# Achievable Rate Region of the Buffer-Aided Two-Way Energy Harvesting Relay Network

Xiaolong Lan, Qingchun Chen, *Senior Member, IEEE*, Xiaohu Tang, *Member, IEEE*, and Lin Cai, *Senior Member, IEEE*

***Abstract*—In this paper, we investigate the buffer-aided two-way wireless energy harvesting (EH) relay network comprising two users and one relay, where two users exchange information-with the help of the relay via three phases of *EH, multiple-access, and broadcast*. By transforming the opportunistic scheduling design problem into an equivalent convex problem, we present the adaptive design for the buffer-aided two-way EH relay network to maximize the long-term achievable rate region. To fulfill the delay sensitive transmission requirements, a delay-aware adaptive transmission (DAAT) scheduling scheme is proposed to guarantee the average end-to-end delivery delay by employing the Lyapunov optimization framework. Our analysis discloses that the average achievable rate region of the two-way wireless EH relay network can be improved when fully considering the potentials by deploying data buffer and energy storage at the relay. There exists an inherent tradeoff among the achievable sum rate, the delivery delay, and the power consumption. It is shown that, when a certain time delay is tolerable, the DAAT scheduling scheme is able to realize a rate region arbitrarily close to the achievable rate region of the buffer-aided two-way wireless EH relay network.**

***Index Terms*—Buffer-aided two-way relay, energy harvesting, achievable rate region.**

## I. Introduction

RECENTLY, as an effective method to prolong the lifetime of energy-constrained networks, energy harvesting based wireless communication has attracted much research attention, in which communication nodes can harvest both energy and information from ambient radio frequency (RF) signals [1], [2]. Since the ambient RF signals can carry energy as well as information at the same time, simultaneous wireless information andpower transfer (SWIPT) has been widely studied in [3]–[9]. The first two practical receiver designs were proposed in [3] to realize SWIPT via either time switching (TS) or power splitting (PS). Some typical transmission design for different wireless network topologies are presented in [4]–[9].

Cooperative communication plays a vital role in improving the throughput/diversity gains, and extending the coverage of wireless networks. However, in the conventional half-duplex relaying protocol, two successive time slots scheduling policy is usually assumed, wherein the relay always receives data in one time slot and forwards it in another. The conventional relaying protocol tends to be inefficient over time-varying channel since the transmission scheduling is independent of the current channel state information (CSI). Recently, buffer-aided relaying was proposed to improve the cooperative relaying networks, and it is shown that using buffer provides a new degree of the freedom to better schedule the relaying transmission. It is disclosed in [10] that the buffer-aided relaying can transform the slow fading two-hop channel into an equivalent end-to-end fast fading channel. In [11], an adaptive link selection scheme was proposed for the buffer-aided relaying network always to select the link with a better CSI. Due to its efficiency, two-way relaying has been considered as an effective relay-assisted solution for the information exchange between two users [12], [13]. It is unveiled in [16], [17] that the buffer-aided two-way relay networks can significantly improve the achievable throughput. Apparently, the system performance of the two-way relay networks with EH nodes will be improved when both data buffer and energy storage are provisioned at the relay. However, previous research efforts of two-way RF-EH relay networks in [18], [19] have been devoted to optimizing either the TS or the PS factor. The influence of buffers was not yet explored. Thus, how to fully exploit the benefits of two-way RF-EH relay networks with data buffer and energy storage and how to design efficient transmission scheme is our primary research motivation in this paper.

Motivated by the enormous potential of energy harvesting technologies and the buffer-enabled scheduling opportunities, a buffer-aided adaptive transmission framework will be presented at first in this paper for the two-way wireless EH relay network, where two users exchange information with the help of the relay via three phases of *EH*, *multiple-access*, and *broadcast*. To reveal the achievable rate region, the opportunistic scheduling design problem is formulated to take into account of transmission mode selection, power allocation, and rate allocation with the peak transmit power and the average transmit power constraints.

The contributions of this paper can be summarized as follows:

- Our analysis discloses that, the average achievable rate region of the two-way wireless EH relay network can be

Manuscript received October 8, 2017; revised June 10, 2018; accepted August 31, 2018. Date of publication September 13, 2018; date of current version November 12, 2018. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61771406, the Natural Sciences and Engineering Research Council of Canada and the Science and Technology Plan of Shenzhen (JCYJ20170818114014753). The work was also partly supported by the Major Frontier Project of Sichuan Province under Grant No. 2015JY0282. *(Corresponding author: Qingchun Chen.)*

X. Lan and X. Tang are with the Southwest Jiaotong University, Chengdu 611756, China (e-mail: xiaolonglan@my.swjtu.edu.cn; xhutang@swjtu.edu.cn).

Q. Chen is with the Guangzhou University, Guangzhou 510006, China (e-mail: qcchen@gzhu.edu.cn).

L. Cai is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: cai@ece.uvic.edu).

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/TVT.2018.2869946

noticeably improved when fully considering the potential by deploying data buffer and energy storage at the relay.
- A delay-aware adaptive transmission scheme was proposed to fulfill the delay sensitive traffic requirements.
- An inherent tradeoff among the achievable rate, delivery delay, and power consumption is disclosed. It is shown that, if a certain time delay is tolerable, the DAAT scheduling scheme is able to realize an average rate region arbitrarily close to the achievable rate region.

The remainder of this paper is organized as follows. In Section II, a literature survey is presented. The system model of the buffer-aided two-way EH relay network is presented in Section III. The achievable rate region analysis and the DAAT scheduling scheme are given in Section IV and Section V, respectively. Simulation results are shown in Section VI. Finally, we conclude this paper in Section VII.

## II. Related Works

The throughput maximization problem of the single user channel with a limited energy storage capacity was studied in [4]. The throughput maximization problem in the classic three-node Gaussian relay model with EH source and relay nodes was addressed in [5] for traffic either with or without delay constraints. SWIPT subject to time-varying co-channel interference was investigated in [6], where the optimal mode switching rule was proposed to realize the rate-energy tradeoff and the outage-energy tradeoff. The multi-user scheduling in RF-EH network was considered in [7], and the "harvest-then-transmit" protocol was recommended, where all users first harvest energy from the downlink broadcast signals, and then use the harvested energy to transmit data to the access point in a time division multiple access manner. It is shown that, by carefully allocating time for downlink wireless power transfer and uplink information transmission, the sum rate for all users can be maximized. A wireless broadcast system with three nodes, where one user harvests energy and another user decodes information from a common transmitter provisioned with multiple antennas was studied in [8], and it was revealed that there exist fundamental tradeoffs between the information rate and the energy harvesting efficiency. Multiple relay network with energy harvesting issue was considered in [9] to show that, for a given energy transfer constraint, a relay selection policy can be utilized to realize the optimal tradeoff in terms of either the maximum capacity or the minimum outage probability.

The time division broadcast (TDBC) and the multiple access broadcast (MABC) are two representative two-way relaying protocols. The TDBC protocol requires three successive phases to complete the information exchange between the source and destination [14]. While it is proposed in the MABC protocol to accomplish the information exchange within two phases, wherein the source-to-relay and the destination-to-relay transmission will be integrated into one multiple access phase [15]. The buffer-aided TDBC two-way relaying protocol was proposed in [16] to maximize the sum-rate via opportunistic scheduling. The transmission mode selection and the power allocation issues were jointly considered in [17] to maximize the achievable rate region. The average delivery delay and the average power consumption tradeoffs in the buffer-aided two-way relay network was addressed in [28].

Recently, two-way relay networks with EH nodes received much attention. TS and power allocation were jointly considered in [18] to maximize the sum rates in two-way RF-EH relay network. A flexible transmission policy for the EH relay with data buffer and energy storage was considered in [20] to maximize the sum rates as well. A directional water-filling algorithm was proposed in [21] to maximize the sum rates of the two-way EH relay networks. An offline optimal joint energy and transmission time allocation scheme was proposed in [22] in the presence of the CSI uncertainty.

Although there are some interesting results about the EH two-way relay network with EH capability, there is no research regarding the achievable rate region of the buffer-aided two-way EH relay network as well as the effective method to approach the rate region. This is the motivation of this paper.

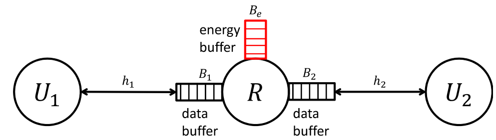


Fig. 1. Two-way wireless energy harvesting relay network consisting of two users $U_1$, $U_2$ and one half-duplex DF relay $R$. It is assumed that there is no direct link between $U_1$ and $U_2$.

## III. System Model and Problem Formulation

### A. Two-Way Wireless EH Relay Network Model

Let us consider a two-way EH relay network, in which user 1 ($U_1$) and user 2 ($U_2$) are supposed to exchange data with the help of a half-duplex decode-and-forward relay $R$, as illustrated in Fig. 1. We assume that the direct link between the two users experience large-scale shadowing such that it does not exist, so a relay is necessary for the system model. Both users and relay are assumed to be equipped with a single antenna. Relay $R$ harvests energy from $U_1$ and $U_2$, and the harvested energy is utilized to either receive data from or forward data to $U_1$ and $U_2$. $R$ is provisioned with one energy storage $B_e$ and two data buffers $B_1$ and $B_2$. $B_e$ is used to store the collected energy from both users, while $B_1$ and $B_2$ are employed to buffer data from $U_1$ and $U_2$, respectively. In this paper, we assume that the energy storage size and both data buffer sizes are large enough such that no overflow happens.[1] Let $E(t)$ denote the amount of energy available in energy buffer $B_e$ in the $t$-th time slot, $Q_1(t)$ and $Q_2(t)$ denote the amount of data in $B_1$ and $B_2$ within the $t$-th time slot, respectively. The block fading channel model is assumed such that the channel coefficients are constant within each time slot but change independently from one to another. It is assumed that all the involved channels are reciprocal. Meanwhile, perfect CSI is assumed,[2] where $h_1(t)$, $h_2(t)$ stands for the channel coefficients of the links $U_1 \leftrightarrow R$ and $U_2 \leftrightarrow R$ within the $t$-th time slot, respectively. Let $\Omega_1 = \mathbb{E}[|h_1(t)|^2]$, $\Omega_2 = \mathbb{E}[|h_2(t)|^2]$, where $\mathbb{E}[\cdot]$ denotes the statistical expectation. The fading gains of $|h_1(t)|^2$ and $|h_2(t)|^2$ are assumed to follow the exponential distribution with parameters $\frac{1}{\Omega_1}$ and $\frac{1}{\Omega_2}$, respectively.

For the two-way EH relay network, three transmission modes are assumed: (i) *EH mode* in which the relay node harvests

[1] In [30], it is shown that, the practical buffer occupancy will be less than the pre-defined control parameter $V$. Thus it is reasonable to assume no overflow if the storage size is larger than $V$, which can be adjusted for different traffic delivery delay requirements.

[2] Since our focus is the achievable rate region analysis, we do not consider the channel estimation problem in this paper. As for the practical channel estimate design issue in the decode-and-forward relaying system, the readers may refer to the related literature, e.g., [27].

energy from the received RF signals of two users, (ii) *multiple-access mode* in which both users transmit their data to the relay simultaneously[3] and relay decodes the received information and store them in the corresponding buffer, and (iii) *broadcast mode* in which the relay forwards the multiplexed data from both buffers back to two users by using the harvested energy. In particular, in the *broadcast mode*, a special broadcast channel with side information is created, in which the relay broadcasts data generated from messages of both data buffers [23]. To characterize the transmission mode selection, binary variables of $\{q_i(t) \in \{0,1\},\ i=1,2,3\}$ are introduced to indicate whether or not the corresponding transmission mode is selected in the $t$-th time slot. $q_i(t)=1$ if the $i$-th transmission mode is selected in the $t$-th time slot, otherwise $q_i(t)=0$. Furthermore, since only one mode can be selected at each time slot, $q_1(t)+q_2(t)+q_3(t)=1,\ \forall t$. We denote $R_{ir}(t)$[4] and $R_{ri}(t) (i=1,2)$ the transmission rate from $U_i$ to the relay and the transmission rate from the relay to $U_i$ in the $t$-th time slot, respectively.

When both $U_1$ and $U_2$ transmit simultaneously to the relay, the received signal at the relay in the $t$-th time slot is

$$Y_r(t) = \sqrt{\frac{P_1(t)}{d_1^m}} h_1(t)x_1(t) + \sqrt{\frac{P_2(t)}{d_2^m}} h_2(t)x_2(t) + n(t), \quad (1)$$

where $d_i (i=1,2)$ denotes the distance between $U_i$ and relay, $m$ is the path loss exponent, $x_i(t)$ stands for the transmitted signal by $U_i$, $\mathbb{E}\{|x_i(t)|^2\}=1$, $P_1(t)$ and $P_2(t)$ denotes the transmit power at $U_1$ and $U_2$, respectively. $n(t)$ represents the additive Gaussian noise with zero mean and variance $\sigma^2$.

1) $\mathcal{M}_1$ *(EH Mode):* In this mode, $Y_r(t)$ is used for energy harvesting, and the amount of harvested energy is

$$E_h(t) = q_1(t)\left(\frac{P_1(t)|h_1(t)|^2}{d_1^m} + \frac{P_2(t)|h_2(t)|^2}{d_2^m}\right)\eta T, \quad (2)$$

where $\eta \in [0,1]$ denotes the energy conversion efficiency, $T$ is the duration of each time slot. Moreover, the harvested energy will be stored in $B_e$, which can be updated as

$$E(t) = E(t-1) + E_h(t). \quad (3)$$

2) $\mathcal{M}_2$ *(Multiple-Access Mode):* In this mode, the relay will try to extract data of $U_1$ and $U_2$ from $Y_r(t)$ by employing successive interference cancellation. Therefore, the achievable rate of $R_{1r}(t)$ and $R_{2r}(t)$ in the $t$-th time slot for a successful transmission must satisfy the following multiple-access capacity region constraints [17]

$$R_{ir}(t) \le q_2(t)C(P_i(t)s_i(t)),\ i=1,2, \quad (4a)$$

$$R_{1r}(t) + R_{2r}(t) \le q_2(t)C(P_1(t)s_1(t) + P_2(t)s_2(t)), \quad (4b)$$

where $C(x)=\log_2(1+x)$ and $s_i(t)=\frac{|h_i(t)|^2}{d_i^m\sigma^2} (i=1,2)$. Please note that the special case that one user keeps silent while the other user transmits data to $R$ (i.e., the traditional point-to-point transmission) can be subsumed by the above capacity region of the multiple-access channel.

