On Energy Harvesting Gain and Diversity Analysis in Cooperative Communications

Meng-Lin Ku, Wei Li, Yan Chen and K. J. Ray Liu

***Abstract*—The use of energy harvesting cooperative relays is a promising solution to battery-limited wireless networks. In this paper, we consider a cooperative system in which one source node transmits data to one destination with the assistance of an energy harvesting decode-and-forward relay node. Our objective is to minimize the long-term average symbol error rate (SER) performance through a Markov decision process (MDP) framework. By doing so, we find the optimal stochastic power control at the relay that adapts the transmission power to the changes of energy harvesting, battery, channel and decoding states. We derive a closed-form expression for the exact SER of the cooperative system. Further insights are gained by analyzing the asymptotic SER and its lower and upper bounds at high signal-to-noise ratio (SNR), and the performance is eventually characterized by the occurrence probability of the relay's actions at the worst channel states in the MDP. We also show that the optimal cooperative policy at asymptotically high SNR follows a threshold-type structure, i.e., the relay spends the harvested energy only when the signal is successfully decoded and the source is faced with the worst channel condition in its direct link. Using these observations to quantify the diversity gain and the energy harvesting gain, we reveal that full diversity is guaranteed if and only if the probability of harvesting zero energy quantum is zero, which can be achieved by reducing the energy quantum size or increasing the energy harvesting capability. Finally, we present several numerical examples to validate the analytical findings.**

***Index Terms*—Energy harvesting, cooperative communication, energy harvesting gain, diversity order.**

I. Introduction

Wireless communications are often vulnerable to small-scale fading caused by multipath channel propagation. In past few years, cooperative communications have gained much interest to mitigate this negative effect through the use of relays to reap the inherent spatial diversity gains [1]. This is particularly attractive when it is unaffordable to install multiple antennas on size-limited communication nodes. Various cooperative techniques have been proposed and analyzed in terms of the outage probability or the symbol error rate (SER), among which decode-and-forward (DF) and amplify-and-forward (AF) are deemed as the most popular ones to provide full diversity gains and to make a more efficient use of transmission power [2], [3]. It has been shown in [1] that the DF protocol, in which the relay first decodes the received signal, re-encodes it and then forwards it to the destination if the decoding is correct, has better performance than the AF protocol, in which the relay simply amplifies the received signal and forwards it.

In many wireless applications, wireless nodes are untethered to an energy infrastructure and can only be powered by batteries with limited capacity. This major limitation requires frequent battery replacement to prolong network lifetime when the battery is exhausted. Such an embarrassment of energy shortage is even challenging for cooperative communications as wireless cooperative nodes are often subject to space limitation to utilize multiple antennas, not to mention the use of a large battery with long lifetime. In addition, the replacement of batteries may be inconvenient, costly or dangerous in some applications, e.g., environmental monitoring in wireless sensor networks. Recently, energy harvesting has become an attractive option to wireless nodes by scavenging ambient energy from environments such as solar, wind, thermoelectric or motion effects, etc [4]. Thus, it is a naturally evolutionary step to consider wireless cooperative nodes powered by energy harvesting devices. In spite of a potentially infinite amount of energy available at nodes, the dynamics of the harvested energy and the limited capacity of rechargeable batteries motivate us to revisit the classical problems of power management and to design more efficient energy usage schemes.

When cooperative communications meet energy harvesting, three interesting questions are raised: (1) Can/How a source that cooperates with an energy harvesting relay node achieve a full diversity gain in reality ? (2) What is the optimal relay transmission policy for achieving the full diversity ? (3) Except for the diversity gain, what is the impact of energy harvesting on the performance gain in terms of signal-to-noise power ratio (SNR) ? In traditional wireless systems, the SNR performance gain is often termed as the coding gain. Here, a new terminology, *energy harvesting gain*, is introduced instead to emphasize the influence of energy harvesting on the coding gain performance of cooperative communications. While energy harvesting has been extensively investigated in the recent literature, the aforementioned questions have not been fully addressed and remain to be answered, and these questions are important toward understanding the fundamental performance limit of the cooperative networks with energy harvesting capability.

Extensive research efforts have been devoted to energy management problems for various energy harvesting communications [5]–[12]. In [5], the authors investigated a directional water-filling algorithm to maximize the short-term throughput for a wireless link with an energy harvesting transmitter by

Meng-Lin Ku is with the Department of Communication Engineering, National Central University, Jung-li 32001, Taiwan (e-mail: mlku@ce.ncu.edu.tw).

