to the solution (9). The following theorem provides a bound of the performance of the proposed algorithm.

Theorem 3: Assuming the related states in the algorithm is i.i.d over time, the average outage function of the optimal strategy obtained by (9) is limited by an upper bound that is independent of the operation time, i.e.,

$$\bar{P}_{out} - P_{out}^* \le \frac{B_1 + 2b_{max}B_{max}}{2V} \tag{27}$$

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

Proof: The proof is similar to that in [8]. Please see the proof of [8, Th. 2] for detail. ■

C. Extension to General Objective Functions

In this subsection, we would like to discuss the extension of Theorem 2 and Theorem 3 to the energy harvesting problems with general objective functions.

Corollary 1: The long-term battery constraint can be satisfied if one of the following conditions holds

- the threshold $A \le \min(\mathbb{E}(E_H), b_{max})$,
- the objective function $f(\omega)$ is continuous with bounded first order of derivative.

Proof:

Case 1: according to (1), we have $b(t) \geq \min(E_H(t), b_{max})$. Then, the long-term battery constraint can be bounded as follows

$$\lim_{t\to+\infty} \frac{1}{t}\sum_{i=1}^{t} b(i) \geq \lim_{t\to+\infty} \frac{1}{t}\sum_{i=1}^{t} \min(E_H(t), b_{max}), \\ = \min(\mathbb{E}(E_H), b_{max}) \geq A. \quad (28)$$

Therefore, when $A \leq \min(\mathbb{E}(E_H), b_{max})$, the long-term battery constraint can be satisfied.

Case 2: with the general objective function $f(\omega)$, (17) becomes

$$\frac{dJ(\omega(t))}{d\omega(t)} = 2\{\omega(t) - b(t) + B(t) + Vf'(\omega(t))\}. \quad (29)$$

When the first order of derivative is bounded, i.e., $f'(\omega(t)) \geq c$, then we have

$$\frac{dJ(\omega(t))}{d\omega(t)} \geq 2\{\omega(t) - b(t) + B(t) + c\}. \quad (30)$$

In such a case, if $B(t)$ starts to exceed $b_{max}-c$, the optimum energy for transmission will be 0 to ensure the decrease of the $B(t)$, which can be called as the charge-only pattern. Similar to the proof in Theorem 2, there exists a maximal length of the charge-only period and thus the $B(t)$ will be bounded. ■

Different from Theorem 2, Theorem 3 is applicable for general energy harvesting problems as long as the system state is i.i.d.. This is determined by the ω-only-policy [27].

D. Diversity Order and Energy Harvesting Gain

We then analyze the diversity order and energy harvesting gain of the DF and AF protocol. The diversity order and the energy harvesting gain can be defined as follows: at the asymptotically high SNR $\Upsilon = \frac{\gamma_0\omega}{N_0}$ where γ_0 is the average channel power, if the outage $P_{out}(\Upsilon) \approx g_E^{-1} \times \Upsilon^{-d}$, then d is the diversity order and g_E is the energy harvesting gain [1] [20].

Without loss of generality, the threshold A is set to be smaller than E_{Hmin} in the analysis below. The results are summarized in Theorem 4.

Theorem 4: The diversity orders of the proposed method with the DF and AF protocols are both equal to 1, and the energy harvesting gains of the DF and AF protocols are bounded as

$$\left\{\max\left[(2^{2R_1}-1), \frac{(2^{2R_2}-1)}{c_1}\right] + \max\left[(2^{2R_2}-1), \frac{(2^{2R_1}-1)}{c_1}\right]\right\}^{-1} < g_{DF} < \left[(2^{R_1+R_2+1}-2)\max(1, c_2^{-1})\right]^{-1} \quad (31)$$

and

$$(2^{2R_1} + 2^{2R_2} - 2)^{-1} < g_{AF} < \frac{\min(c_3, 1)}{2\max[5(2^{2R_1+2R_2}-1), 1]}, \quad (32)$$

where $c_1 = \frac{b_{max}}{T_M P_S}$, $c_2 = \frac{E_{Hmin}}{T_M P_S}$ and $c_3 = \frac{P_R}{P_S}$.

Proof: Please see APPENDIX B for proof and detail. ■

From *Theorem 4*, we can see that our proposed strategy can achieve diversity order one, and the energy harvesting gains can be bounded with the values determined by the transmission rates R_1 and R_2.

E. Energy Distribution

In this subsection, we study the instantaneous optimum energy used for transmission, and the results are shown in the following theorem.

Theorem 5: The optimal energy used for transmission is 0, $b(t) - B(t)$ or $\hat{\omega}(t)$, and the conditions to achieve the optimal energy used for transmission are

1) Silent Strategy ($\omega^*(t) = 0$): The silent strategy is used when one of the following conditions holds

- $b(t) \leq B(t)$ and $\hat{\omega}(t) > b(t)$;
- $b(t) \leq B(t)$ and $V \leq \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t)$;

2) Constraint-Preferred Strategy ($\omega^*(t) = b(t) - B(t)$): The constraint-preferred strategy is utilized when one of the following conditions holds

- $b(t) > B(t)$ and $\hat{\omega}(t) > b(t)$;
- $b(t) > B(t)$ and $V \leq \frac{1}{2}[b(t) - B(t)]^2 + \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t)$;
- $b(t) > B(t)$ and $b(t) - B(t) > \hat{\omega}(t)$;

3) Outage-Preferred Strategy ($\omega^*(t) = \hat{\omega}(t)$): The outage-preferred strategy is used when one of the following conditions hold

- $\hat{\omega}(t) \leq b(t)$, $b(t) \leq B(t)$ and $V > \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t)$;
- $\hat{\omega}(t) \leq b(t)$, $b(t) > B(t)$ and $V > \frac{1}{2}[b(t) - B(t)]^2 + \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t)$.

