Optimal Relaying in Energy Harvesting Wireless Networks With Wireless-Powered Relays

Masoumeh Moradian, *Member, IEEE*, Farid Ashtiani, *Member, IEEE*, and Ying Jun Zhang, *Senior Member, IEEE*

***Abstract*—In this paper, we consider a wireless cooperative network with a wireless-powered energy harvesting (EH) relay. The relay employs a time switching (TS) policy that switches between the EH and data decoding (DD) modes. Both energy and data buffers are kept at the relay to store the harvested energy and decoded data packets, respectively. In this paper, we derive static and dynamic TS policies that maximize the system throughput or minimize the average transmission delay. In particular, in the static policies, the EH or DD mode is selected with a pre-determined probability. In contrast, in a dynamic policy, the mode is selected dynamically according to the states of data and energy buffers. We prove that the throughput-optimal static and dynamic policies keep the relay data buffer at the boundary of stability. More specifically, we show that the throughput-optimal dynamic policy has a threshold-based structure. Moreover, we prove that the delay-optimal dynamic policy is threshold-based and keeps at most one packet at the relay. We notice that unlike the static case, the delay-optimal and throughput-optimal dynamic policies coincide in most cases. Finally, through extensive numerical results, we demonstrate the efficiency of the optimal dynamic policies compared with the static ones.**

***Index Terms*—Energy harvesting, relaying, quasi-birth-death (QBD) process, threshold-based structure, delay, throughput, cooperative network.**

I. Introduction

ENERGY harvesting (EH), a technology to collect energy from the surrounding environment, has received considerable attention as a sustainable solution to prolong the lifetime of wireless networks under energy constraints. EH technology is especially appealing in networks with low-cost battery-powered devices, e.g., wireless sensor networks [1]. Unlike battery-powered networks, EH wireless networks potentially have an unlimited lifetime, thanks to the unlimited energy supply in the environment. Thus, with a relatively low cost, a set of EH relay nodes can be overlaid on an existing non-EH network to enhance the performance and quality of service (QoS).

In addition to conventional harvestable energy sources such as solar, wind, and thermal energy, ambient RF waves have emerged as a source of energy for wireless EH (WEH) nodes. Since radio waves carry both information and energy, simultaneous wireless information and power transfer (SWIPT) has been actively studied in [2], [3], which assume that a WEH node can decode information and harvest energy simultaneously from the same radio signal. However, due to practical hardware and circuit limitation, a radio signal that has been used for energy harvesting may not be reused for data decoding [4]. Correspondingly, time-switching (TS) protocols, which switch between EH and data decoding (DD) modes, have been studied in single-antenna point-to-point channels [5], in two-user MIMO interference channels [6], and in amplify-and-forward [7], [8] and decode-and-forward [9], [10] relaying networks.

In this paper, we focus on a three-node cooperative wireless network with a WEH relay and a direct link between source and destination. The WEH relay uses a TS protocol to switch between the EH and DD modes. The decisions for being in the DD and EH modes are taken at the beginning of the time slots. In the DD mode, the relay decodes each transmitted source packet and then, if no ACK message is broadcast by the destination, the relay stores the decoded packet in its data buffer. In the EH mode, the relay does not decode any packet and only harvests energy from the ambient RF waves, e.g., incoming source packets or co-channel interference [5], to replenish its energy buffer. For such a system, we aim to find the optimal TS policies that maximize the system throughput or minimize the average end-to-end transmission delay. In an EH node with data and energy buffers, the data and energy arrival processes should be balanced such that the desired long-term performance metrics are satisfied. For instance, in an EH node with non-RF energy sources, the energy arrival rate is determined by the environment. Then, to preserve the stability of the data buffer, the energy demand rate, indicated by the data arrival rate and the transmission power, should be less than or equal to the energy arrival rate [11]. Likewise, when satisfying a delay constraint, the distribution of the data inter-arrival time should also be managed in addition to the data arrival rate [12]–[14]. In a WEH node with TS-based receiver, the data and energy arrival processes are coupled to each other

Manuscript received February 4, 2019; revised June 10, 2019; accepted July 23, 2019. Date of publication July 30, 2019; date of current version November 20, 2019. This work was supported in part by the General Research Fund established by Hong Kong Research Grants Council under Project 14208017. The associate editor coordinating the review of this paper and approving it for publication was H. Suraweera. *(Corresponding author: Masoumeh Moradian.)*

M. Moradian was with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran, and also with the Department of Information Engineering, Chinese University of Hong Kong, Hong Kong. She is now with the Department of Computer Science, Institute for Research in Fundamental Sciences, Tehran 19538-33511, Iran (e-mail: mmoradian@ipm.ir).

F. Ashtiani is with the Department of Electrical Engineering, Advanced Communications Research Institute, Sharif University of Technology, Tehran 111554363, Iran (e-mail: ashtianimt@sharif.edu).

Y. J. Zhang is with the Department of Information Engineering, Chinese University of Hong Kong, Hong Kong (e-mail: yjzhang@ie.cuhk.edu.hk).

Digital Object Identifier 10.1109/TGCN.2019.2932040

due to switching between DD and EH circuits in time. Then, the TS policy can balance the arrival processes to data and energy buffers by taking the proper switching decisions. Note that when the TS policy switches to the EH mode, the arrival of data packets is suspended and the energy arrivals become possible. Likewise, when the DD mode is selected, the data arrival is resumed and the energy arrival is suspended. If the switching decisions are not designed carefully at the relay, the energy may be harvested in extra amounts, or an undesirable amount of data may be backlogged in the data buffer. Each of the mentioned phenomena degrades the performance of the relaying. In this paper, we aim to optimize the performance of such a relaying system and derive the optimized TS policies. To tackle this challenge, we propose a quasi-birth-death (QBD) process to analyze the throughput and delay as a function of the TS policy. Optimal static and dynamic TS policies are then derived for throughput maximization and delay minimization, respectively.

The main contributions of this paper are summarized as follow:

- We derive optimal static TS policies, in which the modes are selected based on a pre-determined probability regardless of the states of the buffers. We prove that the throughput-optimal static policy is the one that keeps the data buffer at the WEH relay at the boundary of stability. Furthermore, the delay-optimal static policy is obtained by analyzing an underlying QBD process. We also derive the necessary and sufficient condition under which the non-cooperation policy is delay-optimal.
- We derive optimal dynamic TS policies, where the modes are selected dynamically based on the states of the data and energy buffers. We prove that both the throughput-optimal and delay-optimal policies are threshold-based in terms of the status of the energy buffer. Moreover, the delay-optimal dynamic policy is the one that keeps at most one packet in the data buffer. Interestingly, we observe that the throughput-optimal and delay-optimal dynamic TS policies are the same in most cases. However, this is not the case for static TS policies. By simulations, we validate our analyses and compare the performance of the optimal static and dynamic policies in different conditions.

A. Related Works

In [7]–[10], [15], the maximum achievable throughput of the source is derived in WEH cooperative networks that adopt a TS policy at the relay. However, none of them include a data buffer at the relay. In fact, the information decoded by the relay in a time slot may not be transmitted immediately afterwards, because the energy harvested in a slot may not be sufficient. Thus, a data buffer is necessary at the relay, so that the packet can be stored until enough energy is harvested. Continuous and discrete time EH are investigated in [7] and [8], respectively. In continuous time EH, the energy required for transmission in each slot is harvested in a variable portion of the same slot, while in discrete time EH, the whole slot is dedicated to either EH or data transmission. These two schemes are compared in [15]. In [9], the relay harvests energy from interference signals and source transmissions in a continuous EH manner. The authors in [10] considered discrete EH in a network with multiple relays, where all relays cooperate with each other to transmit the information of the source. In [16], [17], a selected set of relays cooperate to transmit the data of the source while all relays are equipped with finite energy storage. The authors of [16] and [17] propose centralized and distributed relay selection schemes, respectively, to minimize the source outage probability.

The delay performance of cooperative networks with EH relays has been considered in [12], [13], [18]–[20]. In [12], [13], [18], the average transmission delay in a wireless network with cooperative EH relay(s) is minimized. The relay, however, harvests energy from the environmental resources instead of RF waves, and thus the arrival processes of data and energy buffers are not correlated. The authors in [19] minimize the transmission completion time of a fixed amount of data backlogged at the source in a cooperative network with multiple WEH relays. They optimize the transmission order and time allocation between the source and relays to minimize the transmission time of the limited backlogged data. In contrast, the source is continuously backlogged in our scenario. Thus, unlike [19], the data buffer at node R may become unstable due to applying an inappropriate TS policy. In [20], the authors derive the optimal offline and online policies for delay and power minimization under a deadline constraint, when the EH relay is equipped with finite data and energy buffers. In contrast, in this paper, we optimize the long-term performance metrics, e.g., the average transmission delay of the packets while preserving the stability of the data buffer at the relay node.

The authors in [14], [21] investigate scenarios in which the WEH nodes are equipped with data and energy buffers and empowered by a power source (PS) in a TS manner. In particular, the PS transfers energy to the nodes wirelessly in a fraction of each time frame. Then, in the rest of the frame, the nodes transmit data in the uplink. In [14], [21], the data is generated at the data queues of the nodes independently and is not affected by the wireless power transfer protocol, i.e., the data arrival process at the nodes is independent of the energy arrival process. In [22], the authors consider a two-hop cooperative network in which the source communicates with the destination through a WEH relay, equipped with data and energy buffers. Then, the optimal TS policy is derived to maximize the source throughput, while limiting the average delay. In [22], there exists a trade-off between the throughput and the average delay, i.e., higher throughput is achieved at higher delays, since there is not a direct link between the source and destination. The authors employ the Lyapunov optimization to characterize the existing trade-off, while guaranteeing the stability of the data and energy queues. In our scenario, there is a direct link between the source and destination. We show that in this case, the throughput and delay-optimal dynamic policies can be achieved at the same time in most practical situations.

The rest of the paper is organized as follows. In Section II, we introduce the system model, the underlying QBD process,

TABLE I
PARAMETERS

Parameter	Explanation
$\mathrm{Q_d}$	Data buffer at node R
$\mathrm{Q_e}$	Energy buffer at node R
$p_{\mathrm{XY}}^{\mathrm{det}}$	Detection probability of the channel X-Y
N	Size of the energy buffer, $\mathrm{Q_e}$
K	Number of energy units used for a packet transmission
$b_{\max}$	Maximum number of energy units harvested in a slot
Γ	Random variable denoting the number of energy units harvested in a slot in EH mode
γ_i	Probability of harvesting i energy units in a slot
τ	Average transmission delay of the source packets
q_{d}	Number of data packets backlogged at node R
q_{e}	Number of energy units backlogged at node R
$\overline{D_{\mathrm{R}}}$	Average system delay at node R
λ_{S}	Throughput of node S
λ_{id}	Data arrival rate at node R
λ_{od}	Data departure rate at node R
λ_{ie}	Energy arrival rate at node R after blocking
λ_{oe}	Energy departure rate at node R
α_s	Probability of switching to the DD mode at state s
$p_{\mathrm{R}}^{\mathrm{a}}$	Probability that node R is active
p_{b}	Blocking probability of the energy buffer
α^{T}	Throughput-optimal DD probability in the static policy
α^{D}	Delay-optimal DD probability in the static policy
α_s^{T}	Throughput-optimal DD probability at state s in the dynamic policy
α_s^{D}	Delay-optimal DD probability at state s in the dynamic policy

and the desired performance metrics. Section III presents the derivation of optimal static policies. Section IV discusses how we find the optimal dynamic policies. Section V presents the numerical results. Finally, Section VI concludes the paper. Also, the notations are introduced in Table I.

II. System Model

A. Network Scenario

We consider a three-node cooperative wireless network comprised of a source (S), a relay (R), and a destination (D) node, as shown in Fig. 1(a). All nodes are within the transmission range of each other. Suppose that nodes S and D are battery-powered and node S is infinitely backlogged, i.e., it has always a packet to transmit. On the other hand, node R is an energy harvesting (EH) node, which harvests its energy from ambient RF waves, e.g., source transmissions, an RF energy source or interference signals, and is able to store the harvested energy in a rechargeable battery. Suppose that node R is equipped with one antenna and two different circuits, one for EH and the other for DD. Thus, node R can only transmit data packets, receive data packets, or harvest energy at a particular time. Node R may relay the data packets of the source, but it does not have its own traffic. The data and energy buffers at node R are denoted by $\mathrm{Q_d}$ and $\mathrm{Q_e}$, respectively. While the data buffer is assumed to have infinite size, we assume that the energy buffer size is finite due to the finite battery capacity in practice. The assumption of infinite data buffer is to facilitate the stability analysis of the data buffer. The assumption is reasonable when the buffer is sufficiently large in practice.

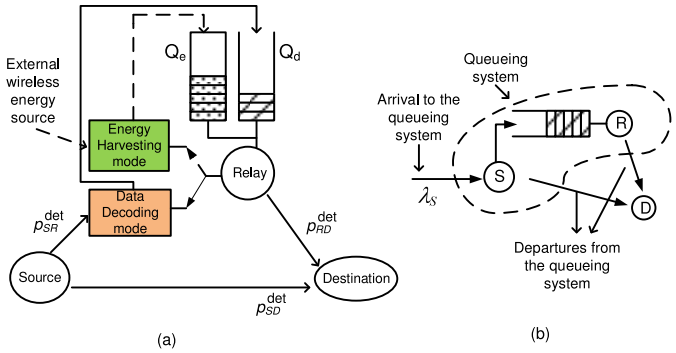


Fig. 1. (a) Network scenario, (b) Equivalent queueing system.