Wei Li, Yan Chen and K. J. Ray Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: wli52140@umd.edu; yan@umd.edu; kjrliu@umd.edu).

Wei Li is also with the Department of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an, 710049, China.

assuming that the harvested energy and channel fading states are known non-causally. By using a deterministic energy harvesting model, packet scheduling problems that aim at maximizing the throughput or minimizing the transmission time were studied in [6] for a point-to-point communication system with an unlimited-capacity battery, while the design was later generalized in [7] with finite energy storage. The work of [8] designed power allocation for throughput maximization over a finite horizon with a preset energy arrival profile. However, all these works require tight prediction on the non-causal side information of energy amount arrivals, and this becomes very challenging when the energy management interval is enlarged. As an alternative, some other works adopted stochastic energy harvesting models under which the energy arrivals are described in a probability sense [8]–[12]. In [8], dynamic programming was employed to find the optimal power allocation scheme that maximizes the throughput according to a Markov random energy harvesting model. When the energy and packet arrivals are formulated as Markov processes, sleep and wake-up strategies were developed for wireless solar-power sensor networks in [10]. The authors in [11] and [12] proposed data-driven stochastic models, and power control and adaptive modulation were jointly designed to maximize the net bit rate through a discounted Markov decision process (MDP).

On the other hand, recent attention has been paid to energy harvesting cooperative communications [13]–[24]. By maximizing the short-term throughput, the authors in [13] investigated directional water-filling power control schemes for an energy harvesting source and a conventional half-duplex relay with constant power in two-hop communication systems. In [14], power allocation problems were addressed for a scenario that both source and half-duplex DF relay nodes are self-sustained with energy harvesting, subject to different data traffic delay constraints. The work in [15] proposed offline and online power allocation algorithms for maximizing throughput in the conventional and buffer-aided single link cooperative systems with energy harvesting source and relay nodes. In [16], the problem of throughput maximization in an energy harvesting two-hop AF relay network was carried out by considering the non-causal or causal knowledge of harvested energy. The optimal energy expenditure schemes were also discussed in [17] and [18] for full-duplex relaying protocols, while two-way relay channels with energy harvesting nodes were considered in [19], [20]. Moreover, when only partial state information about the relay is available at the source node, the transmission scheduling problem was casted as a partially observable Markov decision process (POMDP) in [21]. However, the aforementioned works primarily focused on the data throughput maximization problem and the development of the optimal solution and its property under different network settings. Only few works concentrated on analyzing the outage behavior or the SER performance. In [23], the outage probability for a cooperative network aided by energy harvesting relay nodes was derived based upon a simple on-off stochastic energy harvesting model. In [24], SER performance analysis was performed for relay selection in a cooperative network employing voluntary energy harvesting relays. However, the battery-exhausted probabilities, which depend on transmission actions and stochastic energy arrivals, were assumed to be known in the analysis of these two works, and neither of them discussed the optimal transmission policies for minimizing the outage probability or the SER performance.

Cooperative communications, if successfully implemented, is undoubtedly expected to improve the link quality of wireless networks with energy harvesting. However, a quantitative answer on the impact of energy harvesting on the SER performance as well as the potential diversity gains and energy harvesting gains is still missing. In this paper, we investigate the optimal cooperative transmission policy for an energy harvesting relay node that helps forward the signal from a source node to a destination node via a selective DF protocol. For this purpose, we resort to the MDP as a means to find out the optimal transmission action at the relay with the goal of minimizing the long-term average SER and to analyze the achievable diversity gains and energy harvesting gains of the cooperative networks. Specifically, the novelty and contribution of this paper are summarized as follows:

- A solar-data-driven stochastic energy harvesting model in [12] is utilized in the construction of the MDP design framework, in which the optimal relay transmission policy is designed in order to minimize the long-term average SER performance by adapting the transmission power to the changes of the energy harvesting, battery, channel and decoding states. Unlike the existing works that focus on either finding the optimal solution for the throughput maximization or simply analyzing the outage/SER performance without addressing the optimal transmission policy, the main goal of this paper is to analyze the SER performance of the energy harvesting cooperative communications under a realistic energy harvesting model and to analytically characterize the interplay between the attainable performance and the transmission policy.
- Based on the developed MDP framework, exact and asymptotic SER expressions are derived for the energy harvesting cooperative communications. In particular, we establish the relationship between the asymptotic SER and the occurrence probability for the adopted relay action in the MDP. Furthermore, we analyze upper and lower bounds for the asymptotic SER to quantify the diversity gains as well as the energy harvesting gains of the considered cooperative transmission policy. By theoretical analysis, a theorem regarding the accessibility of a diversity order of two is provided, and it reveals that the full diversity is achievable only if the stationary probability of the relay's actions at the worst channel states for which the decoding is successful but the relay keeps silent goes to zero. To the best of our knowledge, this is the first attempt to comprehensively study the diversity gains and energy harvesting gains of the optimal stochastic transmission policies by means of the MDP.
- We then uncover that the optimal cooperative transmission policy at asymptotically high SNR is degenerated into a threshold-type policy. That is, the relay with non-empty battery spends the harvested energy only when the source node stays at the worst channel condition in its direct link and the relay node can successfully decode the signals. With this elegant characteristic, we further explore an energy