Proof: Please see APPENDIX C for proof and detail. ■

From Theorem 5, we can see that the probability of **Outage-preferred strategy** is small for a small V since it requires the V to be larger than some thresholds. On the contrary, the probability of **Silent strategy** can be large since the condition $V \leq \frac{1}{2}\hat{\omega}^2(t) - b(t)\hat{\omega}(t) + B(t)\hat{\omega}(t)$ is easy to reach, which makes the P_{out} tend to be large. Hence, a larger V is helpful for the entire outage probability. Based on this observation, we also consider a strategy, named "Constraint loosed strategy", where the penalty weight V is infinity as shown in Algorithm 1.

Algorithm 1 Constraint Loosed Strategy

```
Initialize: V, b_max, P_S, N_0, R_1, R_2, b(1) and cooperative patterns
for t = 1 to +∞ do
    Observing E_H(t), γ_1(t) and γ_2(t)
    Calculating the ω̂(t)
    if b(t) ≥ ω̂(t) then
        ω(t) = ω̂(t)
    else
        ω(t) = 0
    end if
    b(t+1) = min{b(t) − ω(t) + E_H(t), b_max}
end for
```

V. Numerical Simulations

In this section, we conduct simulations to evaluate the performance of the proposed methods, in terms of long-term average outage probability, long-term average battery, and long-term average energy used for transmission, by comparing with the state-of-the-art MDP method in [21]. We also illustrate the outage probability under different SNRs, based on which the diversity order can be evaluated. In all the figures, the "Loosed" represents the "Constraint loosed algorithm". This scheme is not as a benchmark to the proposed algorithm. Instead, it is the proposed scheme without the long-term battery constraint.

The simulation settings are shown as below. The policy management duration of the cooperative communications T_M is set to be 300 seconds where an entire process of two-way transmission is finished. The power of the transmission sources is set as $P_S = 35mW$. In the simulations, it is assumed that there exists a positive $\sigma = 0.5$ standing for the ratio between the target rate and the target sum rate as $\sigma R = R_1$, $(1-\sigma)R = R_2$, and the sum rate R is set to be 4. We utilize the data records of the solar irradiance measured in solar site in Elizabeth City State University in June from 2010 to 2012 [45]. The solar panel size is set as $4cm^2$ and the harvesting efficiency is 20%. Capacity of the battery is assumed to be $35 \times 12 \times T_M$ mJ. The channels are assumed to be independent and Rayleigh fading, and for imitating the slow variation of the wireless channel, the Doppler frequency is set to be $f_D = 5 \times 10^{-2}$. 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 [46]. The channel power quantization thresholds for the fading channel in the MDP-based method [21] are $\{0, 0.3, 0.6, 0.9, 2.0, 3.0, +\infty\}$ or $\{0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9, 1.5, 2.0, 2.5, 3.0, +\infty\}$. Since the MDP-based method cannot deal with the continuous energy, the number of the battery energy state is 12 or 24, and the basic energy level is $35 \times T_M$ mJ. The number of solar state is 4 and the discount factor is set as 0.99.

We first compare the average outage performance of the two proposed methods with the MDP method for different numbers of states under various expected normalized SNRs $10 \log \frac{P_S \gamma}{N_0}$ where the threshold $A = 5 * (10 \log \frac{P_S \gamma}{N_0})^{-1}$. The results with the AF and DF protocols are shown in Fig.2 and Fig.3, respectively. From Fig.2, we can see that, with the AF protocol, the proposed methods perform better than the MDP method, and the proposed Loosed algorithm works the best with the SNR varying from 0dB to 40dB. This is because that, at the middle and high SNR regions, a small amount of energy may avoid outage, due to which the managements are better if the basic management level is smaller. Therefore, the proposed methods which are based on the continuous energy management work better, especially at the higher SNR regimes. The reason that the proposed Loosed algorithm performs the best is that it ignores the average battery constraint, due to which the outage performance can be better with the cost of violating the average battery constraint.

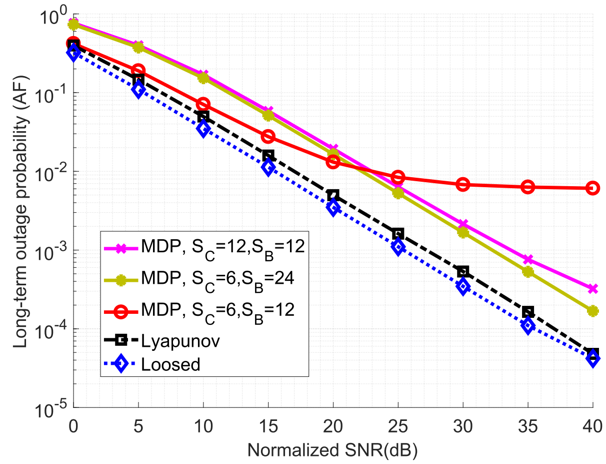


Fig. 2. Comparison of the average outages under different SNRs (AF).

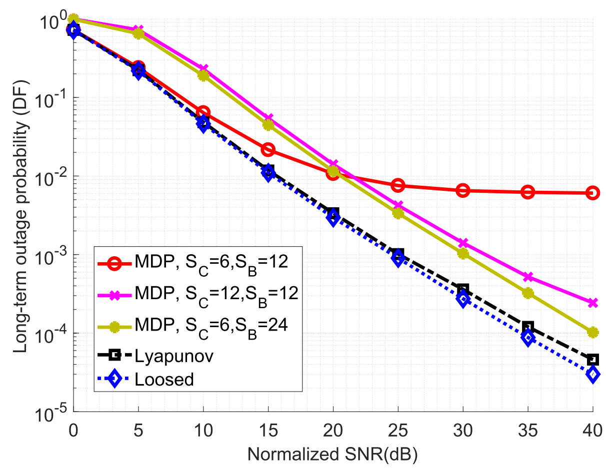


Fig. 3. Comparison of the average outages under different SNRs (DF).