After the relay has successfully restored data from both users, the queue length in $B_1$ and $B_2$ will be updated as

$$Q_i(t) = Q_i(t-1) + R_{ir}(t), i=1,2. \quad (5)$$

3) $\mathcal{M}_3$ *(Broadcast Mode):* In this mode, the relay extracts the message intended for $U_2$ from data buffer $B_1$ and the message intended for $U_1$ from data buffer $B_2$, respectively. By using the coding scheme recommended in [23], the relay generates a coded data $x_r(t)$, which is a function of two codewords that correspond to the messages extracted at $B_1$ and $B_2$, while the transmit power is supplied by the collected energy in $B_e$. Therefore, the received signals at $U_1$ and $U_2$ are given by

$$Y_i(t) = \sqrt{\frac{P_r(t)}{d_i^m}} h_i(t)x_r(t) + n(t),\ i=1,2, \quad (6)$$

where $P_r(t)$ denotes the transmit power at the relay. It is assumed that $\mathbb{E}\{|x_r(t)|^2\}=1$. Since each user knows its own data, the maximal broadcast rate at each link equals to the maximal achievable rate when only that individual link presents, and thus we have [17]

$$R_{r1}(t) \le q_3(t)\min\{C(P_r(t)s_1(t)), Q_2(t-1)\}, \quad (7a)$$

$$R_{r2}(t) \le q_3(t)\min\{C(P_r(t)s_2(t)), Q_1(t-1)\}. \quad (7b)$$

Both data and energy buffers at the relay will be updated as

$$Q_1(t) = Q_1(t-1) - R_{r2}(t), \quad (8a)$$

$$Q_2(t) = Q_2(t-1) - R_{r1}(t), \quad (8b)$$

$$E(t) = E(t-1) - q_3(t)P_r(t)T. \quad (8c)$$

### B. Problem Formulation

In this paper, our objective is to determine the achievable rate region of the buffer-aided two-way EH relay networks with the long-term average and the peak transmit power constraints. Most of the existing analysis in the two-way relay network assumes a fixed scheduling policy to select the transmission mode. However, when the relay is provisioned with the data buffer and its transmit energy is supplied via wireless energy transfer, now the transmission mode can be adaptively selected based on the underlying channel state information (CSI), the data buffer state information (BSI) and energy storage information (ESI). Moreover, the optimal power allocations, rate allocations, and the transmission mode selection are jointly considered for the sake of maximizing the long-term achievable rate region.

For the power allocations, the transmit power at both users and relay are now adjusted based on the underlying CSIs, BSIs, and ESI. Here the following two transmit power constraints should be satisfied: $i$). the long-term transmit power constraints at $U_1$ and $U_2$, $ii$). the peak transmit power constraint at $U_1$, $U_2$ and $R$ within each time slot, i.e.,

$$0 \le P_i(t) \le \hat{P}_i, i \in \{1,2,r\}, \forall t, \quad (9a)$$

$$\bar{P}_i = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \big(q_1(t)+q_2(t)\big)P_i(t) \le \bar{P}_i^{\max},\ i=1,2, \quad (9b)$$

where $\hat{P}_i$ denotes the peak transmit power at node $i \in \{1,2,r\}$, $\bar{P}_i$ and $\bar{P}_i^{\max}$ denotes the time average transmit power and the

[3]We don't consider the timing synchronization issue in this paper. For the practical timing synchronization issue, the readers may refer to the related literature, e.g., [31].

[4]Bit per time slot is assumed in all the rate and capacity in this paper.

maximum allowed time average transmit power budget at $U_i$, respectively. Moreover, the average energy consumed by the relay should not exceed the average energy that it can harvest, i.e.,

$$\bar{P}_r = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} q_3(t) P_r(t) \le \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \frac{E_h(t)}{T}. \quad (9c)$$

In particular, we set $\hat{P}_r = \frac{E(t)}{T}$ such that the energy consumed in each time slot does not exceed the energy in the energy storage. Let $\bar{R}_{1r}, \bar{R}_{r2}, \bar{R}_{2r}$ and $\bar{R}_{r1}$ denote the long-term average transmission rates from user 1-to-relay, relay-to-user 2, user 2-to-relay and relay-to-user 1, *i.e.*,

$$\bar{R}_i = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} R_i(t), \quad i \in \{1r, r2, 2r, r1\}. \quad (10)$$

The transmit rate pair $(R_{1r}(t), R_{2r}(t))$ must satisfy the multiple-access capacity region and the transmit rate pair $(R_{r1}(t), R_{r2}(t))$ has to satisfy the broadcast capacity with side information. Furthermore, Let $\bar{R}_{12}$ and $\bar{R}_{21}$ denotes the average transmission rates from user 1-to-user 2 and user 2-to-user 1, respectively. Obviously, $\bar{R}_{12} = \min\{\bar{R}_{1r}, \bar{R}_{r2}\}, \bar{R}_{21} = \min\{\bar{R}_{2r}, \bar{R}_{r1}\}$.

We focus on the long-term achievable rate region for the energy harvesting two-way relay network, and the maximum long-term sum rate can be obtained when the queues in $B_1$ and $B_2$ are at the edge of non-absorbtion [24], i.e.,

$$\bar{R}_{1r} = \bar{R}_{r2}, \quad \bar{R}_{2r} = \bar{R}_{r1}. \quad (11)$$

As shown in [17], the long-term achievable rate region can be determined by maximizing the following weighted long-term sum rate problem for all weighting coefficients $\theta \in [0, 1]$ of $\mathbf{P1}$ for the optimization variables of $\mathbf{q(t)} = [q_1(t), q_2(t), q_3(t)]$, $\mathbf{P(t)} = [P_1(t), P_2(t), P_r(t)]$ and $\mathbf{R(t)} = [R_{1r}(t), R_{r2}(t), R_{2r}(t), R_{r1}(t)]$

$$\begin{aligned} \mathbf{P1}: \quad & \max_{\mathbf{q(t)},\mathbf{P(t)},\mathbf{R(t)}} \theta \bar{R}_{1r} + (1-\theta)\bar{R}_{2r} \\ & \text{s.t.} \quad (4), (7), (9), (11), \\ & \quad q_1(t) + q_2(t) + q_3(t) = 1, \ \forall t, \\ & \quad q_i(t)(q_i(t) - 1) = 0, \ i = 1, 2, 3 \ \ \forall t. \end{aligned}$$

Unfortunately the optimization problem $\mathbf{P1}$ is mixed-integer non-convex since $F_1(x, y) = yC(x) = y \log_2(1 + x)$ and $F_2(x, y) = xy$ are not convex for $\{(x, y)|x \ge 0, y \ge 0\}$, and $\{q_i(t), i = 1, 2, 3\}$ are binary integer variables. In Section IV, we will transform $\mathbf{P1}$ into a standard convex optimization problem by introducing some new variables.

## IV. The Optimal Adaptive Relaying Design

Since the optimization problem $\mathbf{P1}$ is non-convex, we propose to introduce new variables of $E_i^{M_1}(t) = q_1(t)P_i(t)$, $E_i^{M_2}(t) = q_2(t)P_i(t)$, $i = 1, 2$, and $E_r(t) = q_3(t)P_r(t)$. In fact, $E_i^{M_1}(t)$ and $E_i^{M_2}(t)$ represents the energy consumed by $U_i$ in the EH mode ($\mathcal{M}_1$) and the multiple-access mode ($\mathcal{M}_2$) at slot $t$, respectively, and $E_r(t)$ stands for the energy consumed by the relay in the broadcast mode ($\mathcal{M}_3$) at slot $t$. Let $\mathbf{E(t)} = [E_1^{M_1}(t), E_2^{M_1}(t), E_1^{M_2}(t), E_2^{M_2}(t), E_r(t)]$ denote the collection of the new optimization variables. Therefore, by substituting $\mathbf{E(t)}$ into (4), the multiple-access capacity region can be reformulated as

$$R_{ir}(t) \le C_{ir}(t), \ i = 1, 2, \quad (12a)$$

$$R_{1r}(t) + R_{2r}(t) \le C_r(t), \quad (12b)$$

where $C_{ir}(t) = q_2(t)C(E_i^{M_2}(t)s_i(t)/q_2(t))$, $i = 1, 2$ and $C_r(t) = q_2(t)C((E_1^{M_2}(t)s_1(t) + E_2^{M_2}(t)s_2(t))/q_2(t))$.

Based on the optimal queue condition [24], as $N \to \infty$, the effect of data queue at the relay can be neglected, i.e., the relay always has enough data to be delivered to both users. Thus the broadcast capacity region with side information can be rewritten as

$$R_{ri}(t) \le C_{ri}(t), \ i = 1, 2, \quad (13)$$

where $C_{ri}(t) = q_3(t)C(E_r(t)s_i(t)/q_3(t))$. Similarly, by substituting $\mathbf{E(t)}$ into (9), the time average transmit power constraint and the peak transmit power constraint can be reformulated as

$$0 \le E_i^{M_1}(t) \le q_1(t)\hat{P}_i, \ i = 1, 2, \ \forall t, \quad (14a)$$

$$0 \le E_i^{M_2}(t) \le q_2(t)\hat{P}_i, \ i = 1, 2, \ \forall t, \quad (14b)$$

$$0 \le E_r(t) \le q_3(t)\hat{P}_r, \ \forall t, \quad (14c)$$

$$\bar{E}_i = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \left(E_i^{M_1}(t) + E_i^{M_2}(t)\right) \le \bar{P}_i^{\max}, i = 1, 2, \quad (14d)$$

$$\bar{E}_r = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E_r(t) \le \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \frac{E_h(t)}{T}. \quad (14e)$$

Now we relax $q_i(t) \in [0, 1]$, and the constraint $q_i(t)(q_i(t) - 1) = 0$ can be replaced by $q_i(t)(q_i(t) - 1) \le 0$. $\mathbf{P1}$ can be reformulated as the following optimization problem of $\mathbf{P2}$

$$\begin{aligned} \mathbf{P2}: \quad & \max_{\mathbf{q(t)},\mathbf{R(t)},\mathbf{E(t)}} \theta \bar{R}_{1r} + (1-\theta)\bar{R}_{2r} \\ & \text{s.t.} \ (11), (12), (13), (14), \\ & \quad q_1(t) + q_2(t) + q_3(t) = 1, \ \forall t, \\ & \quad q_i(t)(q_i(t) - 1) \le 0, \ i = 1, 2, 3 \, \forall t. \end{aligned}$$

Note that the right-hand sides in (12) and (13) are of the general form $F_3(x, y) = yC(x/y)$. Moreover, $-F_3(x, y)$ is the perspective of the convex function $-C(x)$, which is convex since the perspective operation preserves convexity [25]. Therefore, with linear objective function as well as convex and linear constraints, now the optimization problem $\mathbf{P2}$ is convex for $\mathbf{q(t)}, \mathbf{R(t)}, \mathbf{E(t)}$.

In the following, we will derive the optimal $\mathbf{q}^*\mathbf{(t)}, \mathbf{R}^*\mathbf{(t)}$ and $\mathbf{E}^*\mathbf{(t)}$. However, before formally stating the optimal power allocation, the optimal rate allocation, as well as the optimal mode selection scheme, we introduce the utilized Lagrange multipliers at first. Let $\mu_{1r}(t), \mu_{r2}(t), \mu_{2r}(t), \mu_{r1}(t)$, and $\mu_r(t)$ denote the Lagrange multipliers associated with the transmission rate constraints $R_{1r}(t) \le C_{1r}(t), R_{r2}(t) \le C_{r2}(t), R_{2r}(t) \le C_{2r}(t), R_{r1}(t) \le C_{r1}(t)$, and $R_{1r}(t) + R_{2r}(t) \le C_r(t)$, respectively. We denote the Lagrange multipliers associated with the peak power constraints (14a), (14b) and (14c) by $\lambda_i^{M_1}(t), \lambda_i^{M_2}(t) (i = 1, 2)$ and $\lambda_r^{M_3}(t)$, respectively. Let $\alpha_1$ and $\alpha_2$ denote the Lagrange multipliers associated with the constraints (11) in the optimization problem $\mathbf{P2}$, respectively. We denote the Lagrange multipliers associated with the long-term average power constraint of $U_1$, $U_2$, and $R$ by $\gamma_1$, $\gamma_2$, and

$\gamma_r$, respectively. Furthermore, the values of $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \gamma_r]$ only depend on the statistical CSIs, power budgets, and the value of $\theta$. Based on the optimization problem $\mathbf{P2}$, we have Theorem 1.