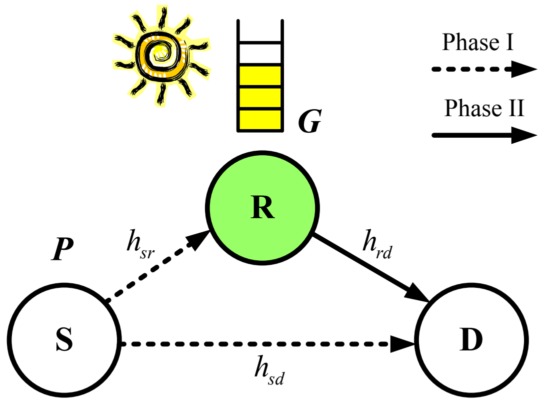


Fig. 1. Energy harvesting cooperative communications with the selective DF protocol.

quantum supporting way, along with the promising structures of policies, for achieving the full diversity order. Our analysis shows that the fully diversity can be reached if and only if the energy quantum outage probability, i.e., the probability of obtaining zero energy quantum, is equal to zero. By linking this result to the solar-data-driven energy harvesting model in [12], we prove that a zero energy quantum outage probability is attainable if a ratio between the energy quantum size and the energy harvesting capability is considerably small. Finally, some numerical examples are offered to justify the analytical derivations and the proposed theorems in this paper.

The rest of this paper is organized as follows. In Section II, we introduce the selective DF cooperative networks with energy harvesting. In Section III, an MDP design framework for finding the optimal cooperative transmission policy is presented, and the main structure results of the policy are also discussed. Section IV is devoted to derive the exact SER and the asymptotic SER, followed by the analysis of the diversity gains and the energy harvesting gains. Furthermore, we address the optimal policy at asymptotically high SNRs and the energy supporting condition for achieving the full diversity. Numerical results are presented in Section V, and conclusions are drawn in Section VI.

II. Energy Harvesting Cooperative Communications

We consider a cooperative relay network in Fig. 1, where a source (S) and a destination node (D) communicate over a wireless fading channel with the assistance of an energy harvesting relay node (R). The relaying protocol involves two signal transmission phases, and the time duration of each phase is T_P. Define h_{sr} and h_{sd} as the channel coefficients from the source to the relay and the destination, respectively, and h_{rd} as the channel from the relay to the destination. Further, the channels h_{sr}, h_{sd}, and h_{rd} are complex white Gaussian random variables with zero mean and variance η_{sr}, η_{sd} and η_{rd}. Let x be the M-ary phase-shift-keying (M-PSK) data modulated symbol of the source node, where $\mathbb{E}\left[|x|^2\right] = 1$ and the operator $\mathbb{E}[\cdot]$ takes the expectation. In the first phase, the source sends the information to the destination, and meanwhile, the information is received by the relay node. The received signals can be expressed as

$$y_{sd} = \sqrt{P_s} h_{sd} x + z_d\,; \tag{1}$$

$$y_{sr} = \sqrt{P_s} h_{sr} x + z_r\,, \tag{2}$$

where z_d and z_r are additive complex white Gaussian noise with zero mean and variance N_0, and P_s is the transmission power of the source node. From (1) and (2), the instantaneous SNRs at the relay and the destination can be calculated as

$$\Upsilon_{sd} = \frac{P_s \zeta_{sd}}{N_0}\,; \tag{3}$$

$$\Upsilon_{sr} = \frac{P_s \zeta_{sr}}{N_0}\,, \tag{4}$$

where we define $\zeta_{sd} = |h_{sd}|^2$ and $\zeta_{sr} = |h_{sr}|^2$ as the instantaneous channel power. In the second phase, the relay can decide whether to forward the decoded data symbol $\hat{x}$ with transmission power P_r or to keep silent with zero power consumption. In this paper, a selective DF strategy in [3] is adopted, and the relay can help forward the re-encoded data symbols, only if it can decode the received data symbols correctly. In practice, this can be implemented by considering an SNR threshold at the relay, and it is reasonable to assume that the relay can successfully decode the data symbols with a negligible error probability if the instantaneous SNR of its received signal is larger than a preset threshold. Hence, the received signal at the destination is written as

$$y_{rd} = \sqrt{P_r} h_{rd} x + \tilde{z}_d\,, \tag{5}$$

where $\tilde{z}_d$ stands for the noise in the second phase with the same statistic as z_d, P_r is the relay transmission power, and $P_r \neq 0$ if the data symbol is correctly decoded and forwarded to the destination by the relay; otherwise, $P_r = 0$.