We consider slotted time and each time slot is further divided into two equal subslots. Moreover, the channel between node X and node Y ($\mathrm{X}, \mathrm{Y} \in \{\mathrm{S}, \mathrm{R}, \mathrm{D}\}$), i.e., the channel X-Y, is a block fading one, i.e., it remains constant during each slot. However, the channel states are not known at node R and the TS policy is designed based on the average values of the packet detection probabilities. In this regard, $p_{\mathrm{XY}}^{\mathrm{det}}$ denotes the average packet detection probability of the channel X-Y. $p_{\mathrm{XY}}^{\mathrm{det}}$ depends on the physical parameters, e.g., transmission power of node X, modulation and coding scheme at node X, channel X-Y distribution, the distances, etc. In Section V, we consider practical block fading channels and derive $p_{\mathrm{XY}}^{\mathrm{det}}$ by considering different physical parameters. The transmission strategies of nodes S, R and D are described as follows.

Node S: Node S transmits a data packet in the first subslot of each time slot and remains silent in the second subslot. If it receives an ACK from either node R or node D in the first subslot, then the current packet is removed from its buffer, and it will transmit a new packet in the next time slot. Otherwise, the current packet will be retransmitted in the next time slot.

Node D: Upon receiving a data packet successfully, it broadcasts an ACK message. Otherwise, it broadcasts a NACK message.

Node R: At the beginning of each slot, node R decides whether to operate in the DD mode or EH mode. When operating in the DD mode, node R decodes the packet of node S in the first subslot. When node R successfully decodes a packet in the DD mode, it listens to the ACK message from node D. If an ACK is received at the end of the first subslot, node R infers that the packet is successfully decoded at node D and discards the packet. On the other hand, if an ACK is not received, node R stores the packet in its data buffer, $\mathrm{Q_d}$, and sends an ACK to node S. When operating in the EH mode, node R harvests energy from an ambient RF power source in the first subslot. In either mode, node R transmits in the second subslot when it has at least one data packet in its data buffer and enough energy units backlogged in its energy buffer. In this case, we say that node R is active.

Suppose that node R consumes *K* units of energy to transmit a data packet. The data packets are transmitted on a first-come first-serve (FCFS) basis. If a packet is not transmitted successfully in a slot, it will remain at the head-of-line of the data buffer and be retransmitted in a following

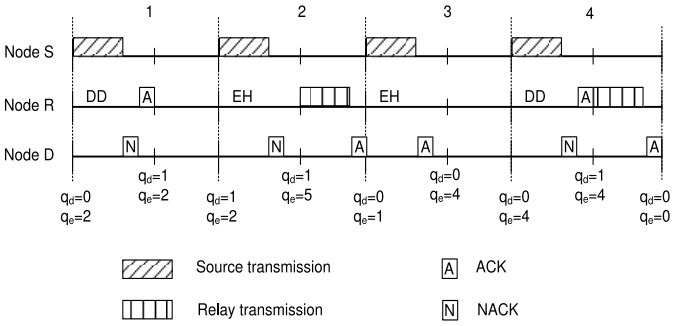


Fig. 2. Transmission protocol ($b_{\max} = 3$, $K = 4$).

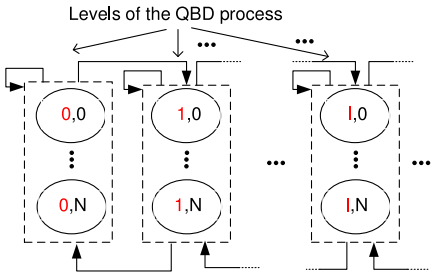


Fig. 3. The Underlying QBD model.

time slot whenever there exists at least K energy units in the energy buffer.

A realization of the transmission strategies of the nodes is illustrated in Fig. 2, where $q_{\rm d}$ and $q_{\rm e}$ denote the number of data packets and energy units in the data and energy buffers, respectively, at the beginning of a time slot. Suppose that there are initially two energy units stored in $\rm Q_e$, $K = 4$, and $p_{\rm SR}^{\rm det} = 1$. As shown in Fig. 2, node R is in the DD mode in the first slot. Upon hearing NACK in the first subslot, node R stores the received source packet in $\rm Q_d$ and transmits an ACK. However, due to the lack of energy, it cannot transmit the packet in the second subslot. In the second slot, node R switches to the EH mode, harvests enough energy in the first subslot, and transmits the backlogged source packet successfully in the second subslot. In the third slot, node R remains in the EH mode to harvest more energy. Finally, it switches to the DD mode in the forth slot, receives a source packet, and transmits it successfully in the same slot.

The energy harvested in each slot is a continuous variable. However, for the sake of queuing analysis, we assume that the energy is harvested in discrete amounts called energy units. Let Γ be the random variable denoting the number of energy units harvested in a slot when node R is in the EH mode. Also, define $\gamma_m = Pr\{\Gamma = m\}$. In Section V, we discretize a practical continuous EH profile to derive the probability mass function of Γ, i.e., γ_m. Note that the stochastic nature of Γ is due to the randomness in wireless channels. Also, by the assumption of block fading channels, Γ is independently and identically distributed over different slots. In the case of practical channel models, e.g., Rayleigh, the harvested energy is proportional with the wireless channel gain and has an exponential distribution. Then, the probability of harvesting large amounts of energy is very low. Thus, we truncate the distribution of the harvested energy and assume that there is a maximum number of energy units that can be harvested in a slot. In fact, we assume that $\gamma_m = 0$ for $m > b_{\max}$, i.e., $b_{\max}$ is the maximum number of energy units that can be harvested in a slot.

It is worth mentioning that in our work, we consider a network-layer scenario, where the relay transmits fixed-size packets with a specified power. The transmission power in such a scenario (e.g., ∼0dBm in a wireless sensor node [23]) is generally higher than the power harvested from RF waves at each slot, which ranges from a few microwatts to a few hundreds of microwatt [24]. This is because the RF waves are severely attenuated due to the path loss and fading. Moreover, a small fraction of the power emitted by a wireless power source at a specific distance can be collected at the antenna of a typical wireless device, e.g., a sensor node. Thus, it is reasonable to assume that $b_{\max} \leq K$. Furthermore, we confirm the validity of this assumption by considering real EH profile for our simulations in Section V and comparing the analytical and simulation results. In addition, we assume that the size of the energy buffer, i.e., N, is large enough to store the energy needed for at least two transmissions, i.e., $N \geq 2K$. This assumption is also valid in scenarios where the capacity of the energy buffer (e.g., rechargeable battery) is not very small. Other than being realistic, the two assumptions, $N \geq 2K$ and $b_{\max} \leq K$, enables us to analytically express the throughput-optimal and delay-optimal transmission policies at node R.

B. Mode Selection and the Underlying QBD Process

Define the system state at the beginning of each time slot by an ordered-pair $(q_{\rm d}, q_{\rm e})$. At each slot, based on the current state, node R switches to either the DD or EH mode. The system state evolves according to a Markov chain and is modeled by a QBD process shown in Fig. 3. According to the convention, the first entry of system state denotes the level of the QBD and the second entry denotes its phase [28]. There are an infinite number of levels and a finite number of phases in the QBD, corresponding to our assumption on infinite data buffer size and finite energy buffer size. Moreover, transitions occur between the states only within the same or adjacent levels. This corresponds to our assumption that at most one data packet arrives at or departs from $\rm Q_d$ at each slot.

Let α_s denote the probability of switching to the DD mode at state s. Moreover, $\mathbf{M}_{ij}^{-}$ and $\mathbf{M}_{ij}$ represent the probabilities that the number of data packets decreases by one and remains unchanged, respectively, in the second subslot, while the energy state changes from i to j, given that $\rm Q_d$ is backlogged at the beginning of the second subslot. $\mathbf{M}^{-}$ and $\mathbf{M}$ are derived in (32) and (33), respectively. Likewise, the energy state in the first subslot changes according to the transition matrix $\mathbf{T}$ in (34), given that node R selects EH mode. Then, the transition probability from a state in the first level of the QBD, i.e., (0, i), to other states (l, j), denoted by $P_{0i \to lj}$, is

derived as (please refer to Appendix A for more details)

$$P_{0i\to lj}=\begin{cases}\alpha_s\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}\mathbf{M}_{ij}^{-}+\alpha_s\big(p_{\rm SD}^{\rm det}\\ +\big(1-p_{\rm SD}^{\rm det}\big)\big(1-p_{\rm SR}^{\rm det}\big)\big)\mathbf{I}_{ij}+(1-\alpha_s)\mathbf{T}_{ij};& l=0,\\ \alpha_s\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}\mathbf{M}_{ij};& l=1,\end{cases}\tag{1}$$

where $\mathbf{I}$ is the identity matrix. Furthermore, define $\mathbf{B}_{ij}^{-}$ and $\mathbf{B}_{ij}$ to be the probabilities that the number of data packets decreases by one and remains unchanged in the second subslot, respectively, and the energy buffer state changes from *i* to *j* in a slot, given that node R selects EH mode in the first subslot and the data buffer at node R is backlogged at the beginning of the second subslot. In particular, $\mathbf{B}=\mathbf{T}\times\mathbf{M}$ and $\mathbf{B}^{-}=\mathbf{T}\times\mathbf{M}^{-}$. Then, the transition probability from state $s=(l,\,i)$ to state (l',j) $(l>0)$, represented by $P_{li\to l'j}$, is written as

$$P_{li\to l'j}=\begin{cases}\alpha_s\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}\mathbf{M}_{ij}^{-}+\alpha_s\big(\big(1-p_{\rm SD}^{\rm det}\big)\big(1-p_{\rm SR}^{\rm det}\big)\\ +p_{\rm SD}^{\rm det}\big)\mathbf{M}_{ij}+(1-\alpha_s)\mathbf{B}_{ij};& l'=l,\\ \alpha_s\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}\mathbf{M}_{ij};& l'=l+1,\\ \alpha_s\big(\big(1-p_{\rm SD}^{\rm det}\big)\big(1-p_{\rm SR}^{\rm det}\big)+p_{\rm SD}^{\rm det}\big)\mathbf{M}_{ij}^{-}\\ +(1-\alpha_s)\mathbf{B}_{ij}^{-};& l'=l-1.\end{cases}\tag{2}$$

C. Throughput and Average Transmission Delay

In this subsection, we define two performance metrics, namely, source throughput and average transmission delay. In particular, source throughput, denoted by $\lambda_{\rm S}$, is defined as the rate at which the packets are successfully detected at node D. By the assumption that node S is infinitely backlogged, $\lambda_{\rm S}$ can be derived as

$$\lambda_{\rm S}=p_{\rm SD}^{\rm det}+p_{\rm R}^{\rm a}p_{\rm RD}^{\rm det},\tag{3}$$

where $p_{\rm R}^{\rm a}$ is the probability that node R is active, i.e., it has at least one data packet and *K* energy units backlogged at the beginning of the second subslot.

Likewise, the average transmission delay is defined as the average time from the moment a packet becomes head-of-line (HoL) at node S until it is successfully detected at node D. Consider node R, i.e., its data queue and its server, as well as the server of node S as a queueing system shown in Fig. 1(b). The reason we consider these components as a queueing system is that the source packets receive their service either from the server of node S, in the case of being successfully transmitted via the direct link, or from both servers of node S and node R, in the case of being relayed. In the latter case, the packets may wait at $\rm Q_d$ and then, be served by the server of the relay. Hence, the average transmission delay of the source packets, denoted by τ, corresponds to the average waiting time of the packets at the aforementioned queueing system, and can be derived according to the Little's Law [31] as

$$\tau=\frac{\overline{q}_{\rm d}+1}{\lambda_{\rm S}},\tag{4}$$

where $\overline{q}_{\rm d}$ is the average number of data packets at node R. Here, the average number of packets in the aforementioned queueing system is equal to $\overline{q}_{\rm d}+1$, because node S is infinitely backlogged and thus, always has one packet in its server. Moreover, the arrival rate to the queueing system is equal to the throughput $\lambda_{\rm S}$ when the system is stable, i.e., the arrival and departure rates are equal.

It is worth to mention that the policies achieving the maximum throughput and minimum delay, respectively, are not necessarily the same. To see this, note that the source throughput is maximized when $p_{\rm R}^{\rm a}$ is maximized. That is, the probability that the data buffer $\rm Q_d$ is backlogged and the energy buffer $\rm Q_e$ has at least *K* energy units is maximized. However, this does not necessarily minimize $\overline{q}_{\rm d}$, the average number of data packets backlogged at node R. The throughput-optimal and delay-optimal TS policies are derived in subsequent sections.

III. Optimal Static Policies

Under a static TS policy, node R selects an operation mode with a fixed probability irrespective of the states of the data and energy buffers. That is, $\alpha_s=\alpha\ \forall s$, where α_s was defined in Section II-B. In this section, we derive the optimal α that maximizes the throughput and minimizes the average transmission delay, respectively.

A. Throughput-Optimal Static Policy

We first give the definition of the boundary of stability as follows.

Definition 1: According to [32], $\rm Q_d$ is stable if and only if its arrival rate is equal to its departure rate. Then, when applying a static (dynamic) policy at node R, the boundary of stability of $\rm Q_d$ is defined as the data arrival rate such that if it is increased by any $\epsilon>0$, no static (dynamic) policy can be found to stabilize $\rm Q_d$, i.e., at any static (dynamic) policy, the arrival rate at $\rm Q_d$ becomes greater than the departure rate.