*Theorem 1:* Given the long-term average transmit power and the peak transmit power constraints, the optimal power allocation, the optimal rate allocation, and the optimal mode selection in terms of the maximal achievable weighted sum rate for the buffer-aided two-way EH relay networks are determined by the following policy.[5]

1) The optimal mode selection policy

$$q_i^*(t) = \begin{cases} 1, & \text{if } i = \arg\max_{i=1,2,3} \Lambda_i(t), \\ 0, & \text{otherwise}, \end{cases} \tag{15}$$

where $\Lambda_1(t)$, $\Lambda_2(t)$ and $\Lambda_3(t)$ are the corresponding mode selection metrics, and

$$\Lambda_1(t) = \lambda_1^{M_1}(t)\hat{P}_1 + \lambda_2^{M_1}(t)\hat{P}_2, \tag{16a}$$

$$\begin{aligned} \Lambda_2(t) = \sum_{i=1}^{2} \Big(&\mu_{ir}(t)C(P_i(t)s_i(t)) \\ &- \frac{\mu_{ir}(t)P_i(t)s_i(t)}{(1+P_i(t)s_i(t))\ln 2}\Big) \\ &- \frac{\mu_r(t)(P_1(t)s_1(t)+P_2(t)s_2(t))}{(1+P_1(t)s_1(t)+P_2(t)s_2(t))\ln 2} \\ &+ \mu_r(t)C(P_1(t)s_1(t)+P_2(t)s_2(t)) \\ &+ \lambda_1^{M_2}(t)\hat{P}_1 + \lambda_2^{M_2}(t)\hat{P}_2\Big|_{P_1(t)=P_1^{M_2^*}(t)}^{P_2(t)=P_2^{M_2^*}(t)}, \end{aligned} \tag{16b}$$

$$\begin{aligned} \Lambda_3(t) = \sum_{i=1}^{2} \Big(&\mu_{ri}(t)C(P_r(t)s_i(t)) \\ &- \frac{\mu_{ri}(t)P_r(t)s_i(t)}{(1+P_r(t)s_i(t))\ln 2}\Big) \\ &+ \lambda_r^{M_3}(t)\hat{P}_r\Big|_{P_r(t)=P_r^*(t)}, \end{aligned} \tag{16c}$$

where $\boldsymbol{\mu}(t) = [\mu_{1r}(t), \mu_{r2}(t), \mu_{2r}(t), \mu_{r1}(t), \mu_r(t)]$ and $\boldsymbol{\lambda}(t) = [\lambda_1^{M_1}(t), \lambda_1^{M_2}(t), \lambda_2^{M_1}(t), \lambda_2^{M_2}(t), \lambda_r^{M_3}(t)]$ are given by

$$\begin{aligned} &(\mu_{1r}(t), \mu_{2r}(t), \mu_r(t)) \\ &= \begin{cases} (0, \mu_2-\mu_1, \mu_1), & \text{if } 0<\mu_1\le\mu_2, \\ (\mu_1-\mu_2, 0, \mu_2), & \text{if } 0<\mu_2\le\mu_1, \\ (\mu_1, 0, 0), & \text{if } \mu_1>0 \wedge \mu_2\le 0, \\ (0, \mu_2, 0), & \text{if } \mu_1\le 0 \wedge \mu_2<0, \\ (0,0,0), & \text{if } \mu_1<0 \wedge \mu_2<0, \end{cases} \end{aligned}$$

$$(\mu_{r1}(t), \mu_{r2}(t)) = (\alpha_2, \alpha_1),$$

$$\lambda_i^{M_1}(t) = \begin{cases} 0, & \text{if } P_i^{M_1}(t) < \hat{P}_i, \\ \left(\frac{\gamma_r|h_i(t)|^2\eta}{d_i^m} - \gamma_i\right)^+, & \text{otherwise}, \end{cases}$$

$$\lambda_i^{M_2}(t) = \begin{cases} 0, & \text{if } P_i^{M_2}(t) < \hat{P}_i, \\ \Big(\frac{\mu_r(t)s_i(t)}{(1+P_1^{M_2}(t)s_1(t)+P_2^{M_2}(t)s_2(t))\ln 2} \\ \quad + \frac{\mu_{ir}(t)s_i(t)}{(1+P_i^{M_2}(t)s_i(t))\ln 2} - \gamma_i\Big)^+, & \text{otherwise}, \end{cases}$$

$$\lambda_r^{M_3}(t) = \begin{cases} 0, & \text{if } P_r(t) < \hat{P}_r, \\ \left(\sum_{i=1}^{2} \frac{\mu_{ri}(t)s_i(t)}{(1+P_r(t)s_i(t))\ln 2} - \gamma_r\right)^+, & \text{otherwise}, \end{cases}$$

where $\wedge$ represents logical AND and $(x)^+ = \max\{x, 0\}$, $\mu_1 = \theta - \alpha_1$ and $\mu_2 = 1 - \theta - \alpha_2$. The values of $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \gamma_r]$ are presented in Proposition 1.

2) The optimal rate allocation policy in both $\mathcal{M}_2$ and $\mathcal{M}_3$:

$$R_{1r}^*(t) = \begin{cases} C(P_1(t)s_1(t)), & \text{if } \mu_{1r}(t)>0 \wedge \mu_{2r}(t)=0, \\ C\left(\frac{P_1(t)s_1(t)}{1+P_2(t)s_2(t)}\right), & \text{otherwise}, \end{cases} \tag{17a}$$

$$R_{2r}^*(t) = \begin{cases} C(P_2(t)s_2(t)), & \text{if } \mu_{2r}(t)>0 \wedge \mu_{1r}(t)=0, \\ C\left(\frac{P_2(t)s_2(t)}{1+P_1(t)s_1(t)}\right), & \text{otherwise}, \end{cases} \tag{17b}$$

$$R_{r1}^*(t) = \min\{C(P_r(t)s_1(t)), Q_2(t)\}, \tag{17c}$$

$$R_{r2}^*(t) = \min\{C(P_r(t)s_2(t)), Q_1(t)\}. \tag{17d}$$

3) The optimal power allocation policy in $\mathcal{M}_1$:

$$P_i^{M_1^*}(t) = \begin{cases} \hat{P}_i, & \text{if } \frac{\gamma_r|h_i(t)|^2\eta}{d_i^m} - \gamma_i \ge 0, \\ 0, & \text{otherwise}. \end{cases} \tag{18}$$

4) The optimal power allocation policy for the two users in $\mathcal{M}_2$: See (19) at the bottom of the next page.

5) The optimal power allocation policy for the relay in $\mathcal{M}_3$:

$$P_r^*(t) = \begin{cases} 0, & \text{if } \mu_{r1}(t)s_1(t)+\mu_{r2}(t)s_2(t) \le \gamma_r\ln 2, \\ \hat{P}_r, & \text{if } \frac{\mu_{r1}(t)s_1(t)}{1+\hat{P}_r s_1(t)} + \frac{\mu_{r2}(t)s_2(t)}{1+\hat{P}_r s_2(t)} \ge \gamma_r \ln 2, \\ \frac{-b+\sqrt{b^2-4ac}}{2a}, & \text{otherwise}, \end{cases} \tag{20}$$

where $a = s_1(t)s_2(t)\gamma_r\ln 2$, $b = (s_1(t)+s_2(t))$, and $c = \gamma_r\ln 2 - \mu_{r1}(t)s_1(t) - \mu_{r2}(t)s_2(t)$.

*Proof:* See Appendix A. ∎

In Theorem 1, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ corresponds to the time average rate and power constraints, respectively. Both $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ only depend on the underlying statistical CSIs, the power budget, and the value of $\theta$, as illustrated in the following proposition.

*Proposition 1:* The $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ in Theorem 1 can be obtained by employing the gradient descent method [25]. In each time

[5]Here we drop time index $t$ for brevity.

slot, the estimates of $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\gamma}(t)$ can be updated as below

$$\alpha_1(t+1) = \alpha_1(t) + \frac{\delta_{12}(t)}{t}\sum_{i=0}^{t}(R_{1r}(i) - R_{r2}(i)), \quad (21a)$$

$$\alpha_2(t+1) = \alpha_2(t) + \frac{\delta_{21}(t)}{t}\sum_{i=0}^{t}(R_{2r}(i) - R_{r1}(i)), \quad (21b)$$

$$\gamma_j(t+1) = \gamma_j(t) + \frac{\delta_j(t)}{t}\left(\sum_{i=0}^{t}P_j(i) - \bar{P}_j^{\max}\right), j = 1, 2, r, \quad (21c)$$

where $\delta_{12}(t)$, $\delta_{21}(t)$ and $\delta_j(t)$ are the utilized step sizes in the $t$-th time slot, which will control the speed of convergence of $\boldsymbol{\alpha}$

$$(P_1^{M_2^*}, P_2^{M_2^*}) = \begin{cases} (0,0), & \text{if } f_1 \le 0 \wedge f_2 \le 0 \\ \left(0, \frac{\mu_{2r}+\mu_r}{\gamma_2 \ln 2} - \frac{1}{s_2}\right), & \text{if } f_2 > 0 \wedge f_4 < 0 \wedge f_3 \ge 0 \\ (0, \hat{P}_2), & \text{if } f_4 \ge 0 \wedge f_8 \le 0 \\ \left(\frac{\mu_{1r}+\mu_r}{\gamma_1 \ln 2} - \frac{1}{s_1}, 0\right), & \text{if } f_1 > 0 \wedge f_6 < 0 \wedge f_5 \ge 0 \\ (\hat{P}_1, 0), & \text{if } f_6 \ge 0 \wedge f_{10} \le 0 \\ \left(\frac{\mu_{1r}s_2}{(\gamma_1 s_2 - \gamma_2 s_1)\ln 2} - \frac{1}{s_1}, \frac{\mu_r}{\gamma_2 \ln 2} - \frac{\mu_{1r}s_1}{(\gamma_1 s_2(t) - \gamma_2 s_1)\ln 2}\right), & \text{if } \mu_{1r} > 0 \wedge f_5 < 0 \wedge f_3 < 0 \wedge f_{11} < 0 \wedge f_{12} < 0 \\ \left(\frac{\mu_r}{\gamma_1 \ln 2} - \frac{\mu_{2r}s_2}{(\gamma_2 s_1 - \gamma_1 s_2)\ln 2}, \frac{\mu_{2r}s_1}{(\gamma_2 s_1 - \gamma_1 s_2)\ln 2} - \frac{1}{s_2}\right), & \text{if } \mu_{1r} = 0 \wedge f_5 < 0 \wedge f_3 < 0 \wedge f_{13} < 0 \wedge f_{14} < 0 \\ \left(\hat{P}_1, \frac{-b_1 + \sqrt{b_2^2 - 4a_2c_2}}{2a_2} - \frac{1}{s_1}\right), & \text{if } \mu_{1r} = 0 \wedge f_{10} > 0 \wedge f_9 < 0 \wedge f_{14} \ge 0 \\ \left(\frac{-b_2 + \sqrt{b_1^2 - 4a_1c_1}}{2a_1} - \frac{1}{s_1}, \hat{P}_2\right), & \text{if } \mu_{1r} > 0 \wedge f_8 < 0 \wedge f_9 < 0 \wedge f_{12} \ge 0 \\ \left(\hat{P}_1, \frac{\mu_r}{\gamma_2 \ln 2} - \frac{1 + \hat{P}_1 s_1}{s_2}\right), & \text{if } \mu_{1r} > 0 \wedge f_{10} > 0 \wedge f_7 < 0 \wedge f_{11} \ge 0 \\ \left(\frac{\mu_r}{\gamma_1 \ln 2} - \frac{1 + \hat{P}_2 s_2}{s_1}, \hat{P}_2\right), & \text{if } \mu_{1r} = 0 \wedge f_8 > 0 \wedge f_9 < 0 \wedge f_{13} \ge 0 \\ (\hat{P}_1, \hat{P}_2), & \text{if } f_9 \ge 0 \wedge f_7 \ge 0 \end{cases}$$

$$f_1 = (\mu_{1r} + \mu_r)s_1 - \gamma_1 \ln 2, \qquad f_2 = (\mu_{2r} + \mu_r)s_2 - \gamma_2 \ln 2,$$

$$f_3 = (\mu_{2r} + \mu_r)s_2(\gamma_1 \ln 2 - \mu_{1r}s_1) - \mu_r s_1 \gamma_2 \ln 2, \qquad f_4 = \frac{(\mu_{2r} + \mu_r)s_2}{1 + \hat{P}_2 s_2} - \gamma_2 \ln 2,$$

$$f_5 = (\mu_{1r} + \mu_r)s_1(\gamma_2 \ln 2 - \mu_{2r}s_2) - \mu_r s_2 \gamma_1 \ln 2, \qquad f_6 = \frac{(\mu_{1r} + \mu_r)s_1}{1 + \hat{P}_1 s_1} - \gamma_1 \ln 2,$$

$$f_7 = \frac{\mu_{2r}s_2}{1 + \hat{P}_2 s_2} + \frac{\mu_r s_2}{1 + \hat{P}_1 s_1 + \hat{P}_2 s_2} - \gamma_2 \ln 2, \qquad f_8 = \mu_{1r}s_1 + \frac{\mu_r s_1}{1 + \hat{P}_2 s_2} - \gamma_1 \ln 2,$$

$$f_9 = \frac{\mu_{1r}s_1}{1 + \hat{P}_1 s_1} + \frac{\mu_r s_1}{1 + \hat{P}_1 s_1 + \hat{P}_2 s_2} - \gamma_1 \ln 2, \qquad f_{10} = \mu_{2r}s_2 + \frac{\mu_r s_2}{1 + \hat{P}_1 s_1} - \gamma_2 \ln 2,$$

$$f_{11} = \frac{\mu_{1r}s_2}{(\gamma_1 s_2 - \gamma_2 s_1)\ln 2} - \frac{1 + \hat{P}_1 s_1}{s_1}, \qquad f_{12} = \frac{\mu_r}{\gamma_2 \ln 2} - \frac{\mu_{1r}s_1}{(\gamma_1 s_2 - \gamma_2 s_1)\ln 2} - \hat{P}_2.$$

$$f_{13} = \frac{\mu_{2r}s_1}{(\gamma_2 s_1 - \gamma_1 s_2)\ln 2} - \frac{1 + \hat{P}_2 s_2}{s_2}, \qquad f_{14} = \frac{\mu_r}{\gamma_1 \ln 2} - \frac{\mu_{2r}s_2}{(\gamma_2 s_1 - \gamma_1 s_1)\ln 2} - \hat{P}_1,$$

$$a_1 = \gamma_1 \ln 2, \quad b_1 = \hat{P}_2 s_2 \gamma_1 \ln 2 - (\mu_{1r} + \mu_r)s_1, \qquad c_1 = -\mu_{1r}\hat{P}_2 s_1 s_2,$$

$$a_2 = \gamma_2 \ln 2, \quad b_2 = \hat{P}_1 s_1 \gamma_2 \ln 2 - (\mu_{2r} + \mu_r)s_2, \qquad c_2 = -\mu_{2r}\hat{P}_1 s_1 s_2, \quad (19)$$

and $\boldsymbol{\gamma}$ towards their optimal values. These step sizes should be chosen to be monotonically decreasing functions of $t$.

From Theorem 1, the optimal mode selection variables $q_i^*(t), \forall i \in M = \{1, 2, 3\}$ can only be obtained at the boundary of the region $[0, 1]$, which corresponds to the binary values before relaxation. Thus, the optimal values of $E_i^{M_i}(t)(i = 1, 2)$ and $E_r(t)$ equal to $P_i(t)(i = 1, 2)$ and $P_r(t)$ when the corresponding mode is selected. We have the following observations:

- In the EH mode $\mathcal{M}_1$, from (18), the optimal transmit power at $U_i$ can only be zero or $\hat{P}_i$. Moreover, in order to reduce energy consumption in the EH mode, two users of $U_1$ and $U_2$ will transmit at the maximum power only when the channel of the corresponding link is in a good condition.
- In the multiple access mode $\mathcal{M}_2$, the CSIs $(s_1(t), s_2(t))$ can be divided into several mutually exclusive regions, each of which corresponds to one optimal power allocation case, as illustrated in (19). From (17), one may notice that the optimal rate allocation $\mathbf{R}^*$ is dependent on the transmit power, the CSIs, $\alpha_1$, $\alpha_2$ and $\theta$. Moreover, it also discloses that, in the successive interference cancellation of the multiple-access mode, the decoding order depends only on the statistical CSIs and the weighting coefficient $\theta$, it is independent of the CSIs of the current time slot. In particular, when $\mu_{1r}(t) > 0$ and $\mu_{2r}(t) = 0$, the relay considers the signal received from $U_1$ as noise and decodes the signal received from $U_2$. Then the relay subtracts the signal from $U_2$ and decodes the signal from $U_1$, and vice versa for $u_{2r}(t) > 0$ and $\mu_{1r}(t) = 0$.
- In the broadcast mode $\mathcal{M}_3$, the optimal power allocation policy can be divided into three mutually independent regions based on the current CSIs $(s_1(t), s_2(t))$. Moreover, from (20), better channel conditions can result in a higher transmit power being allocated at relay.