We assume that the channel state information (CSI) of the wireless links h_{sd} and h_{rd} can be perfectly estimated by the destination node. With the CSI knowledge, a maximum ratio combining (MRC) scheme is utilized for combining the received signals (1) and (5) of the two phases at the destination:

$$y_c = c_1 y_{sd} + c_2 y_{rd}\,. \tag{6}$$

By applying the combining weights $c_1 = \sqrt{P_s} h_{sd}^* / N_0$ and $c_2 = \sqrt{P_r} h_{rd}^* / N_0$ into (6), the SNR of the output of the combiner y_c can be further calculated as

$$\Upsilon_c = \frac{P_s \zeta_{sd} + P_r \zeta_{rd}}{N_0}\,. \tag{7}$$

III. Stochastic Relay Transmission Policy Using Markov Decision Process

The design of the relay transmission policy depends on a couple of factors, like channel conditions H_{rd} and H_{sd} of the wireless links among nodes, battery condition Q_b, energy harvesting condition Q_e, and decodability of the relay node Q_c. Our goal is to find the optimal transmission policy by formulating the problem as an average SER minimization problem through the MDP, while concerning a limited battery recharging rate. Moreover, we intend to study the diversity and energy harvesting gains of the policy which can be formally defined as follows:

Definition 1: Let γ and $P_{SER}(\gamma)$ be the SNR and the SER, respectively. At asymptotically high SNR, if the SER is expressed as $P_{SER} \sim (g_E \cdot \gamma)^{-d}$, the constants g_E and d are the energy harvesting gain and the diversity order of the cooperative communications.

Consider a five-tuple state space $s = (H_{rd}, H_{sd}, Q_b, Q_e, Q_c) \in \mathcal{S}_{rd} \times \mathcal{S}_{sd} \times \mathcal{Q}_b \times \mathcal{Q}_e \times \mathcal{Q}_c \triangleq \mathcal{S}$, where $\mathcal{S}_{rd} = \{0, \ldots, N_{rd}-1\}$, $\mathcal{S}_{sd} = \{0, \ldots, N_{sd}-1\}$, $\mathcal{Q}_b = \{0, \ldots, N_b-1\}$, $\mathcal{Q}_e = \{0, \ldots, N_e - 1\}$, and $\mathcal{Q}_c = \{0, 1\}$. The policy is managed on the time scale of T_M which covers a number of two-phase transmissions. For each state, the relay can select to forward the information from the source to the destination by adjusting its transmission power level or just to keep silent. Detailed descriptions of these sates and actions are provided in the following.

A. Channel States

The instantaneous channel power of ζ_{rd}, and ζ_{sd}, which correspond to the relay-to-destination (RD) and source-to-destination (SD) links, can be quantized into several levels, given by $\Gamma_{rd} = \left\{0 = \Gamma_{rd}^{(0)}, \Gamma_{rd}^{(1)}, \ldots, \Gamma_{rd}^{(N_{rd})} = \infty\right\}$ and $\Gamma_{sd} = \left\{0 = \Gamma_{sd}^{(0)}, \Gamma_{sd}^{(1)}, \ldots, \Gamma_{sd}^{(N_{sd})} = \infty\right\}$, respectively. By ignoring the subscript of the notations "rd" and "sd", we say a channel ζ is in the j^{th} state, if $\Gamma^{(j)} \leq \zeta < \Gamma^{(j+1)}$, for $j = 0, \ldots, N-1$. Furthermore, we assume that the channel gain is quasi-static during the policy management period T_M, and the channel can only transit from the current state to its neighboring states. The stationary probability of the j^{th} channel state can be computed as

$$\begin{aligned} P(H = j) &= \int_{\Gamma^{(j)}}^{\Gamma^{(j+1)}} \frac{1}{\eta} \exp\left(-\frac{\zeta}{\eta}\right) d\zeta \\ &= \exp\left(-\frac{\Gamma^{(j)}}{\eta}\right) - \exp\left(-\frac{\Gamma^{(j+1)}}{\eta}\right), \end{aligned} \tag{8}$$