By the following lemmas, we prove that the throughput-optimal static policy is the policy that keeps $\rm Q_d$, the data buffer at node R, at the boundary of stability.

Lemma 1: $p_{\rm R}^{\rm a}$ is an increasing function of α when $\rm Q_d$ is stable.

Proof: Define $\lambda_{\rm id}$ and $\lambda_{\rm od}$ to be the arrival and departure rates at $\rm Q_d$, respectively. Then

$$\lambda_{\rm id}=\alpha\Big(1-p_{\rm SD}^{\rm det}\Big)p_{\rm SR}^{\rm det},\qquad \lambda_{\rm od}=p_{\rm R}^{\rm a}p_{\rm RD}^{\rm det}.\tag{5}$$

When $\rm Q_d$ is stable, its arrival rate is equal to its departure rate, i.e., $\lambda_{\rm id}=\lambda_{\rm od}$. Thus, from (5), $p_{\rm R}^{\rm a}$ is derived as

$$p_{\rm R}^{\rm a}=\frac{\alpha\Big(1-p_{\rm SD}^{\rm det}\Big)p_{\rm SR}^{\rm det}}{p_{\rm RD}^{\rm det}},\tag{6}$$

which is obviously an increasing function of α. ■

Lemma 2: The blocking probability of the energy buffer $\rm Q_e$, denoted by $p_{\rm b}$, is zero if and only if $\rm Q_d$ is at the boundary of stability or unstable.

Proof: We first prove the "if" part. When $\rm Q_d$ is at the boundary of stability or unstable, there is always a packet in $\rm Q_d$. In fact, the arrival process at $\rm Q_d$ is Bernoulli since in static policy $\alpha_s=\alpha\ \forall s$ and thus, similar to a Geo/G/1 queue, $\rm Q_d$ will be always backlogged at the boundary of stability. Now, to show that the blocking probability is zero, we first show that $q_{\rm e}\le N-K$ when $\rm Q_d$ is always backlogged. To this end, let

$q'_{\rm e}$ denote the number of energy units at the beginning of the second subslot. If $q'_{\rm e} < K$, no transmission happens and thus, $q_{\rm e} = q'_{\rm e} < K$. Moreover, due to our assumption $N \geq 2K$ (see Section II-A), we have $K \leq N - K$. Thus, we conclude that $q_{\rm e} < K \leq N-K$. On the other hand, if $K \leq q'_{\rm e} \leq N$, a packet is transmitted since $\mathrm{Q_d}$ is backlogged and enough energy is available. Therefore, we have $q_{\rm e} = q'_e - K \leq N - K$. This completes the proof of $q_{\rm e} \leq N - K$. Now, observe that by our assumption $b_{\max} \leq K$, at most K energy units are harvested in the first subslot and thus, no blocking occurs.

Next, we prove the "only if" part by contradiction. Let $\lambda_{\rm ie}$ denote the energy arrival rate after blocking at the energy buffer. Likewise, define $\lambda_{\rm oe}$ to be the departure rate from $\mathrm{Q_e}$. Then, we have

$$\lambda_{\rm oe} = \lambda_{\rm ie} = (1-\alpha)E(\Gamma)(1-p_{\rm b}), \tag{7}$$

where $E(\Gamma)$ is the average of Γ. Now, suppose $p_{\rm b} = 0$ when $\mathrm{Q_d}$ is stable. Then, (7) becomes $\lambda_{\rm oe} = \lambda_{\rm ie} = (1-\alpha)E(\Gamma)$, implying that $\lambda_{\rm oe}$ and $\lambda_{\rm ie}$ are decreasing functions of α. Note that a transmission attempt of a data packet at node R consumes K energy units at the same time. Thus, a decrease in $\lambda_{\rm oe}$ is equivalent to a decrease in the data departure rate, i.e., $\lambda_{\rm od}$. On the other hand, due to stability, the arrival and departure rates at $\mathrm{Q_d}$ are equal, i.e., $\lambda_{\rm od} = \lambda_{\rm id} = \alpha(1-p_{\rm SD}^{\rm det})p_{\rm SR}^{\rm det}$ from (5), implying that $\lambda_{\rm od}$ increases with α. This is a contradiction, except when the equation $\lambda_{\rm od} = \lambda_{\rm id}$ does not hold by increasing α, i.e., $\mathrm{Q_d}$ is at the boundary of stability, according to Definition 1, or when the data arrival and departure rates are not equal even before increasing α, i.e., $\mathrm{Q_d}$ is unstable. ■

Lemma 3: $p_{\rm R}^{\rm a}$ is a decreasing function of α when $\mathrm{Q_d}$ is unstable.

Proof: A packet departure from $\mathrm{Q_d}$ consumes $\frac{K}{p_{\rm RD}^{\rm det}}$ energy units on average, since each packet is transmitted $\frac{1}{p_{\rm RD}^{\rm det}}$ times on average. Thus, the energy departure rate, $\lambda_{\rm oe}$, can be written as a function of data departure rate, $\lambda_{\rm od}$, as follows:

$$\lambda_{\rm oe} = \lambda_{\rm od}\frac{K}{p_{\rm RD}^{\rm det}}. \tag{8}$$

From (5), (7), and (8), we have

$$\alpha = 1 - \frac{p_{\rm R}^{\rm a}K}{E(\Gamma)(1-p_{\rm b})}. \tag{9}$$

According to Lemma 2, $p_{\rm b} = 0$ when $\mathrm{Q_d}$ is unstable. Thus, it can be seen from (9) that $p_{\rm R}^{\rm a}$ is decreasing with α when $\mathrm{Q_d}$ is unstable. ■

In the following proposition, we derive the optimum α that maximizes the throughput, represented by $\alpha^{\rm T}$, and show that it keeps $\mathrm{Q_d}$ at the boundary of stability.

Proposition 1: The throughput-optimal static policy keeps $\mathrm{Q_d}$ at the boundary of stability. In addition, $\alpha^{\rm T}$ is given by

$$\alpha^{\rm T} = \frac{E(\Gamma)p_{\rm RD}^{\rm det}}{E(\Gamma)p_{\rm RD}^{\rm det} + K\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}}. \tag{10}$$

Proof: Lemmas 1 and 3 show that $p_{\rm R}^{\rm a}$ increases with α when $\mathrm{Q_d}$ is stable and decreases with α when $\mathrm{Q_d}$ is unstable. Thus, the maximum of $p_{\rm R}^{\rm a}$, which leads to the maximum throughput according to (3), is attained by carefully setting α so that $\mathrm{Q_d}$ is at the boundary of stability. Now, in order to derive α, which keeps $\mathrm{Q_d}$ at the boundary of stability, i.e., $\alpha^{\rm T}$, we use (6) and (9). From these equations, α is written in terms of $p_{\rm b}$ as in the following

$$\alpha = \frac{E(\Gamma)p_{\rm RD}^{\rm det}(1-p_{\rm b})}{E(\Gamma)p_{\rm RD}^{\rm det}(1-p_{\rm b}) + K\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det}}. \tag{11}$$

The above equation holds when $\mathrm{Q_d}$ is stable including the boundary of stability. Indeed, according to Lemma 2, $p_{\rm b} = 0$ at the boundary of stability. Therefore, (10) is obtained. ■

B. Delay-Optimal Static Policy

It is interesting to note that the relay cooperation may degrade the average transmission delay in some cases. For example, this may occur when $p_{\rm RD}^{\rm det} < p_{\rm SD}^{\rm det}$ since in this case, the packets experience a worse channel when being relayed. In this subsection, we first introduce the necessary and sufficient condition for the non-cooperation policy to be delay-optimal. Here, the non-cooperation policy refers to the one that sets $\alpha = 0$. That is, the relay never decodes a packet, and therefore never helps to relay a packet. Then, we derive the optimal α that minimizes the average transmission delay, denoted by $\alpha^{\rm D}$, when the non-cooperation policy is not optimal. Note that the average delay, τ, derived in (4), and the average queue length, $\overline{q}_{\rm d}$, are functions of the DD probability, α, i.e., we have $\tau = \tau(\alpha)$ and $\overline{q}_{\rm d} = \overline{q}_{\rm d}(\alpha)$. However, we drop α whenever it is clear from the context.

Lemma 4: The non-cooperation policy is delay-optimal, i.e., $\alpha^{\rm D} = 0$, if and only if

$$\overline{D}_{\rm R}(\alpha) \geq \frac{1}{p_{\rm SD}^{\rm det}}, \quad \forall \alpha \in \Big(0, \alpha^{\rm T}\Big]. \tag{12}$$

where $\overline{D}_{\rm R}(\alpha)$ denotes the average system delay (comprised of queueing delay and transmission time) at node R when the DD probability is α. Note that the transmission time of a packet at node R starts from the moment it becomes the head-of-line packet in the data buffer ($\mathrm{Q_d}$) and terminates when it is successfully transmitted. Moreover, $\alpha^{\rm T}$ is the throughput-optimal DD probability, derived in (10).

Proof: According to Little's law, we have $\overline{q}_{\rm d}(\alpha) = \overline{D}_{\rm R}(\alpha)\lambda_{\rm id}$. Thus, from (4), τ is written as $\tau(\alpha) = \frac{\overline{D}_{\rm R}(\alpha)\lambda_{\rm id}+1}{\lambda_{\rm S}}$. Then, substituting (5) for $\lambda_{\rm id}$, we have

$$\tau(\alpha) = \frac{\alpha\Big(1-p_{\rm SD}^{\rm det}\Big)p_{\rm SR}^{\rm det}\overline{D}_{\rm R}(\alpha)+1}{\lambda_{\rm S}}. \tag{13}$$

Moreover, since $\mathrm{Q_d}$ is stable (otherwise, $\overline{q}_{\rm d}$ and thus, τ is infinite), $\lambda_{\rm id} = \lambda_{\rm od}$. Thus, from (3), (5), $\lambda_{\rm S}$ can be written as

$$\lambda_{\rm S} = \alpha\Big(1-p_{\rm SD}^{\rm det}\Big)p_{\rm SR}^{\rm det} + p_{\rm SD}^{\rm det}. \tag{14}$$

Substituting (14) for the denominator of (13), we have

$$\tau = \frac{\alpha\Big(1-p_{\rm SD}^{\rm det}\Big)p_{\rm SR}^{\rm det}\overline{D}_{\rm R}(\alpha)+1}{\alpha\big(1-p_{\rm SD}^{\rm det}\big)p_{\rm SR}^{\rm det} + p_{\rm SD}^{\rm det}}. \tag{15}$$

The non-cooperation policy, i.e., $\alpha = 0$, is delay-optimal if and only if $\tau(\alpha) \geq \tau(0)$ for any stabilizing $\alpha > 0$, i.e.,

$\alpha \in (0, \alpha^{\mathrm{T}}]$. Using (15) and regarding that $\tau(0) = \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$, this translates to

$$\frac{\alpha\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}\overline{D}_{\mathrm{R}}(\alpha)+1}{\alpha\big(1-p_{\mathrm{SD}}^{\mathrm{det}}\big)p_{\mathrm{SR}}^{\mathrm{det}}+p_{\mathrm{SD}}^{\mathrm{det}}} \geq \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}, \quad \forall \alpha \in \Big(0, \alpha^{\mathrm{T}}\Big]. \quad (16)$$

By some mathematical manipulation, the above equation holds if and only if $\overline{D}_{\mathrm{R}}(\alpha) \geq \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$ for any $\alpha \in (0, \alpha^{\mathrm{T}}]$. ∎

Remark 1: Suppose that $\overline{D}_{\mathrm{R}}$, the average system delay at node R, is an increasing function of α, or equivalently the data arrival rate at $\mathrm{Q_d}$, i.e., λ_{id} in (5). This is a valid assumption in a queue with general data arrival and service time processes, given that the average service time of the packets does not decrease with the arrival rate [31]. This condition is satisfied in our scenario since by increasing the arrival rate, the EH rate at node R decreases, implying that the average service time does not decrease. Then, the inequality in (12) holds if and only if $\overline{D}_{\mathrm{R}}(\alpha) \geq \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$ when $\alpha \to 0$ since $\overline{D}_{\mathrm{R}}$ is increasing with α. Thus, the inequality

$$\overline{D}_{\mathrm{R}}(\alpha \to 0) \geq \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}, \quad (17)$$

is a necessary and sufficient condition for the non-cooperation policy to be delay-optimal.

If the condition in (12) does not hold, then the delay-optimal DD probability is positive ($\alpha^{\mathrm{D}} > 0$). In particular

$$\alpha^{\mathrm{D}} = \underset{\alpha \in (0, \alpha^{\mathrm{T}}]}{\operatorname{argmin}} \tau(\alpha) = \frac{\alpha\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}\overline{D}_{\mathrm{R}}+1}{\alpha\big(1-p_{\mathrm{SD}}^{\mathrm{det}}\big)p_{\mathrm{SR}}^{\mathrm{det}}+p_{\mathrm{SD}}^{\mathrm{det}}}, \quad (18)$$

where τ is written as in (15). Note that according to Proposition 1, $\mathrm{Q_d}$ is unstable and τ becomes infinite when $\alpha > \alpha^{\mathrm{T}}$. To solve (18), we first prove in Lemma 5 that $\tau(\alpha)$ is convex.

Lemma 5: $\tau(\alpha)$ is convex for the range of α that yields $\overline{D}_{\mathrm{R}}(\alpha) < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$, i.e., when cooperation outperforms non-cooperation (see Lemma 4).