From Fig.3, we can see that with the DF protocol, the long-term average outage performance is almost the same for all methods at the low SNR regions. When the SNR is larger than 5dB, the performance of the MDP method with a small number of states ($S_C = 6$ and $S_B = 12$) starts to saturate, leading to the worst performance. Similar to the results in the AF protocol, the proposed Loosed algorithm performs the best with the cost that the battery constraint may not be satisfied, which will be shown later.

From Fig.2 and Fig.3, we can also observe that the diversity order of the proposed algorithms is one, while the diversity order of the MDP method with a small number of states ($S_C = 6$ and $S_B = 12$) is zero. This is because that, with the MDP method, the probability that the available energy is smaller than basic energy level is positive, i.e., the probability of a silent relay exists. To improve the performance of the MDP method, we need to increase the number of the states in the MDP method, i.e., to decrease the basic level of the MDP method. However, with the increase of the number of the states, the computational complexity of the MDP method grows rapidly, which prevents us to use a large number of states in the MDP method. Moreover, from the figures, we can

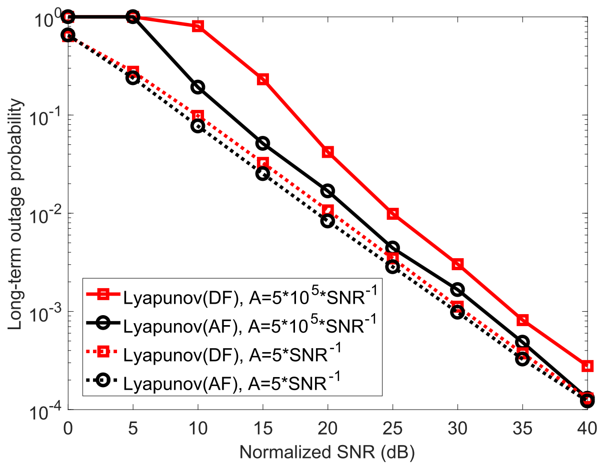


Fig. 4. Comparison of the average outages under different SNRs and the threshold A.

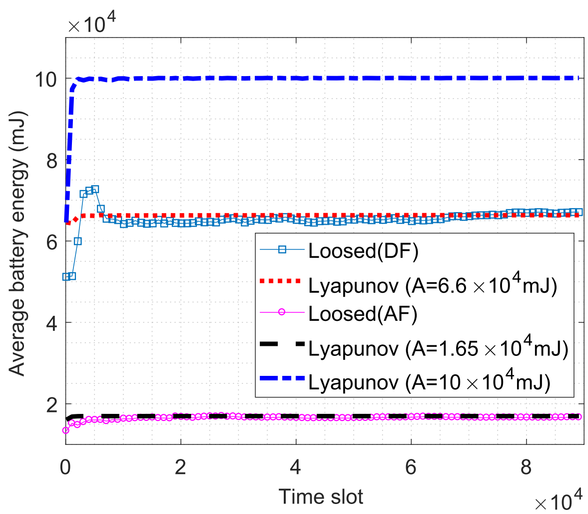


Fig. 5. The evolution of the average battery energy.

see that the energy harvesting gain of the MDP method is much smaller than the proposed methods. One thing we should also notice is that the MDP method does not consider the average battery constraint while the proposed Lyapunov method can guarantee to satisfy the average battery constraint.

The impact of the threshold A on the long-term average outage probability under different SNRs is illustrated in Fig.4. From the figure, we can see that, with both the AF and DF protocols, a larger threshold A will deteriorate the performance of the long-term average outage probability, especially at the low SNR regimes. This is mainly because with a larger threshold A, the average battery constraint is more difficult to satisfy, and thus the performance of the proposed Lyapunov method becomes worse.

The average energy levels stored in the battery of different methods under different protocols are compared in Fig.5. Since the loosed method is not able to guarantee the long-term average battery energy constraint, we set two target battery constraints, where one is near the average battery energy of the loosed method while the other is much larger. The SNR is set to be 0dB. From Fig.5, we can see that the cooperation protocols have an influence on the average battery energy of the loosed one and the average battery energy of the Lyapunov method is near the pre-defined threshold. This validates that the constraint loosed strategy sacrifices the constraint to optimize the outage probabilities.

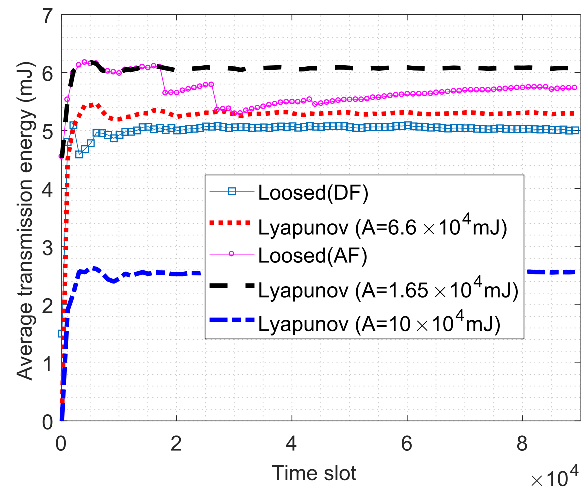


Fig. 6. The evolution of the average energy used for transmission.