According to Theorem 1, we can obtain the optimal adaptive relaying design for the buffer-aided two-way EH relay network, which can achieve the optimal long-term average achievable rate region. When data buffers are deployed at the relay, the relay can adaptively decide which one to transmit to achieve a higher throughput. When energy storage is deployed at the relay, in any given time slot, the relay can adaptively allocate power and decide whether to forward data or harvest energy based on the current CSIs, which can improve the energy efficiency. It is worthwhile to address that, Theorem 1 subsumes the following two special cases: 1) The one-way relaying either from $U_1$ to $U_2$ or from $U_2$ to $U_1$ corresponds to the case of $\theta = 1$ and $\theta = 0$, respectively. Since each user has its own power budget, in order to harvest more energy at the relay, the user without data to deliver still needs to transmit energy to the relay. 2) The maximum sum rate can be achieved when $\theta = \frac{1}{2}$ [17]. In the next section, in order to fulfill the delay sensitive traffic requirements, a delay-aware adaptive transmission (DAAT) scheme was proposed. The optimal adaptive relaying design can be considered as the outer bound of the DAAT scheme.

## V. Delay-Aware Adaptive Transmission Design

In this section, a delay-aware adaptive transmission scheme is proposed to allocate transmit powers adaptively and to determine the transmission mode based on the current BSIs, the ESI, and CSIs for the given delay transmission requirement. Moreover, it is shown that there exists an inherent tradeoff between the average transmission rate and the average delay by using Lyapunov optimization framework. In fact, Lyapunov optimization can be utilized to guarantee the queue stability and to approach the maximum achievable rate region simultaneously for the given average transmit power and the peak transmit power constraints.

### *A. Dynamic Queue Model*

The relay has an energy storage for storing the harvested energy and two data buffers for storing the received data from both users. When scheduled for energy harvesting at slot $t$, the relay stores the harvested RF energy into $B_e$. When the multiple-access mode is selected at slot $t$, the relay stores the received data into $B_1$ and $B_2$. When scheduled for data broadcast transmission at slot $t$, the relay uses the harvested energy to forward the data back to both users. Therefore, the dynamic update process of $B_e$, $B_1$, and $B_2$ can be described as follows

$$E(t+1) = \max\{E(t) + E_h(t) - q_3(t)P_r(t)T, 0\}, \quad (22)$$

$$Q_1(t+1) = \max\{Q_1(t) + R_{1r}(t) - R_{r2}(t), 0\}, \quad (23)$$

$$Q_2(t+1) = \max\{Q_2(t) + R_{2r}(t) - R_{r1}(t), 0\}. \quad (24)$$

Moreover, we also consider the average power consumption states for each user. Let $Z_1(t)$ and $Z_2(t)$ denote the average power consumption states of $U_1$ and $U_2$, respectively, and $Z_i(t), i = 1, 2$ can be given by

$$Z_i(t+1) = \max\{Z_i(t) + \big(q_1(t) + q_2(t)\big)P_i(t) - \bar{P}_i^{\max}, 0\}. \quad (25)$$

In fact, one may readily derive from (25) that

$$Z_i(t+1) \geq Z_i(t) + (q_1(t) + q_2(t))P_i(t) - \bar{P}_i^{\max}. \quad (26)$$

By summing (26) over $t$ time slots and dividing it by $t$, we can obtain

$$\frac{Z_i(t) - Z_i(0)}{t} \geq \frac{1}{t}\sum_{\tau=0}^{t-1}(q_1(\tau) + q_2(\tau))P_i(\tau) - \bar{P}_i^{\max}. \quad (27)$$

Without loss of generality, we assume that the initial state of $Z_i(t)$ is zero, i.e., $Z_i(0) = 0$. By taking expectations on both sides and letting $t \to \infty$, we obtain

$$\lim_{t\to\infty} \frac{\mathbb{E}\{Z_i(t)\}}{t} \geq \bar{P}_i - \bar{P}_i^{\max}. \quad (28)$$

Furthermore, if $Z_i(t)$ is mean rate stable, i.e., $\lim_{t\to\infty} \frac{\mathbb{E}\{|Z_i(t)|\}}{t} = 0$, we can derive $\bar{P}_i \leq \bar{P}_i^{\max}$. Hence the average power consumption constraint of $U_i$ can be satisfied. This implies that we can transform the time average power consumption constraints into a pure queue stability problem. The Lyapunov optimization framework can be utilized to solve problem $P1$.

### *B. Lyapunov Optimization Framework and Delay-Aware Adaptive transmission policy*

Based on the energy storage state, two data buffer states, as well as two virtual power consumption states, we can define the following quadratic Lyapunov function

$$L(\boldsymbol{\Theta}(\mathbf{t})) = \frac{\psi}{2}(\phi - E(t))^2 + \frac{1}{2}\sum_{i=1}^{2}(Q_i^2(t) + Z_i^2(t)), \quad (29)$$

where $\boldsymbol{\Theta}(\mathbf{t}) = [E(t), Q_1(t), Q_2(t), Z_1(t), Z_2(t)]$ denotes the concatenated vector of all queues, $\psi$ is a nonnegative constant and $\phi$ represents perturbation value of the energy storage at the relay. We assume that the energy storage size does not exceed $\phi$. The value $L(\boldsymbol{\Theta}(\mathbf{t}))$ is a scalar measure of the current queue length, which grows larger as the data buffer size $Q_i(t)$ and the virtual power consumption queue size $Z_i(t)$ increases, or the energy storage size $E(t)$ decreases. The Lyapunov drift stands for the expected change in the Lyapunov function between two consecutive time slots, which is given by

$$\Delta(\boldsymbol{\Theta}(\mathbf{t})) = \mathbb{E}\{L(\boldsymbol{\Theta}(\mathbf{t+1})) - L(\boldsymbol{\Theta}(\mathbf{t}))|\boldsymbol{\Theta}(\mathbf{t})\}, \quad (30)$$

where the expectation is taken over the CSIs randomness of the and transmission decisions during time slot $t$ for given the current queue vector $\boldsymbol{\Theta}(\mathbf{t})$. To ensure stability of all queues, our transmission decisions should try to minimize the Lyapunov drift $\Delta(\boldsymbol{\Theta}(\mathbf{t}))$.

The objective of the proposed scheme is to maximize the achievable rate region while stabilizing the queues. To this end, we can follow the Lyapunov optimization framework to minimize the following Lyapunov drift-plus-penalty

$$\Delta(\boldsymbol{\Theta}(\mathbf{t})) - V\mathbb{E}\{R_{sum}(t)|\boldsymbol{\Theta}(\mathbf{t})\}, \quad (31)$$

where $R_{sum}(t) = \theta R_{1r}(t) + (1-\theta)R_{2r}(t)$ and $V$ is a nonnegative control parameter, which can effectively affect the tradeoff between the expected queue size and the achievable rate region.

*Lemma 1:* The Lyapunov drift-plus-penalty function can be upper bounded by

$$\begin{aligned}
\Delta(\boldsymbol{\Theta}(\mathbf{t})) - V\mathbb{E}\big\{R_{sum}(t)|\boldsymbol{\Theta}(\mathbf{t})\big\} &\leq B - V\mathbb{E}\big\{R_{sum}(t)|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
&+ \psi(\phi - E(t))\mathbb{E}\big\{q_3(t)P_r(t)T - E_h(t)|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
&+ \sum_{i=1}^{2} Z_i(t)\mathbb{E}\big\{(q_1(t)+q_2(t))P_i(t) - \bar{P}_i^{\max}|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
&+ Q_1(t)\mathbb{E}\big\{R_{1r}(t) - R_{r2}(t)|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
&+ Q_2(t)\mathbb{E}\big\{R_{2r}(t) - R_{r1}(t)|\boldsymbol{\Theta}(\mathbf{t})\big\}, \forall t, \quad (32)
\end{aligned}$$

where $B$ is a positive constant that satisfies the following constraint

$$\begin{aligned}
B \geq \frac{\psi}{2}\mathbb{E}\big\{\big(q_3(t)P_r(t)T\big)^2 + E_h(t)^2|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
+ \frac{1}{2}\sum_{i=1}^{2}\mathbb{E}\big\{\big((q_1(t)+q_2(t))P_i(t)\big)^2 + (\bar{P}_i^{\max})^2|\boldsymbol{\Theta}(\mathbf{t})\big\} \\
+ \frac{1}{2}\sum_{i=1}^{2}\mathbb{E}\big\{R_{ir}(t)^2 + R_{ri}(t)^2|\boldsymbol{\Theta}(\mathbf{t})\big\}, \ \forall t. \quad (33)
\end{aligned}$$

*Proof:* See Appendix B. ■

In order to ensure that both the queue stability and the maximum achievable rate region can be achieved simultaneously, we have to minimize the drift-plus-penalty in (32). Since Lemma 1 provides us an upper bound on the drift-plus-penalty, we may resort to minimizing the upper bound in (32) instead of directly minimizing the drift-plus-penalty item. At each time slot $t$, given the current queue state $\boldsymbol{\Theta}(\mathbf{t})$ and the current CSIs, decisions on power allocations, rate allocations, and mode selection are made by solving the following optimization problem:

$$\begin{aligned}
\min_{\mathbf{q(t)},\mathbf{P(t)},\mathbf{R(t)}} &: -\Delta_1(t)R_{1r}(t) - \Delta_2(t)R_{2r}(t) - Q_1(t)R_{r2}(t) \\
&- Q_2(t)R_{r1}(t) + \sum_{i=1}^{2} Z_i(t)(q_1(t)+q_2(t))P_i(t) \\
&+ \psi(\phi - E(t))(q_3(t)P_r(t)T - E_h(t)) \\
\text{s.t.} &\quad (4), (7), (9a). \quad (34)
\end{aligned}$$

where $\Delta_1(t) = \theta V - Q_1(t)$ and $\Delta_2(t) = (1-\theta)V - Q_2(t)$. Since the optimization variables $\mathbf{q(t)}$ are binary variables, and only three transmission modes can be selected. We can enumerate the above optimization problem in different cases.

*1) EH Mode $\mathcal{M}_1$:* At time slot $t$, if the EH mode is selected, the relay harvests energy transmitted from both users, we have $q_1(t) = 1$ and $q_2(t) = q_3(t) = 0$. Therefore, the optimization problem of (34) can be equivalently rewritten as below

$$\begin{aligned}
\min_{P_1(t),P_2(t)} &: -\psi(\phi - E(t))E_h(t) + Z_1(t)P_1(t) + Z_2(t)P_2(t) \\
\text{s.t.} &\quad 0 \leq P_i(t) \leq \hat{P}_i, \ i = 1, 2. \quad (35)
\end{aligned}$$

The objective function and the constraint function in (35) are both linear, thus the optimal solution can only be obtained at the boundary. We can obtain the following Lemma 2.

*Lemma 2:* In the energy harvesting mode, the optimal power allocation of $U_i (i = 1, 2.)$ is given by

$$P_i(t) = \begin{cases} \hat{P}_i, & Z_i(t) \leq \psi(\phi - E(t))|h_i(t)|^2 T\eta, \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

*2) Multiple-Access Mode $\mathcal{M}_2$:* In the multiple-access mode, the relay is selected for reception, we have $q_2(t) = 1$ and $q_1(t) = q_3(t) = 0$. Thus, the optimization problem in (34) can be rewritten as below

$$\begin{aligned}
\min_{R_{1r}(t),R_{2r}(t),P_1(t),P_2(t)} &\quad \sum_{i=1}^{2}(-\Delta_i(t)R_{ir}(t) + Z_i(t)P_i(t)) \\
\text{s.t.,} &\quad (4), (9a). \quad (37)
\end{aligned}$$

Obviously, the objective function is linear and the constraint functions in (4) and (9a) are convex, thus the above optimization problem is standard convex. we can derive the optimal solutions by exploiting KKT conditions. Without loss of generality, we assume that $\Delta_1(t) \geq \Delta_2(t)$, however, the following analysis can be extended the case $\Delta_1(t) < \Delta_2(t)$. The optimization problem (37) can be sub-divided into the following three cases based on $(\Delta_1(t), \Delta_2(t))$ to find out the corresponding optimal solution.

*Case 1:* $\Delta_2(t) \leq \Delta_1(t) \leq 0$. In this case, the objective function monotonically increases with respect to $R_{1r}(t)$, $R_{2r}(t)$, $P_1(t)$ and $P_2(t)$, thus the optimal rate and power allocation scheme are satisfied with $R_{ir} = P_i(t) = 0$, $i = 1, 2$.