Proof: From (15), $\frac{d^2\tau}{d\alpha^2}$ is computed as a function of $\overline{D}_R$ and its derivatives as

$$\frac{d^2\tau}{d\alpha^2} = \frac{d^2\overline{D}_{\mathrm{R}}}{d\alpha^2}\frac{\alpha\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}}{A} + 2\frac{d\overline{D}_{\mathrm{R}}}{d\alpha}\frac{p_{\mathrm{SD}}^{\mathrm{det}}\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}}{A^2} + 2\frac{\Big(\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}\Big)^2\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\overline{D}_{\mathrm{R}}\Big)}{A^3},$$
$$A = \alpha\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}} + p_{\mathrm{SD}}^{\mathrm{det}}. \quad (19)$$

Since $\overline{D}_{\mathrm{R}}$ is increasing in terms of α (see Remark 1), we have $\frac{d\overline{D}_{\mathrm{R}}}{d\alpha} > 0$. Moreover, $\mathrm{Q_d}$ acts like a Geo/G/1 queue since the data arrival process at $\mathrm{Q_d}$ is Bernoulli and the service time is general. Thus, its average delay, $\overline{D}_{\mathrm{R}}$, is a convex function of the arrival rate (refer to P-K formula in [31]), i.e., $\frac{d^2\overline{D}_{\mathrm{R}}}{d\alpha^2} > 0$. Thus, the first and second additive terms in (19) are positive. The last one is also positive due to the assumption that $\overline{D}_{\mathrm{R}} < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$. ∎

Although τ is a convex function of α, it is difficult to derive the optimal point, α^{D}, explicitly, since it requires the derivation of $\overline{D}_{\mathrm{R}}$ and consequently $\overline{q}_{\mathrm{d}}$ in terms of α, which is not possible. In this regard, we use an off-the-shelf numerical method for convex optimization, i.e., golden section search [29], as detailed in Algorithm 1, to solve (18). This algorithm shrinks the interval which contains the optimal point, in each iteration, by evaluating τ at the interval extreme points. In order to evaluate τ at a given α, we derive the corresponding $\overline{q}_{\mathrm{d}}$, (see (4)), using the related QBD process. Let $\boldsymbol{\pi}_l$ denote the stationary probability vector of the l-th level of the QBD process, i.e., the states (l, q_{e}), where $q_{\mathrm{e}} \in \{0, 1, \ldots, N\}$. Then, we have $\overline{q}_{\mathrm{d}} = \sum_{l=0}^{\infty} l\boldsymbol{\pi}_l\mathbf{1}$, where $\mathbf{1}$ is the all-one vector. On the other hand, by setting $\alpha_s = \alpha\ \forall s$, the related QBD process becomes homogeneous and thus, we have $\boldsymbol{\pi}_l = \boldsymbol{\pi}_0\mathbf{R}^l$, according to [28], where $\mathbf{R}$[1] is a matrix related to the QBD process. Therefore, $\overline{q}_{\mathrm{d}}$ is written as

$$\overline{q}_{\mathrm{d}} = \sum_{l=0}^{\infty} l\boldsymbol{\pi}_l\mathbf{1} = \boldsymbol{\pi}_0\sum_{l=0}^{\infty} l\mathbf{R}^l\mathbf{1} = \boldsymbol{\pi}_0\mathbf{R}(\mathbf{I}-\mathbf{R})^{-2}\mathbf{1}. \quad (20)$$

Note that $\boldsymbol{\pi}_l$ and $\mathbf{R}$ are computed using some iterative algorithms in [28, Th. 8.3.1], which employ the transition probabilities of the QBD process. Since these transition probabilities, derived in Section II-B, are functions of α, we conclude that $\overline{q}_{\mathrm{d}}$ is also a function of α.

Algorithm 1 Delay-Optimal Static Policy

```
if (12) holds then
    No-cooperation is delay-optimal
else
    a = 0 and b = 1
    while ( (b−a)/b ≥ ϵ ) do
        a_1 = a + 0.382(b − a) and b_1 = b − 0.382(b − a)
        Derive τ(a_1) and τ(b_1) from QBD process
        If τ(a_1) < τ(b_1) then b = b_1. Otherwise, a = a_1.
    end while
end if
```

IV. Optimal Dynamic Policies

In this section, we derive the throughput-optimal and delay-optimal dynamic policies. In a dynamic policy, the mode selection decisions are made at the beginning of each slot based on the states of the data and energy buffers, i.e., s. Unlike the static policy, α_s is not necessarily the same for different states. Thus, the underlying QBD is not homogeneous in general.

A. Throughput-Optimal Dynamic Policy

Define $\{\alpha_s\}_{s \in S}$ to be a dynamic policy, where S is the set of all possible states at node R. Let $\overline{\alpha_s} = \sum_{s \in S} p_s\alpha_s$ denote the average DD probability corresponding to the dynamic policy α_s, where p_s is the probability of node R being in state s at the beginning of a slot under policy α_s. Note that Lemmas 1-3 still hold in the dynamic case by replacing α with $\overline{\alpha_s}$, because all equations in these lemmas are based on average arrival and departure rates. The only part that needs

[1] $\mathbf{R}_{ij}$ is the expected number of visits to the state $(n+1, j)$, before a return to level n or previous levels given that the process starts from state (n, i).

revision, is the proof of "if" part in Lemma 2. In this proof, we have assumed that when $\mathrm{Q_d}$ is at the boundary of stability or unstable, it is always backlogged. Then, we have proved $p_\mathrm{b}=0$. However, in the dynamic case, the data and energy arrivals can be controlled at each state. Thus, $\mathrm{Q_d}$ is not necessarily backlogged at the boundary of stability. In other words, it can be empty in some slots at the boundary of stability (as will be shown in Proposition 2). Thus, in Lemma 6, we reprove the "if" part of Lemma 2 for the dynamic case. Consequently, based on Lemmas 1-3, we can conclude from Proposition 1 that the throughput-optimal dynamic policy also keeps $\mathrm{Q_d}$ at the boundary of stability. Then, $\overline{\alpha_s}$ corresponding to the throughput-optimal dynamic policy is equal to α^T in (10).

Lemma 6: If $\mathrm{Q_d}$ is at the boundary of stability, then $p_\mathrm{b}=0$.

Proof: Note that we derived (11) in Proposition 1 from the balance between the arrival and departure rates at $\mathrm{Q_d}$ and $\mathrm{Q_e}$, assuming the stability of $\mathrm{Q_d}$ (i.e., from (5)-(9)). Thus, (11) is applicable for any stabilizing dynamic policies α_s. That is, we have

$$\overline{\alpha_s}=\frac{E(\Gamma)p_\mathrm{RD}^\mathrm{det}(1-p_\mathrm{b})}{E(\Gamma)p_\mathrm{RD}^\mathrm{det}(1-p_\mathrm{b})+K\big(1-p_\mathrm{SD}^\mathrm{det}\big)p_\mathrm{SR}^\mathrm{det}} \tag{21}$$

where p_b is the blocking probability under policy α_s. Now, we prove the lemma by contradiction. Suppose that there exists a dynamic policy α'_s that keeps $\mathrm{Q_d}$ at the boundary of stability with blocking probability $p'_\mathrm{b}>0$. Then, $\overline{\alpha'_s}<\overline{\alpha_s}|_{p_\mathrm{b}=0}$ since according to (21), $\overline{\alpha_s}$ is decreasing with p_b. On the other hand, $\overline{\alpha_s}|_{p_\mathrm{b}=0}=\alpha^\mathrm{T}$, according to Proposition 1. Thus, we have $\overline{\alpha'_s}<\alpha^\mathrm{T}$. Now, note that $\overline{\alpha'_s}$ corresponds to the boundary of stability, when a dynamic policy is applied at node R. Thus, regarding the definition of the boundary of stability in Section III-A, for the average DD probabilities greater than $\overline{\alpha'_s}$, equivalent to the greater data arrival rates at node R, we cannot find a dynamic policy which stabilizes $\mathrm{Q_d}$. However, α^T, where $\alpha^\mathrm{T}>\overline{\alpha'_s}$, corresponds to the throughput-optimal static policy, according to Proposition 1, i.e., a state-independent dynamic policy that keeps the data buffer at the boundary of stability. This is a contradiction. ∎

Next, we prove that the throughput-optimal dynamic policy is a simple threshold-based policy.

Proposition 2: The following threshold-based policy is a throughput-optimal policy

$$\alpha_s^\mathrm{T}=\begin{cases}1; & s\in\{(0,q_\mathrm{e})|\ q_\mathrm{e}\geq e_\mathrm{th}\},\\ 0; & s\in\{(0,q_\mathrm{e})|\ q_\mathrm{e}<e_\mathrm{th}\},\\ 0; & s\in\{(1,q_\mathrm{e})|\ 0\leq q_\mathrm{e}\leq N\}.\end{cases} \tag{22}$$

for all choices of $e_\mathrm{th}\leq N-b_\mathrm{max}+1$. Consequently, $\overline{\alpha_s^\mathrm{T}}=\sum_{s\in S}p_s\alpha_s^\mathrm{T}=\alpha^\mathrm{T}$, where p_s is the probability of state s under policy α_s^T.

Proof: The policy described in (22) allows at most one data packet to be backlogged at node R at a time since for $s=(1,q_\mathrm{e})$, we have $\alpha_s^\mathrm{T}=0$, i.e., node R remains in the EH mode after a packet is stored. Thus, $\mathrm{Q_d}$ is always stable. Now, if $p_\mathrm{b}=0$ under α_s^T in addition, then α_s^T keeps $\mathrm{Q_d}$ at the boundary of stability according to Lemma 2.

To show p_b is indeed 0 under α_s^T, we show that it is zero in different states as follows.

$s=(0,q_\mathrm{e})$, $q_\mathrm{e}<e_\mathrm{th}$: Here, node R is in the EH mode. Also, from $q_\mathrm{e}<e_\mathrm{th}$ and $e_\mathrm{th}\leq N-b_\mathrm{max}+1$, we have $q_\mathrm{e}\leq N-b_\mathrm{max}$. No blocking occurs in this case since at most b_max energy units can be harvested in the EH mode.

$s=(0,q_\mathrm{e})$, $q_\mathrm{e}\geq e_\mathrm{th}$: Node R is in the DD mode. No energy is harvested and thus, no energy units are blocked. Therefore, the blocking probability is zero in this case.

$s=(1,q_\mathrm{e})$: Node R is in the EH mode. Moreover, when $s=(1,q_\mathrm{e})$ at the beginning of a slot, the data buffer $\mathrm{Q_d}$ is definitely backlogged in the second subslot of the previous slot since no packet enters $\mathrm{Q_d}$ in the second subslot. Using this fact, similar to the proof of the "if" part in Lemma 2, we have $q_\mathrm{e}\leq N-K$. Then, the blocking probability is zero at state $(1,q_\mathrm{e})$ since at most $b_\mathrm{max}\leq K$ energy units can be harvested in the EH mode.

Therefore, we have proved that under policy α_s^T, $\mathrm{Q_d}$ is stable and $p_\mathrm{b}=0$. According to Lemma 2, α_s^T keeps $\mathrm{Q_d}$ at the boundary of stability and thus, is throughput-optimal, i.e., $\overline{\alpha_s^\mathrm{T}}=\alpha^\mathrm{T}$. ∎

Remark 2: It is worth noting that the throughput-optimal policy described in (22) is not unique. For example, when $b_\mathrm{max}=1$, any e_th leads to the optimal throughput. In essence, adopting different e_th in the range $[0,N-b_\mathrm{max}+1]$ leads to the same optimal throughput, albeit different idle time intervals at node R, i.e., the intervals node R waits till q_e exceeds e_th.

B. Delay-Optimal Dynamic Policy

According to (4), the average transmission delay in a dynamic policy is $\tau(\alpha_s)=\frac{\overline{q}_\mathrm{d}(\alpha_s)+1}{\lambda_\mathrm{S}(\alpha_s)}$, where α_s denotes the applied policy at node R. Moreover, using (13) and (14) in the case of dynamic policy, $\tau(\alpha_s)$ can be written similar to (15) as

$$\tau(\alpha_s)=\frac{\overline{\alpha_s}\Big(1-p_\mathrm{SD}^\mathrm{det}\Big)p_\mathrm{SR}^\mathrm{det}\overline{D}_\mathrm{R}(\alpha_s)+1}{\overline{\alpha_s}\big(1-p_\mathrm{SD}^\mathrm{det}\big)p_\mathrm{SR}^\mathrm{det}+p_\mathrm{SD}^\mathrm{det}}. \tag{23}$$

Correspondingly, Lemma 4 and consequently Remark 1 holds in the dynamic case, i.e., non-cooperation is delay-optimal if and only if $\overline{D}_\mathrm{R}(\overline{\alpha_s}\to 0)\geq\frac{1}{p_\mathrm{SD}^\mathrm{det}}$. In Proposition 3, we derive the delay-optimal dynamic policy, where b_max is set to 1 for the time being.