Finally, we illustrate in Fig.6 the performance comparison in terms of the corresponding average energy consumed for transmission. From the figure, we can see that when the constraint is smaller, the energy used for transmission consumption of the Lyapunov algorithm and the loosed one becomes lower. Similar to the average battery energy, the cooperation protocols will affect the average energy used for transmission if the constraint is ignored. Moreover, if the average battery energy levels of the Lyapunov strategy and the constraint loosed method are the same, the energy consumption of the latter strategy is smaller.

VI. Conclusion

In this paper, we studied the energy management problem in an energy harvesting two-way cooperative network. The problem is formulated by minimizing the long-term average outage with a long-term average battery constraint. The Lyapunov optimization theory is employed to transform the long-term optimization problem into the drift-plus-penalty. Theoretic analysis in terms of long-term average battery constraint, average outage performance bound, diversity order and energy harvesting gain, is conducted. Simulation results using real solar irradiance data measured by the solar site in Elizabeth City University show that compared with existing MDP-based method, the proposed method can achieve better outage performance and diversity order.

Appendix A
Proof of Theorem 1

A. Decode-and-forward

With (2), the event that the outage would not happen can be written as

$$\bar{\xi}_{DF} = \left[R_1 \leq \frac{1}{2}\log\left(1+\frac{\gamma_1 P_S}{N_0}\right)\right] \cap \left[R_1 \leq \frac{1}{2}\log\left(1+\frac{\gamma_2 P_R}{N_0}\right)\right]$$

$$\cap\left[R_2 \le \frac{1}{2}\log\left(1+\frac{\gamma_2 P_S}{N_0}\right)\right]$$
$$\cap\left[R_2 \le \frac{1}{2}\log\left(1+\frac{\gamma_1 P_R}{N_0}\right)\right]$$
$$\cap\left[R_1+R_2 \le \frac{1}{2}\log\left(1+\frac{\gamma_1 P_S}{N_0}+\frac{\gamma_2 P_S}{N_0}\right)\right]. \quad (33)$$

From (33), we can see that the outage events are determined by P_S and P_R. Specifically, the outage would not happen if P_S and P_R satisfy

$$\left[\frac{\gamma_1(t)P_S}{N_0} \ge 2^{2R_1}-1\right]\cap\left[\frac{\gamma_2(t)P_S}{N_0} \ge 2^{2R_2}-1\right] \cap\left[\frac{(\gamma_1(t)+\gamma_2(t))P_S}{N_0} \ge 2^{(2R_1+2R_2)}-1\right], \quad (34)$$

and

$$\left[P_R \ge \frac{(2^{2R_1}-1)N_0}{\gamma_2(t)}\right]\cap\left[P_R \ge \frac{(2^{2R_2}-1)N_0}{\gamma_1(t)}\right]. \quad (35)$$

With (34) and (35), the energy used for transmission threshold in DF cooperation protocol can be written as (5).

B. Amplify-and-Forward

With (3), we can see that the first condition of preventing outage is

$$\bar{\xi}^A_{AF} = \left\{R_1 \le \frac{1}{2}\log\left[1+\frac{\gamma_1\gamma_2 P_S P_R}{N_0(\gamma_1 P_S+\gamma_2 P_S+\gamma_2 P_R+N_0)}\right]\right\}, \quad (36)$$

which can be rewritten as:

$$\bar{\xi}^A_{AF} = \left[P_R \ge \frac{(2^{2R_1}-1)N_0(\gamma_1 P_S+\gamma_2 P_S+N_0)}{\gamma_1\gamma_2 P_S-(2^{2R_1}-1)N_0\gamma_2}\right] \cap\left[\frac{\gamma_1 P_S}{N_0} > (2^{2R_1}-1)\right]. \quad (37)$$

Similarly, the second condition of preventing outage can be written as

$$\bar{\xi}^B_{AF} = \left[P_R \ge \frac{(2^{2R_2}-1)N_0(\gamma_1 P_S+\gamma_2 P_S+N_0)}{\gamma_1\gamma_2 P_S-(2^{2R_2}-1)N_0\gamma_1}\right] \cap\left[\frac{\gamma_2 P_S}{N_0} > (2^{2R_2}-1)\right]. \quad (38)$$

Since $\bar{\xi}_{AF} = \bar{\xi}^A_{AF}\cap\bar{\xi}^B_{AF}$, with (37) and (38), the energy used for transmission threshold in AF cooperation protocol can be written as (6).

Appendix B
Proof of Theorem 4

1) Decode-and-Forward: With the outage events in (2), we first consider the relaxed conditions where the outage would not happen by ignoring the sum-rate constraint and letting the instantaneous energy used for transmission become the maximum value, i.e., $P_R = b_{max}/T_M$

$$\bar{\xi}^r_{DF} = \left\{\min\left[\frac{1}{2}\log_2\left(1+\frac{\gamma_1 P_S}{N_0}\right), \frac{1}{2}\log_2\left(1+\frac{\gamma_2 b_{max}}{T_M N_0}\right)\right] \ge R_1\right\}$$
$$\cap\left\{\min\left[\frac{1}{2}\log_2\left(1+\frac{\gamma_2 P_S}{N_0}\right), \frac{1}{2}\log_2\left(1+\frac{\gamma_1 b_{max}}{T_M N_0}\right)\right] \ge R_2\right\}. \quad (39)$$

With (39), the relaxed outage probability can be expressed as

$$P^r_{out} = 1-P\left\{\gamma_1 \ge \max\left[(2^{2R_1}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_2}-1)(\frac{b_{max}}{T_M N_0})^{-1}\right]\right\} \times P\left\{\gamma_2 \ge \max\left[(2^{2R_2}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_1}-1)(\frac{b_{max}}{T_M N_0})^{-1}\right]\right\}. \quad (40)$$