*Case 2:* $\Delta_2(t) \leq 0$ and $\Delta_1(t) > 0$. Firstly, we may derive the optimal $R_{1r}(t)$ and $R_{2r}(t)$ for the given $P_1(t)$ and $P_2(t)$. The objective function monotonically decreases with respect to $R_{1r}(t)$ and monotonically increases with respect to $R_{2r}(t)$, thus the optimal rate allocation policy is achieved when $R_{1r}(t) = \log_2(1 + P_1(t)s_1(t))$ and $R_{2r}(t) = 0$. Then, we can derive the optimal power allocation as follows

$$P_1(t) = \min\left\{\left(\frac{\Delta_1(t)}{Z_1(t)\ln 2} - \frac{1}{s_1(t)}\right)^+, \hat{P}_1\right\}, \quad P_2(t) = 0.$$

*Case 3:* $\Delta_1(t) \geq \Delta_2(t) > 0$. The optimization problem in (37) can be reformulated as

$$\begin{aligned} \min \quad & -\Delta_2(t)(R_{1r}(t) + R_{2r}(t)) + (\Delta_2(t) - \Delta_1(t))R_{1r}(t) \\ & + Z_1(t)P_1(t) + Z_2(t)P_2(t) \\ \text{s.t.,} \quad & (4), (9a). \end{aligned} \tag{38}$$

It is worth noting that the optimal rate allocation scheme can be achieved when $R_{1r}(t) + R_{2r}(t) = \log_2(1 + P_1(t)s_1(t) + P_2(t)s_2(t))$ and $R_{1r}(t) = \log_2(1 + P_1(t)s_1(t))$, thus the optimal rate allocation policy is given by

$$R_{1r}(t) = \log_2(1 + P_1(t)s_1(t)), \tag{39a}$$

$$R_{2r}(t) = \log_2\left(1 + \frac{P_2(t)s_2(t)}{1 + P_1(t)s_1(t)}\right). \tag{39b}$$

By substituting the optimal $R_{1r}(t)$ and $R_{2r}(t)$ into (38), and using the KKT conditions, we can obtain the optimal power allocation policy in (40) shown at the bottom of the page. Different from the traditional water-filling algorithms, the power allocation policy is not only related to the CSIs, but also depen-

$$(P_1(t), P_2(t)) = \begin{cases} (0,0), & \text{if } g_1(t) \leq 0 \wedge g_2(t) \leq 0 \\ \left(0, \frac{\Delta_2(t)}{Z_2(t)\ln 2} - \frac{1}{s_2(t)}\right), & \text{if } g_2(t) > 0 \wedge g_4(t) < 0 \wedge g_8 \geq 0 \\ \left(\frac{\Delta_1(t)}{Z_1(t)\ln 2} - \frac{1}{s_1(t)}, 0\right), & \text{if } g_1(t) > 0 \wedge g_3(t) < 0 \wedge g_7(t) \leq 0 \\ \Big(\frac{(\Delta_1(t) - \Delta_2(t))s_2(t)}{(Z_1(t)s_2(t) - Z_2(t)s_1(t)\ln 2)} - \frac{1}{s_1(t)}, & \\ \quad \frac{\Delta_2(t)}{Z_2(t)\ln 2} - \frac{(\Delta_1(t) - \Delta_2(t))s_1(t)}{(Z_1(t)s_2(t) - Z_2(t)s_1(t))\ln 2}\Big), & \text{if } g_7(t) > 0 \wedge g_8(t) < 0 \wedge g_9(t) > 0 \wedge g_{11}(t) > 0 \\ (0, \hat{P}_2), & \text{if } g_4(t) \geq 0 \wedge g_{10} \leq 0 \\ (\hat{P}_1, 0), & \text{if } g_3(t) \geq 0 \wedge g_5(t) \leq 0 \\ \left(\hat{P}_1, \frac{\Delta_2(t)}{Z_2(t)\ln 2} - \frac{1 + \hat{P}_1 s_1(t)}{s_2(t)}\right), & \text{if } g_5(t) > 0 \wedge g_6(t) < 0 \wedge g_9(t) \leq 0 \\ \left(\frac{-b_3 + \sqrt{b_3^2 - 4a_3c_3}}{2a_3} - \frac{1}{s_1(t)}, \hat{P}_2\right), & \text{if } g_{11}(t) \leq 0 \wedge g_{10}(t) > 0 \wedge g_{12}(t) < 0 \\ (\hat{P}_1, \hat{P}_2), & \text{if } g_6(t) \geq 0 \wedge g_{12}(t) \geq 0 \end{cases}$$

$$\omega_1(t) = \frac{Z_1(t)\ln 2}{s_1(t)}, \qquad \omega_2(t) = \frac{Z_2(t)\ln 2}{s_2(t)},$$

$$g_1(t) = \Delta_1(t) - \omega_1(t), \qquad g_2(t) = \Delta_2(t) - \omega_2(t),$$

$$g_3(t) = \Delta_1(t) - (1 + \hat{P}_1 s_1(t))\omega_1(t), \qquad g_4(t) = \Delta_2(t) - (1 + \hat{P}_2 s_2(t))\omega_2(t),$$

$$g_5(t) = \Delta_2(t) - (1 + \hat{P}_1 s_1(t))\omega_2(t), \qquad g_6(t) = \Delta_2(t) - (1 + \hat{P}_1 s_1(t) + \hat{P}_2 s_2(t))\omega_2(t),$$

$$g_7(t) = D_2(t) - D_1(t)\frac{\omega_2(t)}{\omega_1(t)}, \qquad g_8(t) = \Delta_2(t) - \Delta_1(t) - (\omega_2(t) - \omega_1(t)),$$

$$g_9(t) = \Delta_2(t) - \Delta_1(t) - (1 + \hat{P}_1 s_1(t))(\omega_2(t) - \omega_1(t)), \qquad g_{10}(t) = \Delta_1(t) - \frac{\hat{P}_2 s_2(t)}{1 + \hat{P}_2 s_2(t)}\Delta_2(t) - \omega_1(t),$$

$$g_{11}(t) = \Delta_1(t) - \frac{\omega_1(t)}{\omega_2(t)}\Delta_2(t) - \hat{P}_2 s_2(t)(\omega_2(t) - \omega_1(t)), \qquad a_3 = Z_1(t)\ln 2, b_3 = \hat{P}_2 s_2(t) Z_1(t)\ln 2 - \Delta_1(t)s_1(t)$$

$$g_{12}(t) = \Delta_1(t) - \frac{\hat{P}_2 s_2(t)}{1 + \hat{P}_1 s_1(t) + \hat{P}_2 s_2(t)}\Delta_2(t) - (1 + \hat{P}_1 s_1(t))\omega_1(t), \quad c_3 = (\Delta_2(t) - \Delta_1(t))\hat{P}_2 s_1(t) s_2(t) \tag{40}$$

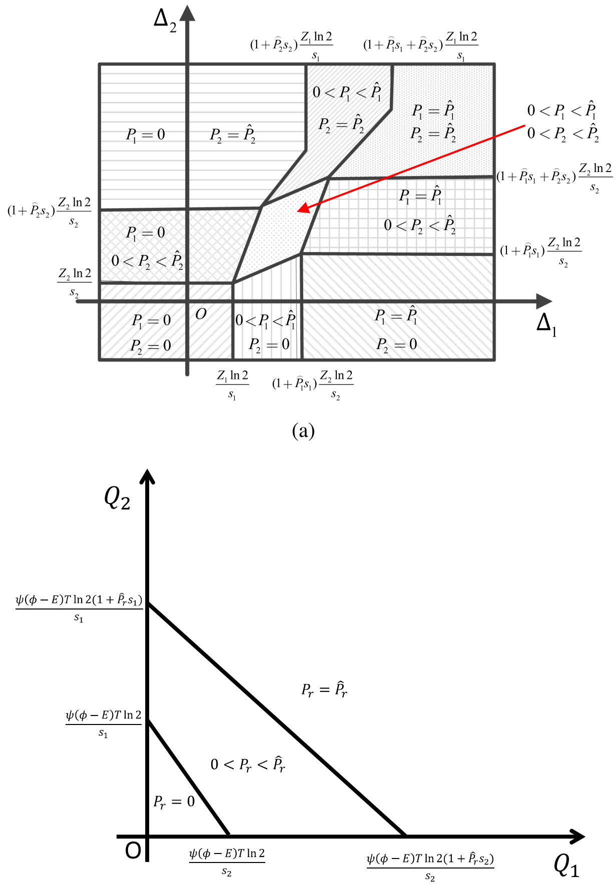


Fig. 2. (a). The optimal transmit power allocation illustration of $U_1$ and $U_2$ in the multiple-access mode. (b) The optimal transmit power allocation illustration of relay in broadcast mode.

dent on the underlying queue states. To gain more insight, the optimal power allocation scheme in the multiple-access mode are demonstrated in Fig. 2(a). We can find that, larger $\Delta_1(t)$ and $\Delta_2(t)$ will result in larger allocated power at $U_1$ and $U_2$, respectively. That is, with the reduction in $Q_1(t)(Q_2(t))$, the transmit power of the corresponding user will be larger. In particular, one may readily observe that $U_i$ will never transmit information to relay when $\Delta_i(t) \leq 0$. In other words, $\theta V$ and $(1-\theta)V$ can be regards as the threshold of the data queue size $Q_1(t)$ and $Q_2(t)$, respectively. Once $Q_1(t)(Q_2(t))$ exceeds $\theta V((1-\theta)V)$, $U_1(U_2)$ will never transmit data to the relay. Thus, we can effectively control the data queue sizes by strictly controlling the value of V. Meanwhile, one may readily observe that the power allocation policy in the multiple-access mode complies with the rate adjusting requirement. If both data queue sizes are larger, two users should remain silent since now there are already enough data at the relay. If one of data queue size is larger, the corresponding user should keep silent for a while to guarantee queue stability. If both data queue sizes are small, both users should transmit data simultaneously to the relay. Moreover, in order to obtain the maximum amount of transmission data, both users should transmit data at maximum power when both data queue sizes are small. On the other hand, one may readily observe that, if the link quality $U_i \leftrightarrow R\,(i=1,2)$ is very poor, i.e., $\frac{Z_i(t)\ln 2}{s_i(t)}\,(i=1,2)$ is very large, which leads to $0 \leq P_r(t) \leq \hat{P}_r$ holds with high probability, $U_i$ tends to remain idle to save transmit power when the corresponding link quality is weak.

*3) Broadcast Mode:* In the broadcast mode, the relay is selected for transmission, we have $q_1(t) = q_2(t) = 0$ and $q_3(t) = 1$. The optimization problem of (34) can be reformulated as

$$\begin{aligned} \min_{R_{r1}(t), R_{r2}(t), P_r(t)} \quad : \quad & \psi(\phi - E(t))P_r(t)T - Q_1(t)R_{r2}(t) \\ & - Q_2(t)R_{r1}(t) \\ \text{s.t.} \quad & (7), (9a). \end{aligned} \tag{41}$$

Similarly, this is a typical optimization problem with linear objective function and convex constraints. By exploiting KKT conditions, we can derive the optimal power allocations and rate allocations in the following Lemma 3.

*Lemma 3:* In the broadcast mode, the optimal power allocation of relay is given by

$$P_r(t) = \begin{cases} 0, \text{ if } \psi(\phi - E(t))T\ln 2 \geq Q_2(t)s_1(t) + Q_1(t)s_2(t), \\ \hat{P}_r, \text{if } \psi(\phi - E(t))T\ln 2 \leq \dfrac{Q_2(t)s_1(t)}{1+\hat{P}_r s_1(t)} \\ \qquad + \dfrac{Q_1(t)s_2(t)}{1+\hat{P}_r s_2(t)}, \\ \dfrac{-b_r + \sqrt{b_r^2 - 4a_r c_r}}{2a_r}, \text{otherwise}, \end{cases}$$

where $a_r = \psi(\phi - E(t))Ts_1(t)s_2(t)\ln 2$, $b_r = \psi(\phi - E(t))T(s_1(t) + s_2(t))\ln 2 - (Q_1(t) + Q_2(t))s_1(t)s_2(t)$, and $c_r = \psi(\phi - E(t))T\ln 2 - Q_1(t)s_2(t) - Q_2(t)s_1(t)$. Moreover, the optimal rate from relay to users are given by

$$R_{r1}(t) = \min\{\log_2(1 + P_r(t)s_1(t), Q_2(t)\}, \tag{42a}$$

$$R_{r2}(t) = \min\{\log_2(1 + P_r(t)s_2(t), Q_1(t)\}. \tag{42b}$$

The optimal transmit power allocation scheme in the broadcast mode is not only related to the CSIs, but also dependent on the BSIs and ESI. As illustrated in Fig. 2(b), we can find that, the increase in both data queue sizes will lead to a larger transmit power at relay. Moreover, with the increase in the collected energy at relay, the value of $\frac{\psi(\phi - E(t))T\ln 2}{s_i}\,(i=1,2)$ will gradually decrease, which leads to $0 \leq P_r(t) < \hat{P}_r$ with smaller probability. In other words, the larger $E(t)$ will result in larger allocated power at relay. Furthermore, when channel quality is very poor, i.e., $\frac{\psi(\phi - E(t))T\ln 2}{s_i}\,(i=1,2)$ is very large, the relay remains silent with much higher probability. In this case, the relay tends to be idle to save power consumption. When both the optimal power allocation and rate allocation scheme in the corresponding mode have been obtained, by substituting them into the objective function of the corresponding optimization problem, we can obtain the optimal mode selection scheme in the following Lemma 4.

*Lemma 4:* The optimal mode selection scheme is given by

$$q_i^*(t) = \begin{cases} 1, \text{if } \tilde{\Lambda}_m(t) = \arg\min_{i=1,2,3} \tilde{\Lambda}_i(t), \\ 0, \text{otherwise}. \end{cases} \tag{43}$$

where $\widetilde{\Lambda}_1(t)$, $\widetilde{\Lambda}_2(t)$ and $\widetilde{\Lambda}_3(t)$ are referred to as the mode selection metrics, which are given by

$$\widetilde{\Lambda}_1(t) = -\psi(\phi - E(t))E_h(t) + Z_1(t)P_1(t) + Z_2(t)P_2(t), \tag{44a}$$

$$\widetilde{\Lambda}_2(t) = \sum_{i=1}^{2}\left(-\Delta_i(t)R_{ir}(t) + Z_i(t)P_i(t)\right), \tag{44b}$$

$$\widetilde{\Lambda}_3(t) = \psi(\phi - E(t))P_r(t)T - Q_2(t)R_{r1}(t) - Q_1(t)R_{r2}(t). \tag{44c}$$

From the mode selection scheme, we may see that, larger data queue sizes are expected when the multiple-access mode is activated, which leads to the increase in $\widetilde{\Lambda}_2(t)$ and the decrease in $\widetilde{\Lambda}_3(t)$, such that the proposed scheme will prefer to selecting the broadcast mode. The data queue size and energy queue size will be decreased when the broadcast mode is selected, such that the energy harvesting mode or multiple-access mode is selected with much higher probability. Therefore, the proposed DAAT scheme can ensure the stability of the data buffer and energy storage.

### C. Performance Analysis

When the DAAT scheme is utilized, we have the analytical bounds on the weighted long-term sum rates and time average queue size in Theorem 2.