Proposition 3: If $b_\mathrm{max}=1$ and $\overline{D}_\mathrm{R}(\overline{\alpha_s}\to 0)<\frac{1}{p_\mathrm{SD}^\mathrm{det}}$, i.e., when cooperation is preferred to non-cooperation (see Remark 1), the delay-optimal policy α_s^D is given by

$$\alpha_s^\mathrm{D}=\begin{cases}1; & s=(0,N)\\ 0; & \text{otherwise.}\end{cases} \tag{24}$$

Proof: Consider the following policy

$$\hat{\alpha}_s(\beta)=\begin{cases}\beta; & s=(0,N)\\ 0; & \text{otherwise.}\end{cases} \tag{25}$$

where $0\leq\beta\leq 1$. In Appendix B, we have proved that $\overline{\hat{\alpha}_s(\beta)}$ is a strictly increasing continuous function of β. On the other hand, for the extreme values $\beta=0$ and $\beta=1$, we have $\overline{\hat{\alpha}_s(0)}=0$ and $\overline{\hat{\alpha}_s(1)}=\alpha^\mathrm{T}$, respectively, where the second equality is written since when $b_\mathrm{max}=1$, $\hat{\alpha}_s(\beta)$ with $\beta=1$ is throughput-optimal, according to Proposition 2. Thus, any

value of $\overline{\hat{\alpha}_s(\beta)} = \overline{\alpha} \in [0, \alpha^{\mathrm{T}}]$ is achieved by a unique $\beta \in [0,1]$. In the following, we denote this unique β by $\beta = \beta(\overline{\alpha})$.

Now, we prove that among all policies with the same average DD probability $\overline{\alpha}$, $\hat{\alpha}_s(\beta)$ with $\beta = \beta(\overline{\alpha})$ minimizes τ. That is, $\hat{\alpha}_s(\beta(\overline{\alpha})) = \mathrm{argmin}\{\tau(\alpha_s) \,|\, \overline{\alpha_s} = \overline{\alpha}\}$, where $\tau(\alpha_s)$ is derived in (23). Note that for the policies with the same average DD probability $\overline{\alpha_s} = \overline{\alpha}$, the minimization of $\tau(\alpha_s)$ in (23) is equivalent to the minimization of $\overline{D_{\mathrm{R}}}(\alpha_s)$. $\overline{D_{\mathrm{R}}}$ is the average queueing delay plus the average transmission time of the packets at the relay. On the one hand, the average queueing delay is zero under $\hat{\alpha}_s(\beta)$ since the packets are received only when $q_{\mathrm{d}} = 0$. On the other hand, all packets have the same average transmission time under $\hat{\alpha}_s(\beta)$ since they observe full energy buffer upon their arrival and the fixed EH mode during their transmission time. Thus, $\overline{D_{\mathrm{R}}} = t_N$, where t_N denotes the average transmission time of a packet under policy $\hat{\alpha}_s(\beta)$. Now, observe that t_N is the minimum transmission time that can be experienced by a packet at node R. This is due to the fact that by observing full energy buffer upon the arrival and fixed EH mode during the transmission time, a packet benefits from the maximum possible energy for its initial transmission as well as probable retransmissions. Therefore, $\overline{D_{\mathrm{R}}}$ and consequently τ is minimized under $\hat{\alpha}_s(\beta)$ with $\beta = \beta(\overline{\alpha})$ given a fixed average DD probability $\overline{\alpha}$.

So far, we have determined the structure of the delay-optimal policy at the given $\overline{\alpha_s} = \overline{\alpha}$, which is $\hat{\alpha}_s(\beta)$ with $\beta = \beta(\overline{\alpha})$. Now, it remains to find the optimal $\overline{\alpha}$ or equivalently the optimal $\beta = \beta(\overline{\alpha})$, which minimizes $\tau(\hat{\alpha}_s(\beta))$. In the following, we show that $\tau(\hat{\alpha}_s(\beta))$ is a decreasing function of β. Then, we conclude that the optimal β is 1 and consequently, the optimal $\overline{\alpha}$ is α^{T}. Thus, the proposition is proved.

Replacing $\overline{D}_{\mathrm{R}} = t_N$ in (23), we have

$$\tau(\hat{\alpha}_s(\beta)) = \frac{\overline{\alpha}\Big(1 - p_{\mathrm{SD}}^{\mathrm{det}}\Big) p_{\mathrm{SR}}^{\mathrm{det}} t_N + 1}{\overline{\alpha}\big(1 - p_{\mathrm{SD}}^{\mathrm{det}}\big) p_{\mathrm{SR}}^{\mathrm{det}} + p_{\mathrm{SD}}^{\mathrm{det}}}, \tag{26}$$

and then

$$\frac{d\tau}{d\beta} = \frac{\Big(1 - p_{\mathrm{SD}}^{\mathrm{det}}\Big) p_{\mathrm{SR}}^{\mathrm{det}} \Big(p_{\mathrm{SD}}^{\mathrm{det}} t_N - 1\Big)}{\big(\overline{\alpha}\big(1 - p_{\mathrm{SD}}^{\mathrm{det}}\big) p_{\mathrm{SR}}^{\mathrm{det}} + p_{\mathrm{SD}}^{\mathrm{det}}\big)^2} \; \frac{d\overline{\alpha}}{d\beta}. \tag{27}$$

As mentioned before, it is proven in Appendix VI that $d\overline{\alpha}/d\beta > 0$ under $\hat{\alpha}_s(\beta)$. Then, to verify that $d\tau/d\beta < 0$, we need to show that $t_N < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$.

When α or $\overline{\alpha_s}$ tends to 0, the average interarrival time of the packets at node R, which is $1/\lambda_{\mathrm{id}}$, goes to infinity almost surely, according to (5). In this case, the data buffer becomes empty and the energy buffer becomes fully backlogged before arrival of a new packet at $\mathrm{Q_d}$. Thus, the packet is transmitted immediately upon its arrival at node R. Moreover, Node R remains in the EH mode during the packet transmission time since $\overline{\alpha_s} \to 0$. Thus, $\overline{D}_R(\overline{\alpha_s} \to 0) = t_N$. Then, by our assumption $\overline{D_{\mathrm{R}}}(\overline{\alpha} \to 0) < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$, we have $t_N < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$. The proof is completed here. It is worth noting that regarding (22) and (24), we deduce that for $b_{\max} = 1$, the delay-optimal dynamic policy is also throughput-optimal. ■

The above proposition describes the delay-optimal policy if $b_{\max} = 1$, which is a threshold-based policy. The structure of the delay-optimal policy when $b_{\max} > 1$ is more complicated. Nevertheless, by giving intuitions, we derive some properties of the delay-optimal policy. Similar to the case $b_{\max} = 1$, for a given $\overline{\alpha}$, the average delay is minimized by minimizing $\overline{D_{\mathrm{R}}}$, according to (23), i.e., the average queueing delay plus the average transmission time of the packet. The average transmission time of a packet will be less if no other packet is received during its transmission. Because, in this case all slots are dedicated to the EH mode, and thus, the maximum possible energy is harvested for transmission of the packet. Moreover, no packet is queued at node R. Thus, the queueing delay is zero. Therefore, it is deduced that for a given $\overline{\alpha_s}$, a delay-optimal policy should receive a packet after successfully transmitting the current packet, i.e., it keeps at most one packet at node R. Consequently, the delay-optimal policy, i.e., the policy corresponding to the optimal $\overline{\alpha_s}$, has the same property. Now we focus on the threshold-based policies with the structure as in (22) since these policies keep at most one packet at node R.

Our numerical results show that in a policy with a structure similar to (22) and for a typical EH profile (e.g., exponential distribution), $\overline{\alpha_s}$ or equivalently the throughput, λ_{S}, increases by decreasing e_{th} from N to $N - b_{\max} + 1$. This is intuitively due to the fact that by decreasing e_{th}, node R will be in the DD mode in more higher energy levels and thus, less energy units are blocked, i.e., p_{b} decreases. Consequently, by decreasing p_{b}, $\overline{\alpha_s}$ increases according to (21). On the other hand, by increasing $\overline{\alpha_s}$, the arrival rate to $\mathrm{Q_d}$ and as a result the average delay at node R, $\overline{D_{\mathrm{R}}}$, increases, according to Remark 1. Since both the nominator and the denominator in (23) increase, there is a chance that τ decreases as e_{th} is decreased from N to $N - b_{\max} + 1$. However, such a phenomenon does not happen when e_{th} is decreased less than $N - b_{\max} + 1$. Because in this case, $\overline{\alpha_s}$ remains constant and equal to α^{T}, according to Proposition 2. However, $\overline{D_{\mathrm{R}}}$ in (23) does not decrease since some of the packets are transmitted at lower initial energy levels. Therefore, τ does not decrease, implying that $e_{\mathrm{th}} < N - b_{\max} + 1$ cannot be a choice for delay-optimal policy. Based on this observations, we guess the structure of the delay-optimal policy for the case $b_{\max} > 1$ in the following conjecture.

Conjecture 1: If the cooperation is helpful, i.e., $\overline{D}_{\mathrm{R}}(\overline{\alpha_s} \to 0) < \frac{1}{p_{\mathrm{SD}}^{\mathrm{det}}}$, the delay-optimal policy, α_s^{D}, is given by

$$\alpha_s^{\mathrm{D}} = \begin{cases} 1; & s \in \{(0, q_{\mathrm{e}}) |\ q_{\mathrm{e}} > e_{\mathrm{th}}\}, \\ \beta; & s \in \{(0, q_{\mathrm{e}}) |\ q_{\mathrm{e}} = e_{\mathrm{th}}\}, \\ 0; & \text{Otherwise}, \end{cases} \tag{28}$$

for some $\beta \in [0,1]$ and $e_{\mathrm{th}} \in \{N - b_{\max} + 1, \ldots, N\}$.

We now propose an algorithm to find the optimal values of β and e_{th}. Consider the Markov chain related to α_s^{D} in (28). The corresponding transition probabilities are written as in (1) and (2) by replacing α_s with α_s^{D}. Then, $\overline{\alpha_s^{\mathrm{D}}}$ and $\overline{q_{\mathrm{d}}}$ are derived

TABLE II
THROUGHPUT-OPTIMAL AND DELAY-OPTIMAL STATIC AND DYNAMIC POLICIES

Policy	Throughput-optimal	Delay-optimal
Static	$\alpha^{\mathrm{T}} = \frac{E(\Gamma)p_{\mathrm{RD}}^{\mathrm{det}}}{E(\Gamma)p_{\mathrm{RD}}^{\mathrm{det}}+K(1-p_{\mathrm{SD}}^{\mathrm{det}})p_{\mathrm{SR}}^{\mathrm{det}}}$	α^{D} is derived in Algorithm 1
Dynamic	$\alpha_s^{\mathrm{T}} = \begin{cases} 1 \ ; & s \in \{(0,q_{\mathrm{e}})\mid q_{\mathrm{e}} > e_{\mathrm{th}}\}, \\ 0 \ ; & s \in \{(0,q_{\mathrm{e}})\mid q_{\mathrm{e}} \le e_{\mathrm{th}}\}, \\ 0 \ ; & s \in \{(1,q_{\mathrm{e}})\mid 0 \le q_{\mathrm{e}} \le N\}. \end{cases}$ $e_{\mathrm{th}} \le N - b_{\max} + 1$	$\alpha_s^{D} = \begin{cases} 1 & ; \quad s \in \{(0,q_{\mathrm{e}})\mid q_{\mathrm{e}} > e_{\mathrm{th}}\}, \\ \beta & ; \quad s \in \{(0,q_{\mathrm{e}})\mid q_{\mathrm{e}} = e_{\mathrm{th}}\}, \\ 0 & ; \quad \text{Otherwise}, \end{cases}$ e_{th} and β are derived in Algorithm 2

as functions of e_{th} and β as follows

$$\overline{\alpha_s^{\mathrm{D}}}(e_{\mathrm{th}},\beta) = \sum_{j=e_{\mathrm{th}}+1}^{N} \pi_{(0,j)} + \beta\pi_{(0,e_{\mathrm{th}})}, \tag{29}$$

$$\overline{q_{\mathrm{d}}}(e_{\mathrm{th}},\beta) = \frac{1}{2}\Big(1-p_{\mathrm{SD}}^{\mathrm{det}}\Big)p_{\mathrm{SR}}^{\mathrm{det}}\left(\sum_{j=e_{\mathrm{th}}+1}^{N}\pi_{(0,j)} + \beta\pi_{(0,e_{\mathrm{th}})}\right) + \sum_{j=0}^{N}\pi_{(1,j)}, \tag{30}$$

where $\pi_{(i,j)}$ is the steady state probability of state (i, j). In (29), $\overline{\alpha}(e_{\mathrm{th}},\beta)$ is the expected value of α under the policy (28). In (30), the first term accounts for the slots in which $\mathrm{Q_d}$ is empty in the first subslot. However, it gets backlogged in the second subslot since node R switches to the DD mode, receives a packet successfully (with probability $(1-p_{\mathrm{SD}}^{\mathrm{det}})p_{\mathrm{SR}}^{\mathrm{det}}$), and stores it at $\mathrm{Q_d}$, in the first subslot. Likewise, the second term stands for the slots in which $\mathrm{Q_d}$ is backlogged from the beginning of the slot. To find optimal e_{th} and β, the following minimization is solved

$$\min_{\substack{e_{\mathrm{th}}\in\{N-b_{\max}+1,\ldots,N\},\\ \beta\in(0,1]}} \tau(\alpha_s^{\mathrm{D}}) = \frac{\overline{q_{\mathrm{d}}}(e_{\mathrm{th}},\beta)+1}{\big(1-p_{\mathrm{SD}}^{\mathrm{det}}\big)p_{\mathrm{SR}}^{\mathrm{det}}\overline{\alpha_s^{\mathrm{D}}}(e_{\mathrm{th}},\beta)+p_{\mathrm{SD}}^{\mathrm{det}}}, \tag{31}$$

where $\tau(\alpha_s^{\mathrm{D}})$ is written using $\tau(\alpha_s) = \frac{\overline{q}_{\mathrm{d}}(\alpha_s)+1}{\lambda_{\mathrm{S}}(\alpha_s)}$ and (23). We prove in Appendix VI that $\tau(\alpha_s^{\mathrm{D}})$ is a convex function of β. Then, we can use the golden section search method in Algorithm 2 to find the optimum value of β for a given e_{th}. Meanwhile, the optimum value of e_{th} is derived by exhausive search in $\{N-b_{\max}+1,\ldots,N\}$. Table II summarizes the throughput-optimal and delay-optimal static and dynamic policies.