Since the channel is assumed to be Rayleigh fading, let $c_1 = \frac{b_{max}}{T_M P_S}$, the outage probability can be written as

$$\begin{aligned} P^r_{out} &= 1-\exp\left\{-\max\left[(2^{2R_1}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_2}-1)(\frac{b_{max}}{T_M N_0})^{-1}\right]\right\} \\ &\quad\times\exp\left\{-\max\left[(2^{2R_2}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_1}-1)(\frac{b_{max}}{T_M N_0})^{-1}\right]\right\} \\ &= 1-\exp\left\{-\max\left[(2^{2R_1}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_2}-1)(\frac{P_S c_1}{N_0})^{-1}\right]\right\} \\ &\quad\times\exp\left\{-\max\left[(2^{2R_2}-1)(\frac{P_S}{N_0})^{-1}, (2^{2R_1}-1)(\frac{P_S c_1}{N_0})^{-1}\right]\right\} \\ &= 1-\exp\left\{-\left\{\max\left[(2^{2R_1}-1),(2^{2R_2}-1)/c_1\right] + \max\left[(2^{2R_2}-1),(2^{2R_1}-1)/c_1\right]\right\}\left(\frac{P_S}{N_0}\right)^{-1}\right\} \\ &= \left(\frac{P_S}{N_0}\right)^{-1}\left\{\max\left[(2^{2R_1}-1),(2^{2R_2}-1)/c_1\right] + \max\left[(2^{2R_2}-1),(2^{2R_1}-1)/c_1\right]\right\} + o\left(\frac{N_0}{P_S}\right). \end{aligned} \quad (41)$$

From (41), we can see that the diversity order of the relaxed outage probability is 1, and the corresponding energy harvesting gain is

$$g^r_{DF} = \left\{\max\left[(2^{2R_1}-1),(2^{2R_2}-1)/c_1\right] + \max\left[(2^{2R_2}-1),(2^{2R_1}-1)/c_1\right]\right\}^{-1}. \quad (42)$$

Then, we consider the strengthened conditions by letting the two transmission rates be larger than R_1+R_2 and assuming the instantaneous energy used for transmission is the minimum value E_{Hmin}. Notice that in such a case the sum-rate constraint is guaranteed to be satisfied and thus can be ignored. The strengthened situations can be written as:

$$\begin{aligned}\bar{\xi}^s_{DF} &= \left\{ \min\left[\frac{1}{2}\log_2\left(1+\frac{\gamma_1 P_S}{N_0}\right), \frac{1}{2}\log_2\left(1+\frac{\gamma_2 E_{Hmin}}{T_M N_0}\right)\right] \geq R_1+R_2 \right\} \\ &\cap \left\{ \min\left[\frac{1}{2}\log_2\left(1+\frac{\gamma_2 P_S}{N_0}\right), \frac{1}{2}\log_2\left(1+\frac{\gamma_1 E_{Hmin}}{T_M N_0}\right)\right] \geq R_1+R_2 \right\}. \quad (43)\end{aligned}$$

Similarly, the outage probability of the strengthened conditions can be written as:

$$\begin{aligned}P^s_{out} &= 1-\exp\left[-2(2^{R_1+R_2}-1)\left(\frac{P_S}{N_0}\right)^{-1}\max\left(1,c_2^{-1}\right)\right] \\ &= 1-\left\{1-\left(\frac{P_S}{N_0}\right)^{-1} 2(2^{R_1+R_2}-1) \times \max\left(1,c_2^{-1}\right)+o\left(\frac{N_0}{P_S}\right)\right\} \\ &= \left(\frac{P_S}{N_0}\right)^{-1} 2(2^{R_1+R_2}-1)\max\left(1,c_2^{-1}\right)+o\left(\frac{N_0}{P_S}\right) \quad (44)\end{aligned}$$

where $c_2 = \frac{E_{Hmin}}{T_M P_S}$.

Hence, the diversity order of the strengthened outage probability is 1 and the corresponding energy harvesting gain is

$$g^s_{DF} = [(2^{R_1+R_2+1}-2)\max(1,c_2^{-1})]^{-1}. \quad (45)$$

In summary, with the proposed algorithm, the diversity order of the DF protocol is 1 and the energy harvesting gain is bounded by

$$\begin{aligned}&\left\{\max\left[(2^{2R_1}-1), \frac{(2^{2R_2}-1)}{c_1}\right] + \max\left[(2^{2R_2}-1), \frac{(2^{2R_1}-1)}{c_1}\right]\right\}^{-1} \\ &< g_{DF} < [(2^{R_1+R_2+1}-2)\max(1,c_2^{-1})]^{-1}. \quad (46)\end{aligned}$$

2) Amplify-and-Forward: With the outage events of AF protocols in (3), we can derive the conditions that the outage would not happen as follows

$$\begin{aligned}\bar{\xi}_{AF} &= \left[(2^{2R_1}-1)^{-1} \geq \frac{N_0}{\gamma_2 P_R}+\frac{N_0}{\gamma_1 P_R}+\frac{N_0}{\gamma_1 P_S}+\frac{N_0^2}{\gamma_1\gamma_2 P_S P_R}\right] \\ &\cap \left[(2^{2R_2}-1)^{-1} \geq \frac{N_0}{\gamma_2 P_R}+\frac{N_0}{\gamma_1 P_R}+\frac{N_0}{\gamma_2 P_S}+\frac{N_0^2}{\gamma_1\gamma_2 P_S P_R}\right]. \quad (47)\end{aligned}$$