*Theorem 2:* When all queues are initially empty and the transmit rates are strictly within the capacity region, for any $V > 0$ and $\epsilon > 0$, the weighted long-term sum rates and the time average queue length of the DAAT scheme fulfill the following constraints

$$R_{sum}^* - \frac{B}{V} \le \lim_{N\to\infty}\frac{1}{N}\sum_{t=0}^{N-1}\mathbb{E}\{R_{sum}(t)\} \le R_{sum}^*,$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N-1}\mathbb{E}\{Q_1(t)+Q_2(t)\} \le \frac{B + V[R_{sum}^* - \Psi(\epsilon)]}{\epsilon},$$

where $R_{sum}(t) = \theta R_{1r}(t) + (1-\theta)R_{2r}(t)$ stands for the weighted sum rate in the $t$-th time slot. $R_{sum}^* = \theta R_{1r}^* + (1-\theta)R_{2r}^*$ is the theoretically optimal weighted long-term sum rate, $R_{1r}^*$ and $R_{2r}^*$ denotes the theoretical optimal time average data rate from $U_1$ to relay and that from $U_2$ to relay, respectively. $\Psi(\epsilon)$ is less than $R_{sum}^*$. $\epsilon > 0$ indicates the gap between the weighted long-term sum rates and the theoretical sum rate region.

*Proof:* See Appendix C. ■

From Theorem 2, we may notice that the average queue length grows linearly with $V$, and the gap between the weighted long-term sum rates achieved by DAAT scheme and $R_{sum}^*$ is inversely proportional to $V$. Namely, there exists $[O(1/V), O(V)]$ tradeoff between rates and average delay.

## VI. Numerical Analysis

In this section, we evaluate the performance of the proposed buffer-aided adaptive transmission policy and the delay-aware adaptive transmission (DAAT) for two-way relay networks in Rayleigh fading through Monte-Carlo simulations. The channel gains ($|h_1(t)|^2$ and $|h_2(t)|^2$) follow exponential distributions with means $\frac{1}{\Omega_1}$ and $\frac{1}{\Omega_2}$, respectively. In all simulations, unless otherwise stated, we set $\Omega_1 = \Omega_2 = 0$ dB, the path loss exponent $m = 2.7$, the energy conversion efficiency $\eta = 0.5$, the noise variances at all nodes are $\sigma^2 = -100$ dBm, the average transmission power at $U_1$ and $U_2$ are $\bar{P}_1^{\max} = \bar{P}_2^{\max} = 30$ dBm, and the peak transmission power at $U_i$ $(i = 1, 2)$ is $\hat{P}_i = 3\bar{P}_i^{\max}$. In particular, we set the peak transmit power at relay $\hat{P}_r = \frac{E(t)}{T}$ such that does not exceed the harvested energy in energy storage. All the presented simulation results are obtained for $N = 10^6$ time slots. For fair comparison purpose, we introduce the benchmark scheme at first.

### A. Benchmark Scheme

We consider a two-way EH wireless relay benchmark network, wherein the relay is neither provisioned with data buffer nor with the energy storage. Thus all the harvested energy at relay will be used for forwarding data in the broadcast phase. We also refer to the benchmark scheme as the conventional scheme, wherein the duration $T$ of each time slot will also be subdivided into three slots of $\alpha T$, $\frac{(1-\alpha)T}{2}$ and $\frac{(1-\alpha)T}{2}$. $\alpha T$ is used for the relay to harvest energy from both users, the first $\frac{(1-\alpha)T}{2}$ is used for both users to transmit data to the relay simultaneously, and the second $\frac{(1-\alpha)T}{2}$ is used for broadcasting information from the relay to both users. Let $E_1$ and $E_2$ denote the total energy consumption at $U_1$ and $U_2$ in each time slot, respectively. Thus, the transmit power at $U_i$ $(i = 1, 2)$ is $P_i = \frac{2E_i}{(1+\alpha)T}$. In the EH phase, the harvested energy at relay is $E_h = \frac{2\alpha\eta(E_1|h_1|^2 d_2^m + E_2|h_2|^2 d_1^m)}{(1+\alpha)d_1^m d_2^m}$. In the multiple-access phase, the achievable rate region must satisfy the multiple-access capacity region. In the broadcast phase, all the harvested energy is used to forward information, thus the average transmit power of relay is $P_r = \frac{4\alpha\eta(E_1|h_1|^2 d_2^m + E_2|h_2|^2 d_1^m)}{(1-\alpha^2)d_1^m d_2^m T}$. Let $R_{12}$ and $R_{21}$ denote the transmission rate from $U_1$ to $U_2$ and from $U_2$ to $U_1$, respectively. And the achievable rate of region for the conventional benchmark system can be derived by solving the following optimization problem

$$\min_{R_{12}, R_{21}, \alpha} : -\theta R_{12} - (1-\theta)R_{21}$$

$$\text{s.t.}, C1 : R_{12} \le \frac{1-\alpha}{2}\log_2\left(1 + \frac{P_1|h_1|^2}{d_1^m\sigma^2}\right),$$

$$C2 : R_{21} \le \frac{1-\alpha}{2}\log_2\left(1 + \frac{P_2|h_2|^2}{d_2^m\sigma^2}\right),$$

$$C3 : R_{12} + R_{21} \le \frac{1-\alpha}{2}\log_2\left(1 + \frac{P_1|h_1|^2}{d_1^m\sigma^2} + \frac{P_2|h_2|^2}{d_2^m\sigma^2}\right),$$

$$C4 : R_{12} \le \frac{1-\alpha}{2}\log_2\left(1 + \frac{P_r|h_2|^2}{d_2^m\sigma^2}\right),$$

$$C5 : R_{21} \le \frac{1-\alpha}{2}\log_2\left(1 + \frac{P_r|h_1|^2}{d_1^m\sigma^2}\right),$$

$$C6 : 0 < \alpha < 1.$$

where $C1$, $C2$ and $C3$ stand for the multiple-access capacity region constraints, $C4$ and $C5$ correspond to the broadcast capacity region constraints, $C6$ is the time assignment ratio constraint. Moreover, the right hand side of constraints $C1$, $C2$ and $C3$ are convex functions with respect to $\alpha$, the right hand side of constraints $C4$ and $C5$ are concave functions with respect to $\alpha$. The objective function and constraint $C6$ are linear function. Thus, the above optimization problem can be solved effectively

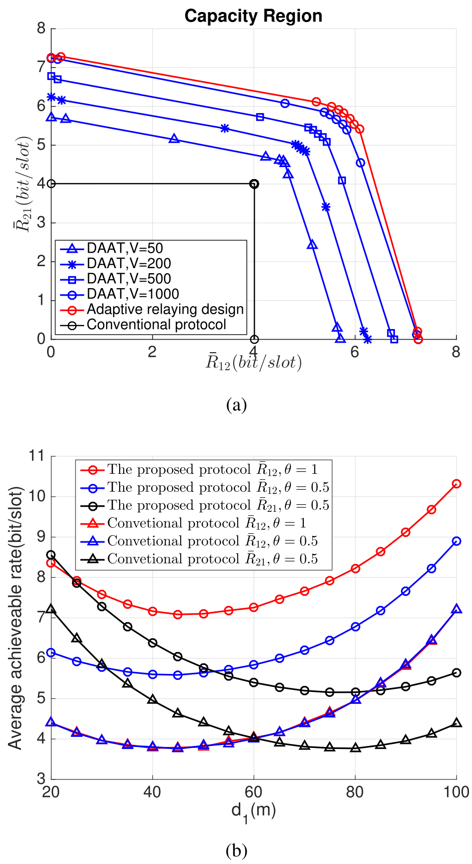


Fig. 3. (a). The achievable average rate region with $\bar{P}_1^{\max} = \bar{P}_2^{\max} = 30$ dBm and $d_1 = d_2 = 60$ m. (b) The achievable average rate in different relay placement scenarios.

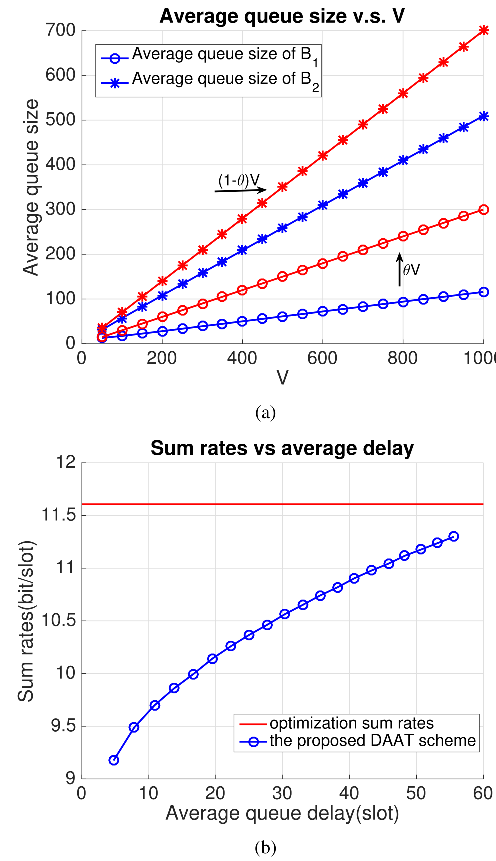


Fig. 4. (a). The average delay with different choice of $V$. (b). The average sum rates versus the average delay.

by using the DC (difference of two convex functions) programming algorithm [29].

## B. Performance Assessment

In Fig. 3(a), we present the achievable long-term rate regions of the proposed DAAT scheme and the benchmark scheme, wherein the distances from both users to the relay are assumed to be identical, i.e., $d_1 = d_2 = 60$ m. We can find that the rate region of the conventional scheme is far inferior to that of the proposed DAAT scheme. In addition, the rate regions of the DAAT scheme with different $V$ are in the interior of the achievable rate region of the adaptive relaying design for the buffer-aided two-way wireless EH relay network. When $V$ is large enough, the achievable rate region of the DAAT scheme will be very close to that theoretical rate region.

The average achievable rates as a function of the distance between $U_1$ and the relay with different $\theta$ are presented in Fig. 3(b). Here, the distance between $U_1$ and $U_2$ is $d_1 + d_2 = 120$ m. Note that the two-way relay network will reduce to the one-way relaying when $\theta = 0$ or $\theta = 1$. As shown in Fig. 3(b), when $d_1$ increases from 20 m to 100 m, the achievable rate $\bar{R}_{12}$ and $\bar{R}_{21}$ will first decrease till the minimum value, and then increases. More interestingly, a larger $\bar{R}_{12}$ can be realized when the relay is placed farther away from $U_1$, while a larger $\bar{R}_{21}$ can be achieved when the relay is closer to the $U_1$, which suggests that the relay should be placed closer to $U_j$ when a larger $\bar{R}_{ij}$ is desired. This can be interpreted as follows: i) The large-scale fading of RF signals decays exponentially with distance, which implies

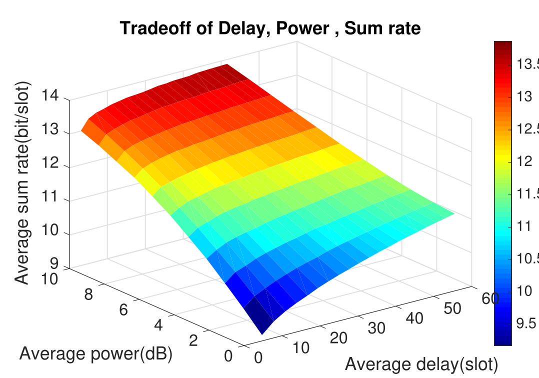


Fig. 5. The tradeoff of transmission delay, power consumption, and sum rate.

that more energy can be harvested by the relay if the relay is placed closer to users; ii) The relay is an energy-constrained node, which limits the system performance by the second hop, so a larger $\bar{R}_{ij}$ can be achieved when the relay is closer to $U_j$.

The impacts of different choice $V$ on the achieved performance in the proposed DAAT scheme are illustrated in Fig. 4. The relationship between average queue size and $V$ is depicted in Fig. 4(a). We may see that, the average queue sizes at both $U_1$ and $U_2$ are proportional to $V$, which are less than $\theta V$ and $(1-\theta)V$, respectively. The relationship between average achievable rate and average queue delay is presented in Fig. 4(b). One may notice that, we may realize higher achievable rate when lager average queue delay is tolerable. Moreover, the achievable rate will gradually approach the optimal rate with the increase in $V$. It suggests us that the increase in average achievable rate can be achieved at the cost of some increase in the average queue delay. All of these simulation results comply with Theorem 2. The inherent tradeoff between the average delay, the average power consumption, and the average achievable rate in the DAAT scheme is shown in Fig. 5. As is shown that, a larger tolerable delay and larger average power consumption will lead to a larger achievable rate, as expected. It also suggests us that the requested delivery rate can be realized with less power consumption if a larger delay is acceptable for the specific traffic requirement.

## VII. Conclusions

In this paper, we have investigated a buffer-aided relaying protocol for the two-way EH relay networks. To reveal the achievable rate region for the specific relay network, the opportunistic scheduling design problem is formulated to take into account of three transmission mode selection, the power allocations and the rate allocations at multiple access mode. Meanwhile, we consider the peak transmit power and the average transmit power constraint at two users and the energy harvesting constraint at relay. The proposed protocols can realize better temporal diversity gains due to the adaptive selection of the transmission mode based on the CSIs and the queue state information. Moreover, the inherent tradeoff between the achievable rate, delivery delay and power consumption indicates the great potential of the buffer-aided relaying design. Finally, it is shown that, when a certain time delay is tolerable, the DAAT scheduling scheme is able to realize a rate region arbitrarily close to the achievable rate region of the buffer-aided two-way wireless EH relay network, which suggests the great potential of the buffer-aided relaying techniques in the wireless relaying system. Since our focus in this paper is to show the achievable rate region and the feasibility to approach it for the buffer-aided two-way EH relay networks, we assume that the energy storage size and both data buffer size are large enough such that no overflow happens. In order to fulfill the delay sensitive traffic requirements, the DAAT scheduling design is addressed. We assume a centralized scheduling scheme, wherein a central control node is able to acquire the underlying CSI, BSI, and ESI. The central node is thus able to perform the scheduling operations and inform all the nodes its decisions. In practical applications, the energy storage and the data buffer size may be limited, and there always exists some imperfections in the acquisition of the underlying state information, we leave these issues for the future analysis. Finally, the secrecy issue in cooperative relaying network has attracted many research attentions [32], how to fulfill the secrecy requirement in the two-way EH relay network would be another interesting research subjects in our future work.