It is worth mentioning that if N is large enough compared to $b_{\max}$, the average transmission time of the packets that are received in the energy states between $N-b_{\max}+1$ and N do not vary considerably (i.e., they are approximately equal to t_N) and thus, $\overline{D}_{\mathrm{R}} = t_N$. Therefore, by replacing $\overline{D}_{\mathrm{R}} = t_N$ in (23) and taking the first derivative, we have $\frac{d\tau}{d\overline{\alpha_s}} = \frac{(1-p_{\mathrm{SD}}^{\mathrm{det}})p_{\mathrm{SR}}^{\mathrm{det}}(p_{\mathrm{S}}^{\mathrm{det}}t_N-1)}{(\overline{\alpha_s}(1-p_{\mathrm{SD}}^{\mathrm{det}})p_{\mathrm{SR}}^{\mathrm{det}}+p_{\mathrm{SD}}^{\mathrm{det}})^2}$. Then, we have $\frac{d\tau}{d\overline{\alpha_s}} < 0$ since $t_N < \frac{1}{p_S^{det}}$ (please refer to Proposition 3). Thus, in order to minimize τ, e_{th} should be chosen such that $\overline{\alpha_s}$ or equivalently the throughput is maximized, i.e., $e_{\mathrm{th}} = N - b_{\max} + 1$, according to Proposition 2. Hence, in this case, the throughput-optimal policy, equivalent to $e_{\mathrm{th}} = N - b_{\max} + 1$ and $\beta = 1$ in (28), is a good approximation for the delay-optimal policy. This has been confirmed by numerical results in the next section.

Algorithm 2 Delay-Optimal Dynamic Policy

1: $e_{\mathrm{th}} = N$ and $\tau_{\min} = 1/p_{\mathrm{SD}}^{\mathrm{det}}$ (i.e., non-cooperation delay)
2: **while** $e_{\mathrm{th}} \ge N - b_{\max} + 1$ **do**
3: $\beta = 0$ and $\beta' = 1$
4: **while** ($\frac{(\beta'-\beta)}{\beta'} \ge \epsilon$) **do**
5: $\beta_1 = \beta + 0.382(\beta'-\beta)$ and $\beta_1' = \beta' - 0.382(\beta'-\beta)$
6: Compute $\tau(e_{\mathrm{th}},\beta_1)$ and $\tau(e_{\mathrm{th}},\beta_1')$ from QBD process
7: if $\tau(e_{\mathrm{th}},\beta_1) < \tau(e_{\mathrm{th}},\beta_1')$ then $\beta' = \beta_1'$. Otherwise, $\beta = \beta_1$.
8: **end while**
9: if $\tau(e_{\mathrm{th}},\beta) \le \tau_{\min}$, then $\tau_{\min} = \tau(e_{\mathrm{th}},\beta)$
10: $e_{\mathrm{th}} = e_{\mathrm{th}} - 1$
11: **end while**

V. NUMERICAL RESULTS

In this section, we validate our analysis through numerical simulations. We also compare optimal dynamic and static policies to investigate the advantages of dynamic policies over the static ones in different conditions.

A. Simulation Setup

We consider practical channel models and then derive the parameters of our system model, including $p_{\mathrm{XY}}^{\mathrm{det}}$ ($\mathrm{XY} \in \{\mathrm{SD},\mathrm{SR},\mathrm{RD}\}$) and γ_m, from the characteristics of the channel models. We assume that the channel S-R is a Rician block fading one due to the existence of line of sight between nodes S and R. However, the channels S-D and R-D are assumed to be Rayleigh block fading ones. Moreover, h_{XY} denotes the coefficient of the channel $\mathrm{X}-\mathrm{Y}$ ($\mathrm{X} \in \{\mathrm{S},\mathrm{R}\},\ Y \in \{\mathrm{R},\mathrm{D}\}$). Also, the nodes S, R, and D are located in (0 m, 0 m), (0 m, 40 m), and (20 m, 10 m). Thus, we have $d_{\mathrm{SD}} = 40$ m and $d_{\mathrm{SR}} = d_{\mathrm{RD}} = 10\sqrt{5}$ m, where d_{XY} denotes the distance between nodes X and Y.

Deriving $p_{\mathrm{XY}}^{\mathrm{det}}$: As mentioned before, $p_{\mathrm{XY}}^{\mathrm{det}}$ is the average packet detection probability of the channel $\mathrm{X}-\mathrm{Y}$ and thus, $1-p_{\mathrm{XY}}^{\mathrm{det}}$ is the corresponding average packet error rate (PER). We use the upper bound presented in [25] for PER in Rayleigh block fading channels. Assuming a BPSK modulation, we can write the mentioned upper bound as $1-e^{-\frac{0.5logL+0.285}{\omega_{\mathrm{XY}}}}$, where

L denotes the packet length in bits and $\omega_{\rm XY}$ is the average signal to noise ratio at node Y. Assuming $L = 256$ bits, we have $p_{\rm XY}^{\rm det} = e^{-\frac{4.285}{\omega_{\rm XY}}}$. Also, $\omega_{\rm XY} = \frac{P_{\rm X} E\{|h_{\rm XY}|^2\}}{\sigma^2 d_{\rm XY}^{\eta}}$, where η indicates the path loss exponent, σ is the noise variance, $P_{\rm X}$ denotes the transmission power at node X ($\rm X \in \{S, R\}$), and $E\{.\}$ indicates the mean operator. The typical values of the parameters are chosen to be $\eta = 3$, $\sigma^2 = -70$ dBm, $P_{\rm S} = 10$ dBm, $P_{\rm R} = 3$ dBm, $E\{|h_{\rm SR}|^2\} = 2 \times 10^{-2}$, $E\{|h_{\rm SD}|^2\} = 2 \times 10^{-3}$, $E\{|h_{\rm RD}|^2\} = 2 \times 10^{-2}$. According to the mentioned typical values, the detection probabilities of the channels S-D and R-D are derived as $p_{\rm SD}^{\rm det} = 0.25$ and $p_{\rm RD}^{\rm det} = 0.88$. For the Rician block fading channel S-R, we use simulations to derive the corresponding PER and then, $p_{\rm SR}^{\rm det}$. In this respect, using the aforementioned setting, we have $p_{\rm SR}^{\rm det} = 0.99$.

Deriving γ_m: We assume that the capacity of the energy buffer is equal to gP_S, where $g \geq 0$ [27]. Then, each unit of energy is equivalent to $\epsilon = \frac{gP_S}{N}$ Joules since the capacity of the energy buffer is divided into N levels. We choose $g = 1$ and $N = 100$. Thus, we have $\epsilon = 10^{-4} J$. Moreover, $K = \frac{P_R}{\epsilon} = 20$ units of energy. We assume that node R harvests its energy from ambient RF waves (e.g., a nearby power source) and the channel node R is harvesting energy from, is a Rician block fading one due to the existence of line of sight. Then, the channel gain has a noncentral-χ^2 distribution with two degrees of freedom. Consequently, the energy harvested in a slot, denoted by T, has the same distribution since it is proportional to the channel gain. Therefore, the cumulative distribution function (CDF) of T is written as $F_T(t) = 1 - {\rm Q}_1(\sqrt{2K'}, \sqrt{\frac{2(K'+1)t}{E\{T\}}})$, where K' is the Rice factor and is defined as the ratio of power of the line-of-sight component to the scattered components. Moreover, $Q_1(.,.)$ is the first order Marcum Q-function [26]. Consequently γ_m, i.e., the probability of harvesting m energy units in a subslot, is derived as $\gamma_m = Pr\{(m - 0.5)\epsilon \leq T < (m + 0.5)\epsilon\} = F_T((m + 0.5)\epsilon) - F_T((m - 0.5)\epsilon)$. By choosing $K' = 6$dB and $E\{T\} = 2.01 \times 10^{-4}$ J, we have $E\{\gamma_m\} = 2$ units of energy. Also, for instance, the first seven components of the EH profile are $\gamma_{[0,\ldots,6]} = [0.067, 0.322, 0.314, 0.179, 0.076, 0.027, 0.008]$. Note that, for $m \geq 20$, we have $\gamma_m \leq 10^{-11}$. Thus, we truncate the EH profile, γ_m, and consider $b_{\max}$ to be 20 units of energy. Note that the probability of harvesting greater than $K = 20$ energy units is very low and thus, our assumption $b_{\max} \leq K$ is a valid assumption. However, for an exact evaluation of our approach, we consider the real EH profile in simulations. Unless otherwise stated, the aforementioned typical values in this section are used in deriving the figures in the following.

B. Simulation and Analytical Results

In Fig. 4, the average transmission delay of source packets, τ, is plotted versus α of the static policy, and $\overline{\alpha_s}$ of the dynamic policy, respectively, in different values of $p_{\rm RD}^{\rm det}$. The value of $p_{\rm RD}^{\rm det}$ is changed by changing $E\{|h_{\rm RD}|^2\}$. For generating different values of $\overline{\alpha_s}$ in dynamic policy, we choose different values for $e_{\rm th}$ and β in (28) and derive the corresponding $\overline{\alpha_s}$ from (29). Then, the static policy is evaluated at the same values of $\overline{\alpha_s}$. Also, note that α (or $\overline{\alpha_s}$) can be directly translated to the throughput, according to (14), i.e., as α increases in horizontal axis of the figures, the throughput increases as well. Thus, α ($\overline{\alpha_s}$) corresponding to the endpoints of the curves results in the optimal throughput. The figures show that τ is a convex function of α or equivalently the data arrival rate at node R (see (5)). Moreover, as shown in Fig. 4(a), when $p_{\rm RD}^{\rm det}$ is low, the number of retransmissions of a packet at node R is so large that the cooperation does not improve the average transmission delay. However, if $p_{\rm RD}^{\rm det}$ is increased to 0.67 and 0.88 as in Fig. 4(b) and 4(c), cooperation reduces the average transmission delay for a certain range of DD probabilities. In particular, when $\alpha < 0.03$ in Fig. 4(b) and $\alpha < 0.05$ in Fig. 4(c), the performance of the static and dynamic policies are the same. This is due to the fact that at these values of α, no queue is formed yet at node R in the static policy. Thus, its performance is similar to the optimal dynamic policy, which prevents any queue built-up by accepting at most one packet at node R, according to Conjecture 1. In addition, in this case, the delay decreases by increasing α, because more packets are transmitted through the better physical channel ($p_{\rm RD}^{\rm det}$). However, when α becomes larger ($\alpha > 0.03$ in Fig. 4(b) and $\alpha > 0.05$ in Fig. 4(c)), the queueing delay at node R in the static policy becomes noticeable, so that an explosive growth in the average transmission delay is observed. This is unlike the optimal dynamic policy that accepts at most one packet at a time and thus, prevents the queue build-up at node R. It is worth noting that in Fig. 4(b) and (c), the throughput-optimal dynamic policy and the delay-optimal dynamic policy are the same. However, when the static policies are deployed, the optimal throughput and delay cannot be attained at the same time. We have also plotted the average delay corresponding to the greedy policy in Fig. 4, where the greedy policy is defined as the policy in (28) with $e_{\rm th} = K$ and $\beta = 1$. In fact, in the greedy policy the relay enters the DD mode as soon as the energy storage exceeds K energy units, when $q_{\rm d} = 0$, and remains in the EH mode during the service time of a packet. As can be seen in Fig. 4(a) and (b), when $p_{\rm RD}^{\rm det}$ is relatively low, the greedy policy performs worse than the no-cooperation policy and static policy in some range of DD probabilities. The reason is that on the one hand, the average number of retransmissions at node R is large at low $p_{\rm RD}^{\rm det}$. On the other hand, since the greedy policy does not store enough energy prior to decoding a packet, the packet has to wait before each (re)transmission till enough energy is accumulated. Moreover, when $p_{\rm RD}^{\rm det}$ increases to 0.88 in Fig. 4(c), the greedy policy performs better than the static policy in all DD probabilities. However, its average delay is less than optimal average delay in the dynamic policy, i.e., the delay at the endpoints of the dynamic curves.