Then, the relaxed conditions can be given as

$$\bar{\xi}^r_{AF} = \left[(2^{2R_1}-1)^{-1} \geq \frac{N_0}{\gamma_1 P_S}\right] \cap \left[(2^{2R_2}-1)^{-1} \geq \frac{N_0}{\gamma_2 P_S}\right]. \quad (48)$$

Similar to (41), the relaxed outage probability can then be derived as

$$\begin{aligned}P^r_{out} &= 1-\exp\left[-(2^{2R_1}-1)\left(\frac{P_S}{N_0}\right)^{-1}\right] \times \exp\left[-(2^{2R_2}-1)\left(\frac{P_S}{N_0}\right)^{-1}\right] \\ &= 1-\exp\left[-(2^{2R_1}+2^{2R_2}-2)\left(\frac{P_S}{N_0}\right)^{-1}\right] \\ &= 1-\left\{1-(2^{2R_1}+2^{2R_2}-2)\left(\frac{P_S}{N_0}\right)^{-1}+o\left(\frac{N_0}{P_S}\right)\right\} \\ &= (2^{2R_1}+2^{2R_2}-2)\left(\frac{P_S}{N_0}\right)^{-1}+o\left(\frac{N_0}{P_S}\right). \quad (49)\end{aligned}$$

From (49), we can see that with the related conditions, the diversity order is 1 and the corresponding energy harvesting gain is

$$g^r_{AF} = (2^{2R_1}+2^{2R_2}-2)^{-1}. \quad (50)$$

On the other hand, the strengthened conditions are achieved by increasing the right side and decreasing the left side as follows

$$\bar{\xi}^{s2}_{AF} = \left[(2^{2R_1+2R_2}-1)^{-1} \geq \frac{N_0}{\gamma_2 P_R}+\frac{N_0}{\gamma_1 P_R}+\frac{N_0}{\gamma_1 P_S}+\frac{N_0}{\gamma_2 P_S} + \frac{N_0^2}{\gamma_1\gamma_2 P_S P_R}\right]. \quad (51)$$

Let $c_3 = \frac{P_R}{P_S}$, we further strengthen the conditions as

$$\bar{\xi}^{s3}_{AF} = \left\{(2^{2R_1+2R_2}-1)^{-1} \geq \frac{N_0}{P_S\times\min(c_3,1)}\left(\frac{2}{\gamma_1}+\frac{2}{\gamma_2}\right) + \left[\frac{N_0}{P_S\times\min(c_3,1)}\right]^2\frac{1}{\gamma_1\gamma_2}\right\}. \quad (52)$$

To simplify our analysis, we also introduce the channel power constraints as

$$\bar{\xi}^s_{1,AF} = \left[\gamma_1 \geq \frac{N_0}{P_S\times\min(c_2,1)}\right] \cap \left[\gamma_2 \geq \frac{N_0}{P_S\times\min(c_3,1)}\right]. \quad (53)$$

With the constraint (53), we have

$$\left[\frac{N_0}{P_S\times\min(c_3,1)}\right]^2\frac{1}{\gamma_1\gamma_2} \leq \frac{N_0}{P_S\times\min(c_3,1)}\frac{1}{\gamma_1} \quad (54)$$

and

$$\left[\frac{N_0}{P_S\times\min(c_3,1)}\right]^2\frac{1}{\gamma_1\gamma_2} \leq \frac{N_0}{P_S\times\min(c_3,1)}\frac{1}{\gamma_2}. \quad (55)$$

With (54) and (55), (52) can be rewritten as

$$\begin{aligned}\bar{\xi}^s_{2,AF} &= \left[(2^{2R_1+2R_2}-1)^{-1} \geq 5\times\frac{N_0}{P_S\times\min(c_3,1)\gamma_1}\right] \\ &\cap \left[(2^{2R_1+2R_2}-1)^{-1} \geq 5\times\frac{N_0}{P_S\times\min(c_3,1)\gamma_2}\right]. \quad (56)\end{aligned}$$