## Appendix A

In this appendix, our goal is to derive the optimal mode selections, power and rate allocations by solving $\mathbf{P2}$. Since the problem $\mathbf{P2}$ is convex for the optimization variables of $\mathbf{E(t)}, \mathbf{q(t)}, \mathbf{R(t)}$ and the feasible set is nonempty, we can obtain the optimal solutions to $\mathbf{P2}$. Let $l_j(t), j \in J = \{1r, r2, 2r, r1\}$ denote the Lagrange multiplier associated with the constraint $R_j(t) \geq 0$, $l_i^{M_1}(t)$, $l_i^{M_2}(t) (i = 1, 2)$, and $l_r^{M_3}(t)$ denote the Lagrange multipliers associated with the constraint $E_i^{M_1}(t) \geq 0$, $E_i^{M2}(t) \geq 0$, and $E_r^{M_3}(t) \geq 0$, respectively. $\upsilon_m(t), m = 1, 2, 3$ and $\upsilon(t)$ stands for the Lagrange multipliers associated with mode selection constraints and the constraint $\sum_{i=1}^{3} q_i(t) = 1$, respectively. Other Lagrange multipliers are defined in Section IV. Now the Lagrangian function for $\mathbf{P2}$ can be rewritten in (45), shown at the bottom of the page.

*1) The Optimal Mode Selection Policy Derivation:* By setting the derivative of the Lagrangian function with respect to $q_i(t)$ equal to zero we have

$$\frac{\partial \mathcal{L}}{q_1(t)} = -\lambda_1^{M_1} \hat{P}_1 - \lambda_2^{M_1} \hat{P}_2 + \upsilon_1(t)(2q_1(t) - 1) + \upsilon(t) = 0, \tag{46a}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{q_2(t)} = &\sum_{i=1}^{2} \left( \frac{\mu_{ir}(t) E_i^{M_2}(t) s_i(t)}{(E_i^{M_2}(t) s_i(t) + q_2(t)) \ln 2} - \frac{\mu_{ir}(t) C_{ir}(t)}{q_2(t)} \right) \\ &+ \frac{\mu_r(t)(E_1^{M_2}(t) s_1(t) + E_2^{M_2}(t) s_2(t))}{(E_1^{M_2}(t) s_1(t) + E_2^{M_2}(t) s_2(t) + q_2(t)) \ln 2} \\ &- \frac{\mu_r(t) C_r(t)}{q_2(t)} - \lambda_1^{M_2}(t) \hat{P}_1 - \lambda_2^{M_2}(t) \hat{P}_2 + \upsilon(t) \\ &+ \upsilon_2(t)(2q_2(t) - 1) = 0,\end{aligned} \tag{46b}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{q_3(t)} = &\sum_{i=1}^{2} \left( -\frac{\mu_{ri}(t) C_{ri}(t)}{q_3(t)} + \frac{\mu_{ri}(t) E_r(t) s_i(t)}{(E_r(t) s_i(t) + q_3(t)) \ln 2} \right) \\ &- \lambda_r^{M_3}(t) \hat{P}_r + \upsilon_3(t)(2q_3(t) - 1) + \upsilon(t) = 0.\end{aligned} \tag{46c}$$

$$\begin{aligned}\mathcal{L}(\boldsymbol{q}, \boldsymbol{E}, \boldsymbol{R}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \boldsymbol{l}, \boldsymbol{\lambda}, \boldsymbol{\upsilon}) = &-\theta \bar{R}_{1r} - (1-\theta)\bar{R}_{2r} + \alpha_1(\bar{R}_{1r} - \bar{R}_{r2}) + \alpha_2(\bar{R}_{2r} - \bar{R}_{r1}) - \sum_{t=1}^{N} \sum_{j \in J} l_j(t) R_j(t) \\ &+ \sum_{t=1}^{N} \sum_{j \in J} \mu_j(t)(R_j(t) - C_j(t)) + \sum_{t=1}^{N} \mu_r(t)(R_{1r}(t) + R_{2r}(t) - C_r(t)) + \sum_{k \in \{1,2,r\}} r_k(\bar{E}_k - \bar{P}_k^{\max}) \\ &- \sum_{t=1}^{N} \sum_{i=1}^{2} \sum_{j=1}^{2} \left( l_i^{M_j}(t) E_i^{M_j}(t) - \lambda_i^{M_j}(t)(E_i^{M_j}(t) - q_j(t)\hat{P}_i) \right) - \sum_{t=1}^{N} \left( l_r^{M_3}(t) E_r(t) - \lambda_r^{M_3}(t)(E_r(t) - q_3(t)\hat{P}_r) \right) \\ &+ \sum_{t=1}^{N} \sum_{i=1}^{3} \upsilon_i(t) q_i(t)(q_i(t) - 1) + \sum_{t=1}^{N} \upsilon(t) \left( \sum_{i=1}^{3} q_i(t) - 1 \right).\end{aligned} \tag{45}$$

Now we will prove that the optimal $q_i^*(t), \forall i \in M = \{1,2,3\}$ can only be attained at the boundary of the region $[0,1]$. If there exists an optimal mode selection $q_n^*(t) \in (0,1)$, $n \in M$, there must be another $q_{n'}^*(t) \in (0,1)$, $n' \in M (n' \neq n)$ to satisfy the constraint of $\sum_{i=1}^{3} q_i(t) = 1$. Without loss of generality, we assume that $q_1^*(t)$ and $q_3^*(t) \in (0,1)$, the similar analysis can be readily extended to other cases. According to the complementary slackness of KKT conditions, the associated Lagrange multipliers are zeros when the related inequalities do not hold, i.e., $\upsilon_1(t) = \upsilon_3(t) = 0$. Now (46a) and (46c) can be rewritten as follows

$$\Lambda_1(t) - \upsilon(t) = 0, \tag{47a}$$

$$\sum_{i=1}^{2}\left[\frac{\mu_{ri}(t)C_{ri}(t)}{q_3(t)} - \frac{\mu_{ri}(t)E_r(t)s_i(t)}{(E_r(t)s_i(t)+q_3(t))\ln 2}\right] + \lambda_r^{M_3}(t)\hat{P}_r - \upsilon(t) = 0. \tag{47b}$$

We assume that channel gains $|h_1(t)|^2$ and $|h_2(t)|^2$ have continuous probability density functions. Besides, the Lagrange multipliers $\lambda_1^{M_1}(t)$, $\lambda_2^{M_1}(t)$, and $\lambda_r^{M_3}(t)$ are related to fading gains. Hence, we can hardly find such a $\upsilon(t)$ to satisfy the above two constraints at the same time for (47a) and (47b), so $q_i(t), i = 1,2,3$ must be attained at the boundary of the region $[0,1]$.

Next, we will show the necessary conditions for $q_1^*(t) = 1$ and $q_2^*(t) = q_3^*(t) = 0$, and the similar analysis can be readily generalized to the case of $q_i^*(t) = 1, i = 2,3$. By substituting $q_1^*(t) = 1, q_2^*(t) = q_3^*(t) = 0$ into (46) and considering that $\frac{\partial \mathcal{L}}{q_i(t)}$ is a monotonically increasing function with respect to $q_i(t)$ ($\mathcal{L}$ is convex), we have

$$0 = \frac{\partial \mathcal{L}}{\partial q_1(t)}|_{q_1^*=1} = -\Lambda_1(t) + \upsilon_1(t) + \upsilon(t),$$

$$0 = \frac{\partial \mathcal{L}}{\partial q_2(t)}|_{q_2^*=0} \leq \frac{\partial \mathcal{L}}{\partial q_2(t)}|_{q_2=1} = -\Lambda_2(t) - \upsilon_2(t) + \upsilon(t),$$

$$0 = \frac{\partial \mathcal{L}}{\partial q_3(t)}|_{q_3^*=0} \leq \frac{\partial \mathcal{L}}{\partial q_3(t)}|_{q_3=1} = -\Lambda_3(t) - \upsilon_3(t) + \upsilon(t).$$

Now we may derive $\Lambda_1(t) - \Lambda_2(t) \geq \upsilon_1(t) + \upsilon_2(t)$ and $\Lambda_1(t) - \Lambda_3(t) \geq \upsilon_1(t) + \upsilon_3(t)$. Since the Lagrange multipliers of the inequality constraints must be non-negative, i.e., $\upsilon_i(t) \geq 0, i = 1,2,3$, the necessary condition for $q_1^*(t) = 1$ is $\Lambda_1(t) = \max\{\Lambda_1(t), \Lambda_2(t), \Lambda_3(t)\}$. In the same way, we may have the necessary conditions for $q_2^*(t) = 1$ and $q_3^*(t) = 1$, which corresponds to the optimal mode selection policy in (15).

*2) The Optimal Rate Allocation Policy Derivation:* The optimal rate allocation scheme in the multiple-access and the broadcast mode should fulfill the following conditions

$$\frac{\partial \mathcal{L}}{R_{ir}(t)} = -\mu_i + \mu_{ir}(t) + \mu_r(t) - l_{ir}(t) = 0, \tag{48a}$$

$$\frac{\partial \mathcal{L}}{R_{ri}(t)} = -\alpha_j + \mu_{ri}(t) - l_{ri}(t) = 0, i,j = 1,2, i \neq j, \tag{48b}$$

where $\mu_1 = \theta - \alpha_1$ and $\mu_2 = 1 - \theta - \alpha_2$. Firstly, we assume $R_{1r}(t) > 0$ and $R_{2r}(t) > 0$, then $l_{1r}(t) = l_{2r}(t) = 0$. Either $\mu_{1r}(t)$ or $\mu_{2r}(t)$ is equal to zero due to the limit of multiple-access channel. If $\mu_{1r}(t) > 0$ and $\mu_{2r}(t) = 0$, we have

$$R_{1r}(t) = C(P_1(t)s_1(t)), \tag{49a}$$

$$R_{2r}(t) = C(P_2(t)s_2(t)/(1 + P_1(t)s_1(t))), \tag{49b}$$

$$\mu_{1r}(t) = \mu_1 - \mu_2, \quad \mu_r(t) = \mu_2. \tag{49c}$$

Similarly, if $\mu_{2r}(t) > 0$ and $\mu_{1r}(t) = 0$, we have

$$R_{1r}(t) = C(P_1(t)s_1(t)/(1 + P_2(t)s_2(t))), \tag{50a}$$

$$R_{2r}(t) = C(P_2(t)s_2(t))), \tag{50b}$$

$$\mu_{2r}(t) = \mu_2 - \mu_1, \quad \mu_r(t) = \mu_1. \tag{50c}$$

The same analysis can be extended to the case of $R_{1r}(t) > 0, R_{2r}(t) = 0$ and $R_{1r}(t) = 0, R_{2r}(t) > 0$. Moreover, based on the Lagrange multipliers $\mu_j \geq 0, \forall j \in J$, we can obtain the optimal rate allocation scheme in the multiple-access mode. Similarly, we can obtain the optimal rate allocation scheme from the relay to users, which is given in (17).

*3) The Optimal Power Allocation Derivation:* Since $q_i^*(t), i = 1,2,3$ is either one or zero. We can obtain the optimal transmit power in the corresponding transmission mode when $q_i^*(t) = 1$. Based on the KKT conditions, the optimal $E_i^{M_1}(t), i = 1,2$, satisfies the following equations

$$l_i^{M_1} E_i^{M_1}(t) = 0 \text{ and } \lambda_i^{M_1}(E_i^{M_1}(t) - \hat{P}_i) = 0, \tag{51a}$$

$$\gamma_i - \frac{\gamma_r |h_i(t)|^2 \eta}{d_i^m} - l_i^{M_1}(t) + \lambda_i^{M_1}(t) = 0, \ i = 1,2. \tag{51b}$$

From (51a) we know that $l_i^{M_1}(t) = \lambda_i^{M_1}(t) = 0$ if $0 < E_i^{M_1}(t) < \hat{P}_i$. And we have $\gamma_i = \frac{\gamma_r |h_i(t)|^2 \eta}{d_i^m}$. Nonetheless, it is impossible to attain $\gamma_i = \frac{\gamma_r |h_i(t)|^2 \eta}{d_i^m}$ since either $|h_1(t)|^2$ or $|h_2(t)|^2$ is time varying. Thus, the optimal $E_i^{M_1}(t)$ can only be zero or $\hat{P}_i$. We have $l_i^{M_1}(t) = 0$ and $\lambda_i^{M_1}(t) \geq 0$ if $E_i^{M_1}(t) = \hat{P}_i$. Similarly, we have $l_i^{M_1}(t) \geq 0$ and $\lambda_i^{M_1}(t) = 0$ if $E_i^{M_1}(t) = 0$. So the optimal power allocation in the EH mode is given in (18).

$$E_i^{M_1}(t) = \begin{cases} \hat{P}_i, & \text{if } \frac{\gamma_r |h_i(t)|^2 \eta}{d_i^m} - \gamma_i \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{52}$$

To derive $E_i^{M_2}(t)$ of $U_i$ in the multiple-access mode, we have

$$\lambda_i^{M_2}(t)(E_i^{M_2}(t) - \hat{P}_i) = 0 \text{ and } l_i^{M_2}(t)E_i^{M_2}(t) = 0, \tag{53a}$$

$$\gamma_i + \lambda_i^{M_2}(t) - l_i^{M_2}(t) - \frac{\mu_{ir}(t)s_i(t)}{(1 + E_i^{M_2}(t)s_i(t))\ln 2} - \frac{\mu_r(t)s_i(t)}{(1 + E_1^{M_2}(t)s_1(t) + E_2^{M_2}(t)s_2(t))\ln 2} = 0, \ i = 1,2. \tag{53b}$$

There are three possibilities: i). $E_i^{M_2}(t) = 0$, ii). $0 < E_i^{M_2}(t) < \hat{P}_i$, iii). $E_i^{M_2}(t) = \hat{P}_i$. We can separately analysis the optimal transmit powers of the users in the multiple-access mode based on nine cases of the values of $(E_1^{M_2}(t), E_2^{M_2}(t))$, which are given in (19) and we omit the detailed proof process here to save space. In fact, by the KKT conditions, the optimal transmit power of relay should satisfy the following conditions

$$-\frac{\mu_{r1}(t)s_1(t)}{(1+E_r(t)s_1(t))\ln 2}-\frac{\mu_{r2}(t)s_2(t)}{(1+E_r(t)s_2(t))\ln 2}+\gamma_r - l_r^{M_3}(t)+\lambda_r^{M_3}(t)=0, \quad (54a)$$

$$l_r^{M_3}(t)E_r(t)=0 \text{ and } \lambda_r^{M_3}(t)(E_r(t)-\hat{P}_r)=0. \quad (54b)$$

We can derive that $E_r(t)=0$ if $l_r^{M_3}(t)\geq 0$ and $\lambda_r^{M_3}(t)=0$, $E_r(t)=\hat{P}_r$ if $l_r^{M_3}(t)=0$ and $\lambda_r^{M_3}(t)\geq 0$. Moreover, we have $l_r^{M_3}(t)=\lambda_r^{M_3}(t)=0$ if $0<E_r(t)<\hat{P}_r$. And the left hand side of (54a) is monotonically increasing in $E_r(t)$ with the minimum value $\gamma_r-\frac{\mu_{r1}(t)s_1(t)+\mu_{r2}(t)s_2(t)}{\ln 2}$ and the maximum value $\gamma_r-\frac{\mu_{r1}(t)s_1(t)}{(1+\hat{P}_r s_1(t))\ln 2}-\frac{\mu_{r2}(t)s_2(t)}{(1+\hat{P}_r s_2(t))\ln 2}$. One root of (54a) will be positive while the other one is negative if $\frac{\mu_{r1}(t)s_1(t)}{1+\hat{P}_r s_1(t)}+\frac{\mu_{r2}(t)s_2(t)}{1+\hat{P}_r s_2(t)}<\gamma_r\ln 2<\mu_{r1}(t)s_1(t)+\mu_{r2}(t)s_2(t)$. Otherwise both roots are negative or greater than $\hat{P}_r$. Thus, we have the optimal transmit power of relay in (20).