where $\eta = \mathbb{E}[\zeta]$ is the average channel power. Define a function $h(\zeta) = f_D \sqrt{2\pi\zeta/\eta} \exp(-\zeta/\eta)$, where f_D is the maximum Doppler frequency, normalized by the policy management period T_M. The state transition probabilities are then determined by [25]

$$\begin{aligned} &P(H = j' | H = j) \\ &= \begin{cases} \frac{h(\Gamma^{(j+1)})}{P(H=j)}, & j' = j+1, \quad j = 0, \ldots, N-2; \\ \frac{h(\Gamma^{(j)})}{P(H=j)}, & j' = j-1, \quad j = 1, \ldots, N-1; \\ 1 - \frac{h(\Gamma^{(j)})}{P(H=j)} - \frac{h(\Gamma^{(j+1)})}{P(H=j)}, & j' = j, \quad j = 1, \ldots, N-2, \end{cases} \end{aligned} \tag{9}$$

and the transition probabilities of $P(H = j | H = j)$ at the boundary states are given by

$$\begin{aligned} &P(H = 0 | H = 0) = 1 - P(H = 1 | H = 0); \\ &P(H = N-1 | H = N-1) = \\ &\qquad 1 - P(H = N-2 | H = N-1). \end{aligned} \tag{10}$$

B. Decoding States

Two decoding states are taken into consideration in the MDP: "Success" and "Fail". If $Q_c = 1$, it means that data symbols are correctly decoded by the relay during the first phase. On the contrary, the state $Q_c = 0$ indicates that the relay fails to decode the message from the source. In general, the decoding probability can be characterized by the instantaneous SNR of the source-to-relay (SR) link $\frac{P_s \zeta_{sr}}{N_0}$ and the decoding capability of the relay node which is specified by a threshold γ. We say that data symbols can be decoded if $\frac{P_s \zeta_{sr}}{N_0} \geq \gamma$, and it is noted that the relay with a smaller threshold has relatively better decoding capability. For simplicity, the channel variations of the SR link are assumed to be independent across the management periods but quasi-static within the period. Thus, the transition probability of the decoding states is irrelevant to the previous state, and it can be expressed as

$$\begin{aligned} &P(Q_c = m' | Q_c = m) \\ &= \begin{cases} P\left(\zeta_{sr} \geq \frac{\gamma N_0}{P_s}\right) = \exp\left(-\frac{\gamma N_0}{P_s \eta_{sr}}\right), & m' = 1; \\ 1 - \exp\left(-\frac{\gamma N_0}{P_s \eta_{sr}}\right), & m' = 0. \end{cases} \end{aligned} \tag{11}$$

C. Relay Actions

Let $P_r = aG$, where G is a constant transmission power level. We consider N_a possible actions for the relay node in the second phase, i.e., $a \in \mathcal{A} = \{0, \ldots, N_a - 1\}$, and the relay complies with the selective DF strategy while playing these actions. When $a \neq 0$, it means that the relay selects to help forward the information by consuming a total amount of $\frac{1}{2} a G T_M$ energy in the battery; otherwise, the relay takes no action.

D. Energy Harvesting States

A solar-data-driven energy harvesting model in [12] is adopted here. We assume that there are N_e underlying energy harvesting states, each of which stands for a meaningful energy harvesting condition like "Excellent", "Good", "Bad", etc. Each state is governed by a state transition probability $P(Q_e = l' | Q_e = l)$, for $l, l' = 0, \ldots, N_e - 1$, and it is associated with an energy harvesting probability in terms of the number of energy quanta that can be obtained from the solar power during the management period T_M. Here, one energy quantum, E_U, is defined as the total amount of energy with respect to the transmission power G during half the management period $\frac{1}{2} T_M$, i.e., $E_U = \frac{1}{2} G T_M$. In other words, the relay action is also operated in units of energy quanta. At the l^{th} state, the probability of harvesting w energy quanta is given by $P(E = w | Q_e = l)$, for $l = 0, \ldots, N_e - 1$ and $w = 0, \ldots, \infty$.