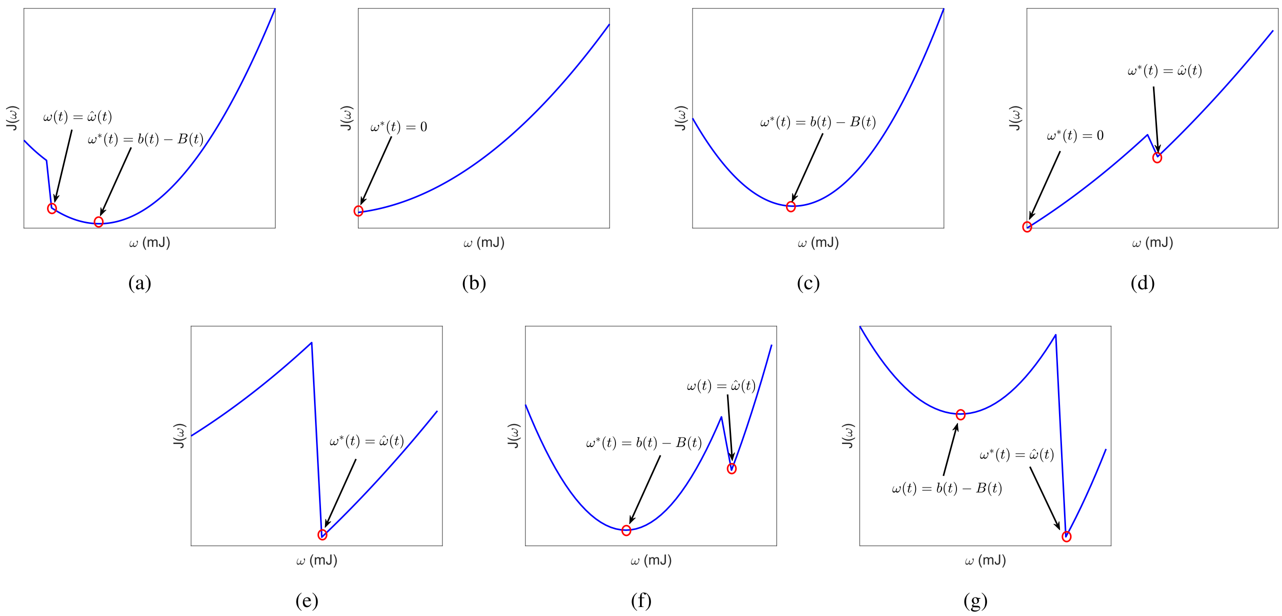


Fig. 7. An illustration of the $J(\omega(t))$: the conditions of achieving the optimal energy used for transmission.

By combing (56) and (53), the final strengthened conditions become

$$\begin{aligned}\bar{\xi}^s_{AF} &= \bar{\xi}^s_{1,AF} \cap \bar{\xi}^s_{2,AF}\\ &= \left[\gamma_1 \geq \frac{5(2^{2R_1+2R_2}-1)N_0}{P_S \times \min(c_3,1)}\right]\\ &\cap \left[\gamma_2 \geq \frac{5(2^{2R_1+2R_2}-1)N_0}{P_S \times \min(c_3,1)}\right]\\ &\cap \left[\gamma_1 \geq \frac{N_0}{P_S \times \min(c_3,1)}\right] \cap \left[\gamma_2 \geq \frac{N_0}{P_S \times \min(c_3,1)}\right]\end{aligned} \tag{57}$$

Therefore, the strengthened outage probability is

$$\begin{aligned}&P^s_{out}\\ &= 1-\exp\left\{-\max\left[\frac{5(2^{2R_1+2R_2}-1)N_0}{P_S\times\min(c_3,1)}, \frac{N_0}{P_S\times\min(c_3,1)}\right]\right\}\\ &\quad \times \exp\left\{-\max\left[\frac{5(2^{2R_1+2R_2}-1)N_0}{P_S\times\min(c_3,1)}, \frac{N_0}{P_S\times\min(c_3,1)}\right]\right\}\\ &= 1-\exp\left\{-2\left(\frac{P_S}{N_0}\right)^{-1}\frac{\max[5(2^{2R_1+2R_2}-1),1]}{\min(c_3,1)}\right\}\\ &= 1-\left\{1-2\frac{\max[5(2^{2R_1+2R_2}-1),1]}{\min(c_3,1)}\left(\frac{P_S}{N_0}\right)^{-1}+o\left(\frac{N_0}{P_S}\right)\right\}\\ &= 2\frac{\max[5(2^{2R_1+2R_2}-1),1]}{\min(c_3,1)}\left(\frac{P_S}{N_0}\right)^{-1}+o\left(\frac{N_0}{P_S}\right)\end{aligned} \tag{58}$$

With (49) and (58), we can see that the diversity order of the AF protocol under the proposed algorithm is 1 and the corresponding energy harvesting gain is bounded by

$$(2^{2R_1}+2^{2R_2}-2)^{-1} < g_{AF} < \frac{\min(c_3,1)}{2\max[5(2^{2R_1+2R_2}-1),1]}. \tag{59}$$

This completes the proof.

Appendix C
Proof of Theorem 5

From (16), we can see that the drift-plus-penalty $J(\omega(t))$ is the sum of a quadratic function and a step function. Thus, the optimal energy used for transmission may be 0, $b(t)-B(t)$ or $\hat{\omega}(t)$. In the following, we discuss the conditions of achieving such optimal energy used for transmission as illustrated in Fig. 7.

- If $b(t)-B(t) > \hat{\omega}(t)$, as illustrated in Fig. 7 (a), in such a case the transition of the step function appears before the minimal point of the quadratic function. Thus, the optimal energy is $\omega^*(t) = b(t) - B(t)$.
- If $\hat{\omega}(t) > b(t)$ and $b(t) \leq B(t)$, as shown in Fig. 7 (b), since $\omega(t) \in [0, b(t)]$, the optimal energy is $\omega^*(t) = 0$.
- If $\hat{\omega}(t) > b(t)$ and $b(t) > B(t)$, as shown in Fig. 7 (c), the optimal energy is the minimal point of the quadratic function, i.e., $\omega^*(t) = b(t) - B(t)$.
- If $\hat{\omega}(t) \leq b(t)$ and $b(t) \leq B(t)$, in such a case, the optimal energy is achieved at either 0 or $\hat{\omega}(t)$. As illustrated in Fig. 7 (d), if $2V < \hat{\omega}^2(t) - 2b(t)\hat{\omega}(t) + 2B(t)\hat{\omega}(t))$, $\omega^*(t) = 0$. On the other hand, if $2V \geq \hat{\omega}^2(t) - 2b(t)\hat{\omega}(t) + 2B(t)\hat{\omega}(t)$, as shown in Fig. 7 (e), $\omega^*(t) = \hat{\omega}(t)$.
- If $\hat{\omega}(t) \leq b(t)$, $b(t) > B(t)$, and $\hat{\omega}(t) > b(t) - B(t)$, in such a case, the optimal energy is achieved at either $b(t)-B(t)$ or $\hat{\omega}(t)$. As illustrated in Fig. 7 (f), if $-[b(t)-B(t)]^2 + 2V < \hat{\omega}^2(t) - 2b(t)\hat{\omega}(t) + 2B(t)\hat{\omega}(t))$, $\omega^*(t) = b(t)-B(t)$. On the other hand, if $-[b(t)-B(t)]^2+2V \geq \hat{\omega}^2(t) - 2b(t)\hat{\omega}(t) + 2B(t)\hat{\omega}(t))$, as shown in Fig. 7 (g), $\omega^*(t) = \hat{\omega}(t)$.