## Appendix B

It is worth noting that

$$\left(\phi-\max\{E(t)+E_h(t)-q_3(t)P_r(t)T,0\}\right)^2 \leq (\phi-E(t)-E_h(t)+q_3(t)P_r(t)T)^2. \quad (55)$$

From (29), (30), and (31), we have

$$\begin{aligned}\Delta(\mathbf{\Theta}(\mathbf{t})) \leq\ & B+\psi(\phi-E(t)\mathbb{E}\{q_3(t)P_r(t)T-E_h(t)|\mathbf{\Theta}(\mathbf{t})\}\\ &+\sum_{i=1}^{2}Z_i(t)\mathbb{E}\{(q_1(t)+q_2(t))P_i(t)-\bar{P}_i^{\max}|\mathbf{\Theta}(\mathbf{t})\}\\ &+Q_1(t)\mathbb{E}\{R_{1r}(t)-R_{r2}(t)|\mathbf{\Theta}(\mathbf{t})\}\\ &+Q_2(t)\mathbb{E}\{R_{2r}(t)-R_{r1}(t)|\mathbf{\Theta}(\mathbf{t})\}. \end{aligned} \quad (56)$$

By adding the penalty item to both sides of (56), we derive Lemma 1.

## Appendix C

Since we assume that the transmit rate is strictly interior of capacity region and channel states are $i.i.d$ over each time slot, there exits a stationary randomized mode selection, power allocation, and rate allocation policy, which is independent of $\mathbf{\Theta}(\mathbf{t})$ [26] satisfying

$$\mathbb{E}\{R_{sum}(t)|\mathbf{\Theta}(\mathbf{t})\}=\mathbb{E}\{R_{sum}(t)\}=\Psi(\epsilon), \quad (57a)$$

$$\mathbb{E}\{q_3(t)P_r(t)T-E_h(t)|\mathbf{\Theta}(\mathbf{t})\} =\mathbb{E}\{q_3(t)P_r(t)T-E_h(t)\}\leq-\epsilon, \quad (57b)$$

$$\mathbb{E}\{(q_1(t)+q_2(t))P_i(t)-\bar{P}_i^{\max}|\mathbf{\Theta}(\mathbf{t})\} =\mathbb{E}\{(q_1(t)+q_2(t))P_i(t)-\bar{P}_i^{\max}\}\leq-\epsilon, i=1,2, \quad (57c)$$

$$\mathbb{E}\{R_{ir}(t)-R_{rj}(t)|\mathbf{\Theta}(\mathbf{t})\} =\mathbb{E}\{R_{ir}(t)-R_{rj}(t)\}\leq-\epsilon, i,j=1,2, \text{ and } i\neq j. \quad (57d)$$

By substituting (57) into (32), we obtain

$$\Delta(\mathbf{\Theta}(\mathbf{t}))-V\mathbb{E}\{R_{sum}(t)|\mathbf{\Theta}(\mathbf{t})\} \leq B-V\Psi(\epsilon)-\psi(\phi-E(t))\epsilon-\sum_{i=1}^{2}(Q_i(t)+Z_i(t))\epsilon. \quad (58)$$

Using the law of iterated expectation and taking conditional expectations over $\mathbf{\Theta}(\mathbf{t})$ yields:

$$\mathbb{E}\{L(\mathbf{\Theta}(\mathbf{t+1}))-L(\mathbf{\Theta}(\mathbf{t}))\}-V\mathbb{E}\{R_{sum}(t)\}\leq B-V\Psi(\epsilon)-\psi(\phi-\mathbb{E}\{E(t)\})\epsilon-\sum_{i=1}^{2}\mathbb{E}\{Q_i(t)+Z_i(t)\}\epsilon. \quad (59)$$

Then, summing over $t\in\{0,1,...,N-1\}$ and dividing by $N$, we have

$$\frac{\mathbb{E}\{L(\mathbf{\Theta}(\mathbf{N}))-L(\mathbf{\Theta}(\mathbf{0}))\}}{N}-\frac{V}{N}\sum_{t=0}^{N-1}\mathbb{E}\{R_{sum}(t)\}\leq B -V\Psi(\epsilon)-\frac{\epsilon}{N}\sum_{t=0}^{N-1}\left(\psi(\phi-\mathbb{E}\{E(t)\})+\sum_{i=1}^{2}\mathbb{E}\{Q_i(t)+Z_i(t)\}\right).$$

The time average expectation for $R_{sum}(t)$ of the DAAT scheme can not exceed the optimal $R^*_{sum}$, i.e., $\lim_{N\to\infty}\frac{1}{N}\sum_{t=0}^{N-1}\mathbb{E}\{R_{sum}(t)\}\leq R^*_{sum}$. Moreover, using the fact $L(\mathbf{\Theta}(\mathbf{N}))\geq 0$, $L(\mathbf{\Theta}(\mathbf{0}))=0$, $Q_i(t)\geq 0$, $Z_i(t)\geq 0$, and $\phi-E(t)\geq 0$ yields

$$\frac{1}{N}\sum_{t=0}^{N-1}\mathbb{E}\{Q_1(t)+Q_2(t)\}\leq\frac{B+V[R^*-\Psi(\epsilon)]}{\epsilon}. \quad (60)$$

Moreover, we note that the queue length is not less than zero so that we have

$$\Psi(\epsilon)-\frac{B}{V}\leq\frac{1}{N}\sum_{t=1}^{N-1}\mathbb{E}\{R_{sum}(t)\}. \quad (61)$$

Taking a limit as $N\to\infty$ and $\Psi(\epsilon)\to R^*_{sum}$ as $\epsilon\to 0$ for (60) and (61), we can conclude Theorem 2.

## References


[1] H. Nishimoto *et al.*, "Prototype implementation of ambient RF energy harvesting wireless sensor networks," in *Proc. IEEE Sensors*, 2010, pp. 1282–1287.

[2] X. Lu, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Wireless charger networking for mobile devices: Fundamentals, standards, and applications," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 126–135, Apr. 2015.

[3] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2012.

[4] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1180–1189, Mar. 2012.

[5] C. Huang, R. Zhang, and S. Cui, "Throughput maximization for the Gaussian relay channel with energy harvesting constraints," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1469–1479, Aug. 2013.

[6] L. Liu, R. Zhang, and K. C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 288–300, Jan. 2013.

[7] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[8] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[9] D. S. Michalopoulos, H. A. Suraweera, and R. Schober, "Relay selection for simultaneous information transmission and wireless energy transfer: A tradeoff perspective," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1578–1594, Aug. 2015.

[10] B. Xia *et al.*, "Buffering in a three-node relay network," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4492–4496, Nov. 2008.

[11] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-aided relaying with adaptive link selection," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1530–1542, Aug. 2013.

[12] 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.

[13] P. Liu and I. M. Kim, "Performance analysis of bidirectional communication protocols based on decode-and-forward relaying," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2683–2696, Sep. 2010.

[14] Y. Wu, P. A. Chou, and S. Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proc. 39th Ann. Conf. Inf. Sci. Syst.*, 2005, pp. 1–6.

[15] I. Krikidis, "Relay selection for two-way relay channels with MABC DF: A diversity perspective," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4620–4628, Nov. 2010.

[16] H. Liu *et al.*, "Sum-rate optimization in a two-way relay network with buffering," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 95–98, Jan. 2013.

[17] V. Jamali, N. Zlatanov, A. Ikhlef, and R. Schober, "Achievable rate region of the bidirectional buffer-aided relay channel with block fading," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7090–7111, Nov. 2014.

[18] K. Xiong, P. Fan, and K. B. Letaief, "Time-switching based SWPIT for network-coded two-way relay transmission with data rate fairness," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, South Brisbane, QLD, 2015, pp. 5535–5539.

[19] R. Jiang *et al.*, "Outage performance of SWIPT-enabled two-way relay networks," in *Proc. IEEE Int. Workshop High Mobility Wireless Commun.*, 2015, pp. 106–110.

[20] B. Varan and A. Yener, "The energy harvesting two-way decode-and-forward relay channel with stochastic data arrivals," in *Proc. IEEE GlobalSIP*, Austin, TX, 2013, pp. 371–374.

[21] B. Varan and A. Yener, "Delay constrained energy harvesting networks with limited energy and data storage," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1550–1564, May 2016.

[22] I. Ahmed, A. Ikhlef, D. W. K. Ng, and R. Schober, "Optimal resource allocation for energy harvesting two-way relay systems with channel uncertainty," in *Proc. IEEE GlobalSIP*, Austin, TX, 2013, pp. 345–348.

[23] S. J. Kim *et al.*, "Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6405–6418, Oct. 2011.

[24] V. Jamali, N. Zlatanov, A. Ikhlef, and R. Schobert, "Adaptive mode selection and power allocation in bidirectional buffer-aided relay networks," in *Proc. IEEE GLOBECOM*, 2013, pp. 1933–1938.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[26] M. J. Neely, *Stochastic Network Optimization With Application to Communication and Queueing Systems* (Synthesis Lectures on Communication Networks). San Rafael, CA, USA: Morgan and Claypool, 2010.

[27] F. Gao, T. Cui, and A. Nallanathan, "Maximum likelihood channel estimation in decode-and-forward relay networks," in *Proc. IEEE Int. Symp. Inf. Theory*, 2008, pp. 1233–1237.

[28] X. Lan, Q. Chen, and X. Tang, "On the transmit power and delivery delay tradeoff in buffer-aided two-way relay networks," in *Proc. IEEE/CIC Int. Conf. Commun. China*, 2016, pp. 1–6.

[29] T. D. Quoc and M. Diehl, "Sequential convex programming methods for solving nonlinear optimization problems with DC constraints," unpublished paper, 2011. [Online]. Available: https://arxiv.org/abs/1107.5841

[30] Y. Liu, Q. Chen, X. Tang, and L. X. Cai, "On the buffer energy aware adaptive relaying in multiple relay networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6248–6263, Sep. 2017.

[31] X. Li, Y. Wu, and E. Serpedin, "Timing synchronization in decode-and-forward cooperative communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1444–1455, Apr. 2009.

[32] L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis, "Secrecy cooperative networks with outdated relay selection over correlated fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7599–7603, Aug. 2017.


**Xiaolong Lan** received the B.E. degree in mathematics and applied mathematics from the Chengdu University of Technology, Chengdu, China, in 2012. He is currently working toward the Ph.D. degree in the School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China. In 2017, he was a Visiting Ph.D. Student with the University of Victoria. His current research interests include buffer-aided communication and energy-harvesting wireless communication.

**Qingchun Chen** (SM'14) received the B.Sc. and M.Sc. (with Hons.) degrees from Chongqing University, Chongqing, China, in 1994 and 1997, respectively, and the Ph.D. degree from Southwest Jiaotong University, Chengdu, China, in 2004. He was with Southwest Jiaotong University from 2004 to 2018. He is currently a Full Professor with Guangzhou University, Guangzhou, China. He has authored and coauthored more than 100 research papers, two book chapters, and 40 patents. His research interest includes wireless communication, wireless network, information coding, and signal processing. He was the recipient of the 2016 IEEE GLOBECOM Best Paper Award. He was an Associate Editor for IEEE ACCESS (2015–present).

**Xiaohu Tang** received the B.S. degree in applied mathematics from the Northwest Polytechnic University, Xi'an, China, the M.S. degree in applied mathematics from the Sichuan University, Chengdu, China, and the Ph.D. degree in electronic engineering from the Southwest Jiaotong University, Chengdu, China, in 1992, 1995, and 2001, respectively. From 2003 to 2004, he was a Research Associate with the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology. From 2007 to 2008, he was a Visiting Professor with the University of Ulm, Germany. Since 2001, he has been with the School of Information Science and Technology, Southwest Jiaotong University, where he is currently a Professor. His research interests include coding theory, network security, distributed storage, and information processing for big data.

Dr. Tang was the recipient of the National excellent Doctoral Dissertation award in 2003 (China), the Humboldt Research Fellowship in 2007 (Germany), and the Outstanding Young Scientist Award by NSFC in 2013 (China). He is an Associate Editor for several journals including the IEEE TRANSACTIONS ON INFORMATION THEORY and *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, and was on a number of technical program committees of conferences.

**Lin Cai** (S'00–M'06–SM'10) received the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2002 and 2005, respectively. Since 2005, she has been with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, where she is currently a Professor. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic and Internet of things.

She was a recipient of the NSERC Discovery Accelerator Supplement (DAS) Grants in 2010 and 2015, respectively, and the Best Paper Awards of IEEE ICC 2008 and IEEE WCNC 2011. She has founded and chaired IEEE Victoria Section Vehicular Technology and Communications Joint Societies Chapter. She was a member of the Steering Committee of the IEEE TRANSACTIONS ON BIG DATA, an Associate Editor of the IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *EURASIP Journal on Wireless Communications and Networking*, *International Journal of Sensor Networks*, and *Journal of Communications and Networks*, and is the Distinguished Lecturer of the IEEE VTS Society. She was a TPC Symposium Co-Chair for IEEE Globecom'10 and Globecom'13. She is a registered Professional Engineer of British Columbia, Canada.