E. Battery States

The battery state transition is determined by both the transmission action and the number of harvested energy quanta at the relay. While $Q_c = 1$, the feasible action set at the b^{th} battery state is given as $\mathcal{A}_{1,b} = \{0, \ldots, \min\{b, N_a - 1\}\}$, since the maximum number of affordable energy quanta is subject to b. Otherwise, $\mathcal{A}_{0,b} = \{0\}$ for $Q_c = 0$. When the action a is taken and the number of harvested energy quanta is w, the battery state will transit from the state b to the state $b' = \max(b - a + w, N_b - 1)$ due to the finite battery storage capacity. Hence, the battery state transition probability at the l^{th} energy harvesting state is given by

$$\begin{aligned} &P_a\left(Q_b = b' \middle| Q_b = b, Q_e = l\right) \qquad (12) \\ &= \begin{cases} P\left(E = b' - b + a \middle| Q_e = l\right), & b' < N_b - 1\,; \\ P\left(E \geq N_b - 1 - b + a \middle| Q_e = l\right), & b' = N_b - 1\,, \end{cases} \end{aligned}$$

where $b' \geq b - a$, for any $a \in \mathcal{A}_{m,b}$.

F. Reward Functions

The SER of the cooperative system is adopted here to serve as the reward function in the MDP. Let $\tilde{x}$ be the decoded data symbol at the destination. The obtained reward $R^{(a)}$ at the state $s = (H_{rd}, H_{sd}, Q_b, Q_e, Q_c) = (j, k, b, l, m)$ with respect to the action $a \in \mathcal{A}_{m,b}$ is defined as

$$\begin{aligned} &R^{(a)}\left(s = (j,k,b,l,m)\right) \qquad (13) \\ &\triangleq 1 - P^{(a)}\left(\tilde{x} \neq x \middle| s = (j,k,b,l,m)\right) \\ &= \begin{cases} 1 - P^{(a)}\left(\tilde{x} \neq x \middle| (H_{rd}, H_{sd}, Q_b, Q_e) = (j,k,b,l)\,,\right. \\ \qquad\qquad \left. \hat{x} = x\right), \quad m = 1; \\ 1 - P\left(\tilde{x} \neq x \middle| H_{sd} = k\right), \quad m = 0. \end{cases} \end{aligned}$$

When the M-PSK modulation scheme is applied in the cooperative system, the first term in (13) can be calculated by substituting $P_r = aG$ into (7), as shown in (14) at the top of the next page, where $g_{sd}(\theta, x) = \exp\left(-\left(\frac{c_M P_s \eta_{sd}}{N_0 \sin^2\theta} + 1\right)\frac{x}{\eta_{sd}}\right)$, $g_{rd}(\theta, x) = \exp\left(-\left(a\frac{c_M G \eta_{rd}}{N_0 \sin^2\theta} + 1\right)\frac{x}{\eta_{rd}}\right)$, and $c_M = \sin^2\left(\frac{\pi}{M}\right)$ is a modulation-specific parameter [26]. From (3) and (8), the second term in (13) is calculated as shown in (15) at the top of the next page. When comparing (14) with (15), we can find that if the action is zero, the conditional SER $P^{(a)}\left(\tilde{x} \neq x \middle| (H_{rd}, H_{sd}, Q_b, Q_e) = (j,k,b,l), \hat{x} = x\right)$ is degenerated to $P\left(\tilde{x} \neq x \middle| H_{sd} = k\right)$, since only the direct link is active in this case.

G. Optimization of Relay Transmission Policy

Define $\pi(s) : \mathcal{S} \rightarrow \mathcal{A}_{m,b}$ as the policy that specifies the relay transmission action at the states. The long-term expected reward in an infinite horizon is formulated as

$$V_\pi(s_0) = \mathbb{E}_\pi\left[\sum\nolimits_{k=0}^{\infty} \lambda^k R^{(\pi(s_k))}(s_k)\right], \quad s_k \in \mathcal{S}, \;\; \pi(s_k) \in \mathcal{A}_{m,b}\,, \qquad (16)$$

where $V_\pi(s_0)$ is the long-term expected reward starting from the initial state s_0 and following the policy π from then on, and $0 \leq \lambda < 1$ is a discount factor. The policy that can maximize the long-term expected reward is referred to as the optimal policy, i.e., $\pi^\star = \arg\max\limits_{\pi} V_\pi(s_0)$. By assuming that the states of the Markov chain are recurrent, and thereby, unrelated to the initial state, the optimal policy can be found through the Bellman's equation [27]:

$$V_{\pi^\star}(s) = \max_{a \in \mathcal{A}_{m,b}}\left(R^{(a)}(s) + \lambda \sum_{s' \in S} P_a\left(s' \middle| s\right) V_{\pi^\star}(s')\right), \quad s \in \mathcal{S}\,, \qquad (17)$$

which can be implemented by a value iteration algorithm as follows:

$$V_{n+1}^{(a)}(s) = R^{(a)}(s) + \lambda \sum_{s' \in \mathcal{S}} P_a\left(s' \middle| s\right) V_n(s'), \quad s \in \mathcal{S}, \;\; a \in \mathcal{A}_{m,b}\,; \qquad (18)$$

$$V_{n+1}(s) = \max_{a \in \mathcal{A}_{m,b}}\left\{V_{n+1}^{(a)}(s)\right\}, \quad s \in \mathcal{S}\,, \qquad (19)$$

where the state transition probability, $P_a\left(s' \middle| s\right)$, is given by

$$\begin{aligned} &P_a\left(s' = (j',k',b',l',m') \middle| s = (j,k,b,l,m)\right) \\ &= P\left(H_{rd} = j' \middle| H_{rd} = j\right) P\left(H_{sd} = k' \middle| H_{sd} = k\right) \\ &\quad \cdot P\left(Q_e = l' \middle| Q_e = l\right) P\left(Q_c = m' \middle| Q_c = m\right) \\ &\quad \cdot P_a\left(Q_b = b' \middle| Q_b = b, Q_e = l\right). \qquad (20) \end{aligned}$$

Without loss of generality, the value of $V_0(s)$ in (18) can be initialized as zero. For the purpose of simple notations, an expectation form for the summation term in (18) will be used in the subsequent sections by applying (12) and (20) and making changes of variables:

$$\begin{aligned} &\sum_{s' \in \mathcal{S}} P_a\left(s' \middle| s\right) V_n(s') \qquad (21) \\ &= \sum_{j',k',l',m'} P\left(H_{rd} = j' \middle| H_{rd} = j\right) P\left(H_{sd} = k' \middle| H_{sd} = k\right) \\ &\quad \cdot P\left(Q_e = l' \middle| Q_e = l\right) P\left(Q_c = m' \middle| Q_c = m\right) \\ &\quad \cdot \sum_{w=0}^{\infty} P\left(E = w \middle| Q_e = l\right) \\ &\qquad \cdot V_n\left(s = (j', k', \min(b - a + w, N_b - 1), l', m')\right) \\ &\triangleq \mathbb{E}_{j,k,l,m}\left[V_n\left(s = (j', k', \min(b - a + w, N_b - 1), l', m')\right)\right]. \end{aligned}$$

H. Main Structure Results of Optimal Relay Transmission Policy

Some important properties of the optimal relay transmission policy are discussed in this subsection. These fundamental results are helpful when we analyze the SER performance in the next section, even though some of these have been explored in various MDP problems. First, we point out that when the battery of the relay contains more residual energy, the cooperative transmission has a larger value of $V_n(s)$, making less contribution to the overall SER performance degradation.

Theorem 1: For a fixed channel, energy harvesting and decoding state $(H_{rd}, H_{sd}, Q_e, Q_c) = (j, k, l, m)$, at the n^{th} iteration, we have $V_n(s = (j,k,b,l,m)) \geq V_n(s = (j,k,b',l,m))$, for $b \geq b'$.

Proof: To prove this theorem, we need to first show that

$$V_n^{(a)}(s = (j,k,b,l,m)) \geq V_n^{(a)}(s = (j,k,b-1,l,m))\,, \qquad (22)$$

for any $a \in \mathcal{A}_{m,b-1}$ and $b \geq 1$. This result can be proved by induction. When $n = 1$, we get $V_1^{(a)}(s = (j,k,b,l,m)) = V_1^{(a)}(s = (j,k,b-1,l,m)) = R^{(a)}(s = (j,k,b,l,m))$ because $V_0(s = (j,k,b,l,m)) = 0$ and the reward function only depends on the channel state and the relay action. Assuming that $n = i$ holds for any $j \in \mathcal{S}_{rd}$, $k \in \mathcal{S}_{sd}$,