In Fig. 5, the optimal average transmission delay of the source packets is plotted versus $p_{\rm RD}^{\rm det}$ in two different values of $E\{\gamma_m\}$. For changing $p_{\rm RD}^{\rm det}$, we have chosen $E\{|h_{RD}|^2\} \in [0, 1]$. When $p_{\rm RD}^{\rm det}$ is small, both static and dynamic policies decide not to cooperate due to large transmission delay at node R. However, when $p_{\rm RD}^{\rm det}$ is large, cooperation leads to much better performance than no-cooperation (22% improvement is observed in dynamic policy at $p_{\rm RD}^{\rm det} = 1$ and

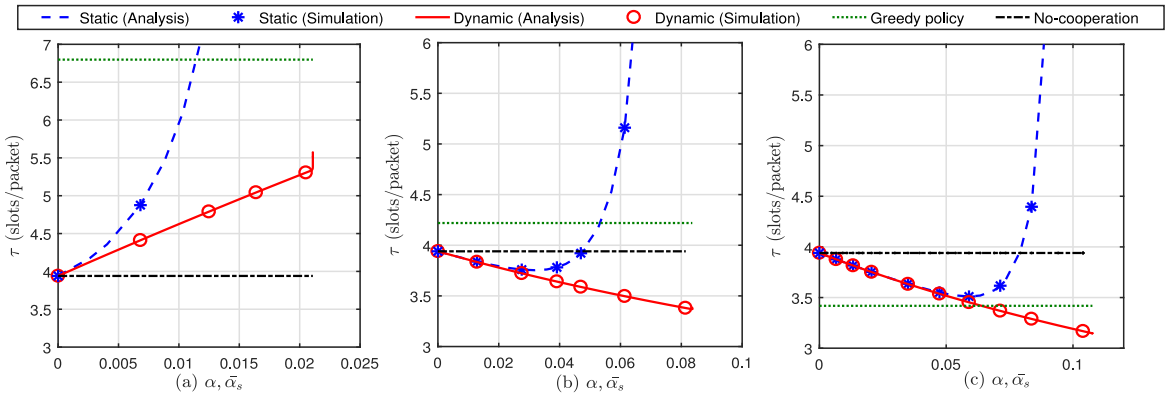


Fig. 4. Average transmission delay of source packets vs. DD probability at the relay node (a) $p_{\mathrm{RD}}^{\mathrm{det}} = 0.16$ $(E\{|h_{\mathrm{RD}}|^2\} = 1.3 \times 10^{-3})$ (b) $p_{\mathrm{RD}}^{\mathrm{det}} = 0.67$ $(E\{|h_{\mathrm{RD}}|^2\} = 6 \times 10^{-3})$ (c) $p_{\mathrm{RD}}^{\mathrm{det}} = 0.88$ $(E\{|h_{\mathrm{RD}}|^2\} = 20 \times 10^{-3})$.

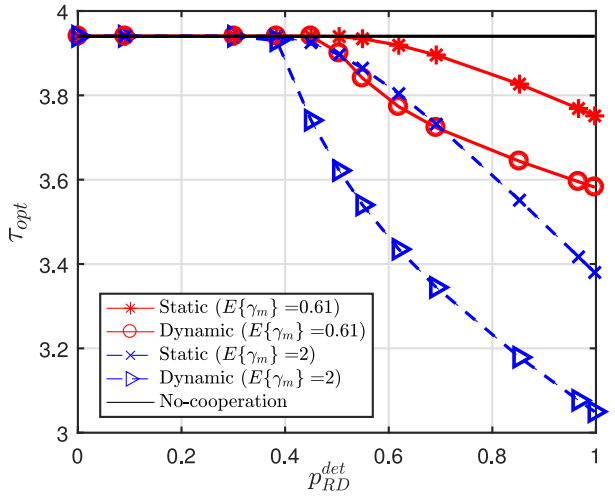


Fig. 5. Optimum average transmission delay of source packets vs. detection probability of R-D channel.

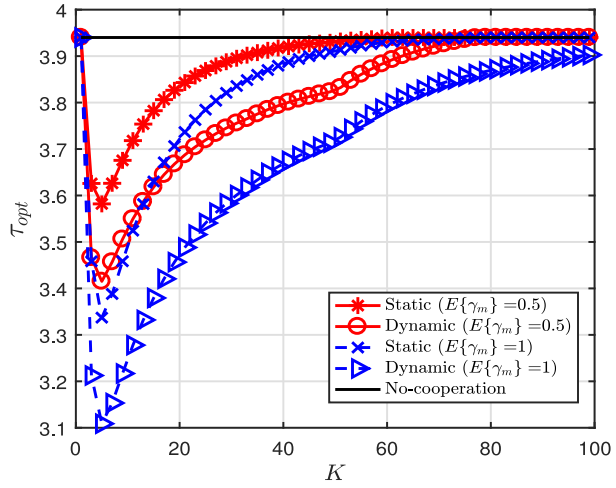


Fig. 6. Optimum average transmission delay of source packets vs. the number of energy units used at node R for transmission of a packet.

$E\{\gamma_m\} = 2$). Moreover, the increase in the energy arrival rate at node R leads to a decrease in the average transmission delay as expected.

In Fig. 6, the optimal average transmission delay is plotted versus the number of energy units needed for transmitting a single packet, i.e., K. For changing K, we have chosen $P_{\mathrm{R}} \in [-30, 6.9]$ dBm. As can be seen in Fig. 6, the optimum average transmission delay decreases by increasing K, when K is small ($K < 5$). In fact, the energy required for a transmission becomes available soon at small K and thus, the value of K is not a bottleneck. However, $p_{\mathrm{RD}}^{\mathrm{det}}$ is low and thus, a packet should be retransmitted several times. When K is increased in such a condition, $p_{\mathrm{RD}}^{\mathrm{det}}$ increases in an exponential manner (see formula of $p_{\mathrm{RD}}^{\mathrm{det}}$ in Section V-A) leading to the considerable decrease in the average number of retransmissions. Thus, the optimum average delay decreases. However, when K is large ($K > 5$ in Fig. 6), the packets wait long in average until the energy is available for transmission and thus, the value of K becomes a bottleneck. On the other hand, $p_{\mathrm{RD}}^{\mathrm{det}}$ is close to 1 and the average number of retransmissions does not change noticeably with K. Therefore, by increasing K, the average waiting time becomes longer and consequently, the optimum average delay increases.

In Fig. 7(a), the average DD probability of the policy α_s^{T} in (22), i.e., $\overline{\alpha_s^{\mathrm{T}}}$, is plotted versus e_{th}, i.e., the energy threshold to switch to the DD mode. As can be seen, in all values of N, $\overline{\alpha_s^{\mathrm{T}}}$ and consequently, the throughput increases by decreasing e_{th} from N to $N - b_{\max} + 1$. However, it remains constant for $e_{\mathrm{th}} \leq N - b_{\max} + 1$ as proved in Proposition 2. Note that since the DD probability corresponding to the optimal throughput, i.e., α^{T} in (10), is independent of N, the maximum throughput in all values of N is the same. In Fig. 7(b), the average delay at node R, i.e., $\overline{D}_{\mathrm{R}}$, of the policy α_s^{D} in (28) is plotted versus e_{th} when $\beta = 1$. For simple illustration, we only plotted $\overline{D}_{\mathrm{R}}$ for $e_{\mathrm{th}} \in [N - b_{\max} + 1, N]$, i.e., for $e_{\mathrm{th}} \in [N - 19, N]$. It can be observed that $\overline{D}_{\mathrm{R}}$ decreases as e_{th} increases. This is due to the fact that at larger e_{th}, the packets are received in higher energy levels, leading to the less average transmission time. Moreover, as can be observed in Fig. 7(b), the difference between $\overline{D}_{\mathrm{R}}$ at $e_{\mathrm{th}} = N$ and $e_{\mathrm{th}} = N - 19$ (the difference between $\overline{D}_{\mathrm{R}}$ at the endpoints of the curves) becomes less as N increases. In sufficiently large N (e.g., $N = 200$), the difference is negligible, which means that $\overline{D}_{\mathrm{R}}$ remains almost constant

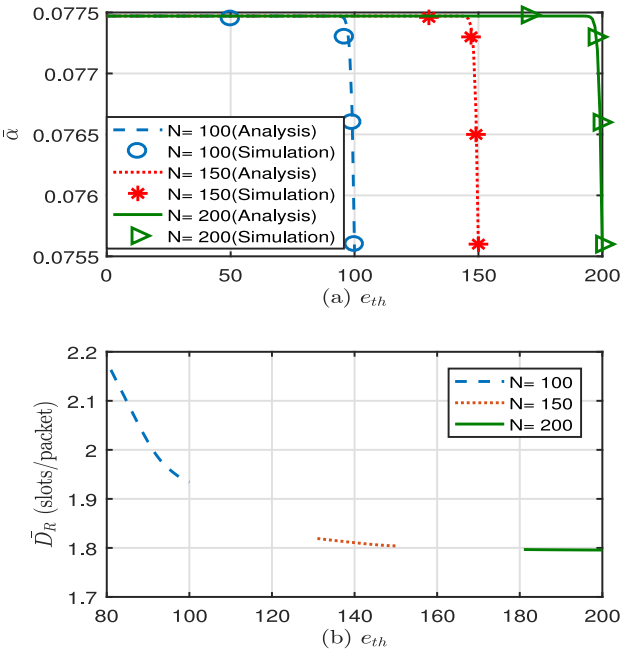


Fig. 7. (a) Average DD probability, $\overline{\alpha_s}$, vs. $e_{\rm th}$ in (22), (b) Average delay at node R, $\overline{D_{\rm R}}$, vs. $e_{\rm th}$ in (28) ($p_{\rm RD}^{\rm det} = 0.62$).

for $e_{\rm th} \in \{N - b_{\max} + 1, \ldots, N\}$. This implies that the average transmission time of a packet does not change considerably when being transmitted at different initial energy levels $\{N - b_{\max} + 1, \ldots, N\}$, i.e., it is approximately equal to t_N. Thus, regarding the explanation in the last paragraph of Section IV-B, τ is minimized when $\overline{\alpha_s}$ and consequently, the throughput has its maximum value. This happens when $e_{\rm th} = N - b_{\max} + 1$ and $\beta = 1$ in (28), according to Proposition 2. Therefore, in this case, the delay-optimal policy is also throughput-optimal.

We also derived the optimal $e_{\rm th}$ in (28) for different detection probabilities at node R, $p_{\rm R}^{\rm det} = [0.38, 0.54, 0.69, 0.85, 0.96, 0.99]$. The corresponding optimal $e_{\rm th}$s are $e_{\rm th} = [100, 98, 96, 94, 90, 84]$. It can be seen that the optimal $e_{\rm th}$ decreases from $N = 100$ to 84 as $p_{\rm RD}^{\rm det}$ increases from 0.38 to 0.99.

VI. Conclusion

We have derived the throughput-optimal and delay-optimal static and dynamic TS policies in a cooperative network with a WEH relay. We have proved that the throughput-optimal policy is obtained when the data buffer at the relay is at the boundary of stability, where the blocking probability at energy buffer is zero and all the harvested energy is consumed for transmission of data packets. Unlike the static policy, we have proved that the throughput-optimal dynamic policy has a threshold-based structure that keeps at most one data packet at the relay node. The delay-optimal static policy has been derived by analyzing an underlying QBD process. Likewise, we have shown that the delay-optimal dynamic policy is threshold-based. In particular, the policy allows at most one data packet to be backlogged at node R. Our analysis has been validated by extensive simulations, through which have also investigated the performance of the proposed policies under various settings of system parameters.

Appendix A
Transition Probabilities of the Underlying QBD Process

$\mathbf{M}_{ij}^{-}$ and $\mathbf{M}_{ij}$ are given by

$$\mathbf{M}_{ij}^{-} = \begin{cases} p_{\rm RD}^{\rm det}; & i \geq K, j = i - K, \\ 0; & \text{otherwise.} \end{cases} \tag{32}$$

$$\mathbf{M}_{ij} = \begin{cases} 1; & i < K, j = i, \\ 1 - p_{\rm RD}^{\rm det}; & i \geq K, j = i - K, \\ 0; & \text{otherwise.} \end{cases} \tag{33}$$

The first case in (32) indicates that the transmission occurs in the second subslot only when at least K energy units are available in the energy buffer and then, the number of data packets decreases by one if the transmitted packet is detected at node D with probability $p_{\rm RD}^{\rm det}$. Also, the first case in (33) states that the number of energy units and data packets does not change in the second subslot if there is not enough energy for transmission. Moreover, according to the second case of (33), a transmission occurs when enough energy is available. However, the number of data packets remains unchanged with probability $1 - p_{\rm RD}^{\rm det}$. Finally, $\mathbf{T}_{ij}$ is written as

$$\mathbf{T}_{ij} = \begin{cases} 0; & j < i, \\ \gamma_{j-i}; & i \leq j < N, \\ \sum_{l=N-i}^{b_{\max}} \gamma_l; & j = N. \end{cases} \tag{34}$$

According to the above equation, the number of energy units does not decrease in the first subslot, since node R does not transmit. Also, due to limited capacity of $\rm Q_e$, i.e., N, the energy buffer becomes full if more than $N - i$ energy units arrive when the energy state is i.