By combining the conditions of achieving the same optimal energy used for transmission, we can obtain the results in Theorem 5.

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] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[3] 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.

[4] 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.

[5] M. R. Javan, N. Mokari, F. Alavi, and A. Rahmati, "Resource allocation in decode-and-forward cooperative communication networks with limited rate feedback channel," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 256–267, Jan. 2017.

[6] G. Zhao, W. Shi, Z. Chen, and Q. Zhang, "Autonomous relaying scheme with minimum user power consumption in cooperative multicast communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2509–2522, Apr. 2016.

[7] Z. Chen, L. X. Cai, Y. Cheng, and H. Shan, "Sustainable cooperative communication in wireless powered networks with energy harvesting relay," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8175–8189, Dec. 2017.

[8] C. Qiu, Y. Hu, and Y. Chen, "Lyapunov optimized cooperative communications with stochastic energy harvesting relay," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1323–1333, Apr. 2018.

[9] W. Li, M.-L. Ku, Y. Chen, and K. J. R. Liu, "On the achievable sum rate for two-way relay networks with stochastic energy harvesting," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Atlanta, GA, USA, Dec. 2014, pp. 133–137.

[10] Q. Li, Q. Zhang, and J. Qin, "Beamforming in non-regenerative two-way multi-antenna relay networks for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5509–5520, Oct. 2014.

[11] 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.

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

[13] 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, Apr. 2015.

[14] 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.

[15] P. Blasco, D. Gündüz, 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.

[16] 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.

[17] 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.

[18] 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.

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synth. Lect. Commun. Netw.*, vol. 3, no. 1, pp. 1–211, 2010.

[28] L. Huang and M. J. Neely, "Utility optimal scheduling in energy-harvesting networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 4, pp. 1117–1130, Aug. 2013.

[29] M. Dong, W. Li, and F. Amirnavaei, "Online joint power control for two-hop wireless relay networks with energy harvesting," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 463–478, Jan. 2018.

[30] X. Wang, Y. Zhang, T. Chen, and G. B. Giannakis, "Dynamic energy management for smart-grid-powered coordinated multipoint systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1348–1359, May 2016.

[31] X. Wang, T. Chen, X. Chen, X. Zhou, and G. B. Giannakis, "Dynamic resource allocation for smart-grid powered MIMO downlink transmissions," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3354–3365, Dec. 2016.

[32] X. Wang, T. Ma, R. Zhang, and X. Zhou, "Stochastic online control for energy-harvesting wireless networks with battery imperfections," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8437–8448, Dec. 2016.

[33] X. Chen, W. Ni, T. Chen, I. B. Collings, X. Wang, and G. B. Guanacos, "Real-time energy trading and future planning for fifth generation wireless communications," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 24–30, Aug. 2017.

[34] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. 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.

[35] J. He, Z. Xue, D. Wu, D. O. Wu, and Y. 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.

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

[37] F. Amirnavaei and M. Dong, "Online power control optimization for wireless transmission with energy harvesting and storage," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4888–4901, Jul. 2016.

[38] D. Zhang, Z. Chen, M. K. Awad, N. Zhang, H. Zhou, and X. S. Shen, "Utility-optimal resource management and allocation algorithm for energy harvesting cognitive radio sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3552–3565, Dec. 2016.

[39] H. Zhang, H. Xing, J. Cheng, A. Nallanathan, and V. C. M. Leung, "Secure resource allocation for OFDMA two-way relay wireless sensor networks without and with cooperative jamming," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1714–1725, Oct. 2016.

[40] J. Huang and H. Gharavi, "Performance analysis of relay-based two-way D2D communications with network coding," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6642–6646, Jul. 2018.

[41] M. Gan, Z. Ding, and X. Dai, "Application of analog network coding to MIMO two-way relay channel in cellular systems," *IEEE Signal Process. Lett.*, vol. 20, no. 7, pp. 641–644, Jul. 2013.

[42] X. Wang, X. Chen, T. Chen, L. Huang, and G. B. Giannakis, "Two-scale stochastic control for integrated multipoint communication systems with renewables," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1822–1834, May 2018.

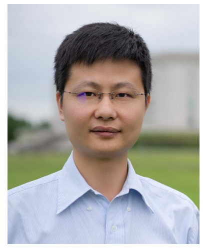
[43] Q. Li, S. H. Ting, A. Pandharipande, and Y. Han, "Adaptive two-way relaying and outage analysis," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3288–3299, Jun. 2009.

[44] X. Lin, M. Tao, Y. Xu, and R. Wang, "Outage probability and finite-SNR diversity–multiplexing tradeoff for two-way relay fading channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3123–3136, Sep. 2013.

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

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

Chengrun Qiu received the B.S. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2018, where he is currently pursuing the master's degree with the School of Information and Communication Engineering. His research interests include reinforcement learning, stochastic optimization and green communications.



Yan Chen (SM'14) received the bachelor's degree from the University of Science and Technology of China in 2004, the M.Phil. degree from The Hong Kong University of Science and Technology 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., 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. His 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 at the PCM in 2017, the Best Student Paper Award at the IEEE ICASSP in 2016, the Best Paper Award at the IEEE GLOBECOM in 2013, the Future Faculty Fellowship and the Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, respectively, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, University of Maryland, in 2011, and the Chinese Government Award for Outstanding Students Abroad in 2010.

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.