Appendix B
Proof of $\frac{d\overline{\alpha}}{d\beta} > 0$ in Proposition 3

Consider the Markov chain corresponding to the policy $\hat{\alpha}_s(\beta)$ in (25), in which $s = (q_{\rm d}, q_{\rm e}) \in \{0, 1\} \times \{0, 1, \ldots, N\}$. The transition probabilities of the mentioned Markov chain can be derived from (1) and (2) by replacing α_s with $\hat{\alpha}_s(\beta)$. Let $\mathbf{F}$ denote the corresponding transition matrix such that $\mathbf{F}_{q_{\rm d}+q_{\rm e}, q'_{\rm d}+q'_{\rm e}}$ is the transition probability from state $(q_{\rm d}, q_{\rm e})$ to the state $(q'_{\rm d}, q'_{\rm e})$. In fact, by this notation, we have assigned the first $N + 1$ rows of $\mathbf{F}$ to the transition probabilities of the states $(0, i)$ $(i \in \{0, 1, \ldots, N\})$ and the next $N + 1$ rows to the states $(1, i)$ $(i \in \{0, 1, \ldots, N\})$. Note that according to $\hat{\alpha}_s(\beta)$ in (25), the transition probabilities in $\mathbf{F}$ are all independent of β except those in $N + 1^{th}$ row since they correspond to the transitions from state $s = (0, N)$ to other states. Moreover, these transition probabilities are linear functions of β. Thus, $\mathbf{F}$ can be written as $\mathbf{F} = \hat{\mathbf{F}} + \beta\tilde{\mathbf{F}}$, where all entries in $\tilde{\mathbf{F}}$ are zero except those in $(N + 1)^{th}$ row. Also, let $\boldsymbol{\nu}$ denote the eigenvector of $\mathbf{F}$ associated to eigenvalue 1, i.e., $\boldsymbol{\nu}\mathbf{F} = \boldsymbol{\nu}$. Also, we normalize $\boldsymbol{\nu}$ such that the $N + 1^{th}$ coordinate of $\boldsymbol{\nu}$ is equal to 1, i.e., $\boldsymbol{\nu}_{N+1} = 1$. Then, by substituting $\mathbf{F} = \hat{\mathbf{F}} + \beta\tilde{\mathbf{F}}$ to

$\boldsymbol{\nu}\mathbf{F} = \boldsymbol{\nu}$ and taking the second derivative, we have

$$\frac{d^2\boldsymbol{\nu}}{d\beta^2} = \frac{d}{d\beta}\left(\boldsymbol{\nu}\tilde{\mathbf{F}} + \frac{d\boldsymbol{\nu}}{d\beta}\hat{\mathbf{F}} + \beta\frac{d\boldsymbol{\nu}}{d\beta}\tilde{\mathbf{F}}\right) = \frac{d\boldsymbol{\nu}}{d\beta}\tilde{\mathbf{F}} + \frac{d^2\boldsymbol{\nu}}{d\beta^2}\hat{\mathbf{F}} = \frac{d^2\boldsymbol{\nu}}{d\beta^2}\hat{\mathbf{F}}. \tag{35}$$

Here, to verify the second equality we note that by the normalization $\boldsymbol{\nu}_{N+1} = 1$, we have $\frac{d\boldsymbol{\nu}_{N+1}}{d\beta} = 0$, i.e., the $(N+1)^{th}$ coordinate of the vector $\frac{d\boldsymbol{\nu}}{d\beta}$ is zero. As a result, $\beta\frac{d\boldsymbol{\nu}}{d\beta}\tilde{\mathbf{F}}$ vanishes since all rows of $\tilde{\mathbf{F}}$ except row $N+1$ are zero. Then by (35) the vector $\frac{d^2\boldsymbol{\nu}}{d\beta^2}$ is either zero or an eigenvector of $\hat{\mathbf{F}}$ with eigenvalue 1. Assume the latter case. Also, let $\hat{\boldsymbol{\pi}}$ denote the stationary probability vector of $\hat{\mathbf{F}}$, i.e., $\hat{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}}\hat{\mathbf{F}}$. Then, from (35), we have $\frac{d^2\boldsymbol{\nu}}{d\beta^2} = c\hat{\boldsymbol{\pi}}$, where c is a constant. Consequently, $\frac{d^2\boldsymbol{\nu}_{N+1}}{d\beta^2} = c\hat{\boldsymbol{\pi}}_{N+1}$. Now, on the one hand, we have $\frac{d^2\boldsymbol{\nu}_{N+1}}{d\beta^2} = 0$ due to our assumption $\boldsymbol{\nu}_{N+1} = 1$. On the other hand, $\hat{\boldsymbol{\pi}}_{N+1} = 1$ since $\hat{\mathbf{F}}$ is the transition matrix of the policy $\hat{\alpha}_s(\beta)$ with $\beta = 0$, according to $\mathbf{F} = \hat{\mathbf{F}} + \beta\tilde{\mathbf{F}}$, for which node R always remains in EH mode and thus, the steady state is $(0, N)$. Therefore, we conclude that $c = 0$. As a result, $\frac{d^2\boldsymbol{\nu}}{d\beta^2} = 0$ and thus, $\boldsymbol{\nu} = \beta\mathbf{a} + \mathbf{b}$, where $\mathbf{a}_{N+1}$ is zero and $b_{N+1} = 1$ since $\boldsymbol{\nu}_{N+1} = 1$.

Now, let $\boldsymbol{\pi}$ be the stationary probability vector of $\mathbf{F}$, i.e., it is an eigenvector of $\mathbf{F}$, where the summation of its entries is equal to one. Also, recall that $\boldsymbol{\nu}$ is also an eigenvector of $\mathbf{F}$. Thus, we have $\boldsymbol{\pi} = \frac{\boldsymbol{\nu}}{\boldsymbol{\nu}^T\mathbf{1}}$. Therefore, regarding that $\boldsymbol{\nu} = \beta\mathbf{a} + \mathbf{b}$, $\boldsymbol{\pi}_{N+1}$ is given by $\boldsymbol{\pi}_{N+1} = \frac{\mathbf{b}_{N+1}}{\beta\,\mathbf{a}^T\mathbf{1} + \mathbf{b}^T\mathbf{1}}$. Thus, the average DD probability corresponding to policy $\hat{\alpha}_s(\beta)$ in (25) is $\overline{\hat{\alpha}_s(\beta)} = \beta\boldsymbol{\pi}_{N+1} = \frac{\beta\,\mathbf{b}_{N+1}}{\beta\,\mathbf{a}^T\mathbf{1} + \mathbf{b}^T\mathbf{1}}$. Then, we have $\frac{d\overline{\hat{\alpha}_s(\beta)}}{d\beta} = \frac{\mathbf{b}^T\mathbf{1}\mathbf{b}_{N+1}}{(\beta\,\mathbf{a}^T\mathbf{1} + \mathbf{b}^T\mathbf{1})^2}$. Note that $\mathbf{b}$ is a scaled version of $\hat{\pi}$ since it corresponds to $\beta = 0$. Therefore, its components are non-negative such that $\mathbf{b}^T\mathbf{1} > 0$. Consequently, $\frac{d\overline{\hat{\alpha}_s(\beta)}}{d\beta} > 0$, meaning that $\hat{\alpha}_s(\beta)$ is an increasing function of β.

APPENDIX C
PROOF OF COVEXITY OF $\tau(e_{\rm th}, \beta)$ VERSUS β

If Conjecture 1 holds, then $\tau(e_{\rm th}, \beta)$, i.e., the average delay corresponding to $\alpha_s^{\rm D}$ in (28), is a convex function of β. In order to prove the convexity, we should prove the following inequality for any $\theta \in [0, 1]$

$$\theta\tau(e_{\rm th}, \beta_1) + (1-\theta)\tau(e_{\rm th}, \beta_2) > \tau(e_{\rm th}, \theta\beta_1 + (1-\theta)\beta_2), \tag{36}$$

where $\beta_1, \beta_2 \in [0, 1]$. The left-hand side of (36) is the average transmission delay in the policy obtained by time sharing between two policies with the same structure as α_s^D in (28) but different DD probabilities at state $(0, e_{\rm th})$, i.e., β_1 and β_2, respectively. Also, observe that the average DD probability of the time-sharing policy at state $(0, e_{\rm th})$ is equal to $\theta\beta_1 + (1-\theta)\beta_2$. The time sharing policy is not necessarily stationary. However, due to the unichain Markovian structure of our problem, a stationary policy exists with the same delay as the delay of time-sharing policy [30]. On the other hand, the right-hand side of (36) is the minimum delay among all stationary policies with the average DD probability $\theta\beta_1 + (1-\theta)\beta_2$ at state $(0, e_{\rm th})$. Thus, (36) holds.

REFERENCES

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

[2] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1612–1616.

[3] P. Grover and A. Sahai, "Shannon meets tesla: Wireless information and power transfer," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2363–2367.

[4] K. Huang and E. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5927–5941, Dec. 2013.

[5] 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.

[6] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4210–4221, Aug. 2013.

[7] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[8] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Block-wise time-switching energy harvesting protocol for wireless-powered AF relays," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 80–85.

[9] Y. Gu and S. Aïssa, "Interference aided energy harvesting in decode-and-forward relaying systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2014, pp. 5378–5382.

[10] H. Wang, W. Wang, and Z. Zhang, "Opportunistic wireless information and energy transfer for sustainable cooperative relaying," in *Proc. IEEE Int. Conf. Commun.*, Sep. 2015, pp. 160–165.

[11] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.

[12] M. Moradian and F. Ashtiani, "Optimal relaying in a slotted aloha wireless network with energy harvesting nodes," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1680–1692, Aug. 2015.

[13] M. Moradian and F. Ashtiani, "On the trade-off between collision and cooperation in a random access wireless network with energy harvesting nodes," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2501–2513, Mar. 2018.

[14] J. Yang, Q. Yang, K. S. Kwak, and R. R. Rao, "Power–delay tradeoff in wireless powered communication networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3280–3292, Apr. 2017.

[15] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Wireless-powered relays in cooperative communications: Time-switching relaying protocols and throughput analysis," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1607–1622, May 2015.

[16] K.-H. Liu, "Performance analysis of relay selection for cooperative relays based on wireless power transfer with finite energy storage," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5110–5121, Jul. 2016.

[17] Y. Gu, H. Chen, Y. Li, Y.-C. Liang, and B. Vucetic, "Distributed multi-relay selection in accumulate-then-forward energy harvesting relay networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 74–86, Mar. 2018.

[18] M. M. I. Rajib and A. Nasipuri, "Delay performance of intermittently connected wireless sensor networks with cooperative relays," in *Proc. IEEE Int. Conf. Commun.*, Sep. 2015, pp. 1994–1999.

[19] L. Tang, X. Zhang, and X. Wang, "Joint data and energy transmission in a two-hop network with multiple relays," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 2015–2018, Nov. 2014.

[20] L. P. Qian, G. Feng, and V. C. M. Leung, "Optimal transmission policies for relay communication networks with ambient energy harvesting relays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3754–3768, Dec. 2016.

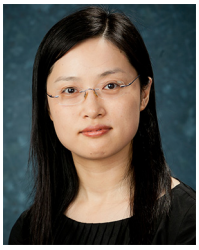
[21] K. Li, W. Ni, L. Duan, M. Abolhasan, and J. Niu, "Wireless power transfer and data collection in wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2686–2697, Mar. 2018.

[22] Y. Liu, Q. Chen, and X. Tang, "Adaptive buffer-aided wireless powered relay communication with energy storage," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 2, pp. 432–445, Jun. 2018.

[23] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
[24] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.
[25] Y. Xi, A. Burr, J. Wei, and D. Grace, "A general upper bound to evaluate packet error rate over quasi-static fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1373–1377, May 2011.
[26] D. Mishra, S. De, and C.-F. Chiasserini, "Joint optimization schemes for cooperative wireless information and power transfer over Rician channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 554–571, Feb. 2016.
[27] I. Krikidis, S. Timotheou, and S. Sasaki, "RF energy transfer for cooperative networks: Data relaying or energy harvesting?" *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1772–1775, Nov. 2012.
[28] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Philadelphia, PA, USA: SIAM, 1999.
[29] K. P. Chong and S. H. Zak, *An Introduction to Optimization*. New York, NY, USA: Wiley, 2007.
[30] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 1994.
[31] L. Kleinrock, *Queueing Systems*, vol. 1. New York, NY, USA: Wiley, 1975.
[32] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. San Rafael, CA, USA: Morgan and Claypool, 2010.

Masoumeh Moradian (S'11–M'16) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2007, 2010, and 2016, respectively. She was a Visiting Scholar with the Chinese University of Hong Kong in 2015. She is currently a Post-Doctoral Researcher with the Institute for Research in Fundamental Sciences (IPM), Tehran. Her research interests include energy harvesting communication networks, queueing theory, and network stochastic optimization.

Farid Ashtiani (S'02–M'03) received the Ph.D. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2003. From 1995 to 1999, he was with the Power Research Center and Niroo Research Institute of Iran. From 1999 to 2001, he was a Research Staff Member with the Advanced Communication Science Research Laboratory, Iran Telecommunication Research Center, Tehran. Since 2003, he has been with the Department of Electrical Engineering, Sharif University of Technology, where he is currently an Associate Professor and a Technical Member of Advanced Communications Research Institute. His research interests include queueing theory, modeling, analysis, and design of different types of wireless networks, and mobility modeling.



Ying Jun (Angela) Zhang (S'00–M'05–SM'10) received the Ph.D. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2004. Since 2005, she has been with the Department of Information Engineering, Chinese University of Hong Kong, where she is currently an Associate Professor. Her research interests include mainly wireless communications systems and smart power systems, in particular optimization techniques for such systems. She has served as the Chair of the Executive Editor Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She has also served many years as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, SECURITY AND COMMUNICATIONS NETWORKS (Wiley), and a Feature Topic in *IEEE Communications Magazine*. She has served on the organizing committee of major IEEE conferences, including ICC, GLOBECOM, SmartgridComm, VTC, CCNC, ICCC, and MASS. She was a recipient of the Young Researcher Award from the Chinese University of Hong Kong in 2011 and the co-recipient of the 2014 IEEE ComSoc APB Outstanding Paper Award, the 2013 IEEE SmartgridComm Best Paper Award, and the 2011 IEEE Marconi Prize Paper Award on Wireless Communications. She is currently the Chair of IEEE ComSoc Emerging Technical Committee on Smart Grid. She was the Co-Chair of the IEEE ComSoc Multimedia Communications Technical Committee and the IEEE Communication Society GOLD Coordinator. As the only winner from engineering science, she has won the Hong Kong Young Scientist Award 2006, conferred by the Hong Kong Institution of Science. She is a fellow of IET and a Distinguished Lecturer of IEEE ComSoc.