$$P^{(a)}\left(\tilde{x}\neq x\right|\left(H_{rd},H_{sd},Q_b,Q_e\right)=\left(j,k,b,l\right),\hat{x}=x\right)\triangleq\frac{P^{(a)}\left(\left(H_{rd},H_{sd}\right)=\left(j,k\right),\tilde{x}\neq x\right|\hat{x}=x\right)}{P\left(H_{rd}=j\right)P\left(H_{sd}=k\right)}$$
$$=\frac{1}{P\left(H_{rd}=j\right)P\left(H_{sd}=k\right)}\cdot\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\int_{\Gamma_{rd}^{(j)}}^{\Gamma_{rd}^{(j+1)}}\int_{\Gamma_{sd}^{(k)}}^{\Gamma_{sd}^{(k+1)}}\exp\left(-\frac{c_M\left(P_s\zeta_{sd}+aG\zeta_{rd}\right)}{N_0\sin^2\theta}\right)\cdot p\left(\zeta_{sd}\right)p\left(\zeta_{rd}\right)d\zeta_{sd}d\zeta_{rd}d\theta$$
$$=\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\frac{g_{sd}\left(\theta,\Gamma_{sd}^{(k)}\right)-g_{sd}\left(\theta,\Gamma_{sd}^{(k+1)}\right)}{\left(\frac{c_MP_s\eta_{sd}}{N_0\sin^2\theta}+1\right)\left(\exp\left(-\frac{\Gamma_{sd}^{(k)}}{\eta_{sd}}\right)-\exp\left(-\frac{\Gamma_{sd}^{(k+1)}}{\eta_{sd}}\right)\right)}$$
$$\cdot\frac{g_{rd}\left(\theta,\Gamma_{rd}^{(j)}\right)-g_{rd}\left(\theta,\Gamma_{rd}^{(j+1)}\right)}{\left(a\frac{c_MG\eta_{rd}}{N_0\sin^2\theta}+1\right)\left(\exp\left(-\frac{\Gamma_{rd}^{(j)}}{\eta_{rd}}\right)-\exp\left(-\frac{\Gamma_{rd}^{(j+1)}}{\eta_{rd}}\right)\right)}d\theta\,. \tag{14}$$

$$P\left(\tilde{x}\neq x\right|H_{sd}=k\right)=\frac{1}{P\left(H_{sd}=k\right)}\cdot\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\int_{\Gamma_{sd}^{(k)}}^{\Gamma_{sd}^{(k+1)}}\exp\left(-\frac{c_MP_s\zeta_{sd}}{N_0\sin^2\theta}\right)p\left(\zeta_{sd}\right)d\zeta_{sd}d\theta$$
$$=\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\frac{g_{sd}\left(\theta,\Gamma_{sd}^{(k)}\right)-g_{sd}\left(\theta,\Gamma_{sd}^{(k+1)}\right)}{\left(\frac{c_MP_s\eta_{sd}}{N_0\sin^2\theta}+1\right)\left(\exp\left(-\frac{\Gamma_{sd}^{(k)}}{\eta_{sd}}\right)-\exp\left(-\frac{\Gamma_{sd}^{(k+1)}}{\eta_{sd}}\right)\right)}d\theta\,. \tag{15}$$

$l\in\mathcal{Q}_e$ and $m\in\mathcal{Q}_c$, we have $V_i^{(a)}\left(s=(j,k,b-1,l,m)\right)\leq V_i^{(a)}\left(s=(j,k,b,l,m)\right)$, for all $b\in\mathcal{Q}_b\backslash\{0\}$. It implies that

$$\begin{aligned}V_i\left(s=(j,k,b,l,m)\right)&=\max_{a\in\mathcal{A}_{m,b}}\left\{V_i^{(a)}\left(s=(j,k,b,l,m)\right)\right\}\\&\geq\max_{a\in\mathcal{A}_{m,b-1}}\left\{V_i^{(a)}\left(s=(j,k,b-1,l,m)\right)\right\}\\&=V_i\left(s=(j,k,b-1,l,m)\right)\,.\end{aligned}\tag{23}$$

Using (18) and (21), we then prove that for $n=i+1$:

$$\begin{aligned}&V_{i+1}^{(a)}\left(s=(j,k,b,l,m)\right)-V_{i+1}^{(a)}\left(s=(j,k,b-1,l,m)\right)=\\&\lambda\cdot\mathbb{E}_{j,k,l,m}\left[V_i\left(s=\left(j',k',\min\left(b-a+w,N_b-1\right),l',m'\right)\right)\right.\\&\quad\left.-V_i\left(s=\left(j',k',\min\left(b-1-a+w,N_b-1\right),l',m'\right)\right)\right]\\&\geq0\,.\end{aligned}\tag{24}$$

Similar to (23), it concludes that $V_{i+1}\left(s=(j,k,b,l,m)\right)\geq V_{i+1}\left(s=(j,k,b-1,l,m)\right)$, for $b\geq1$. ■

The simplicity of a structural policy makes it attractive for hardware implementation in power-hungry energy harvesting relay nodes. Typically, two types of structures are discussed for the optimal policy in the literature, and they are defined in the following [12], [27].

Definition 2: A policy is called a threshold-type policy in the battery states with the threshold $\varepsilon\left(j,k,l,m\right)$, if

$$\pi\left(s=(j,k,b,l,m)\right)\begin{cases}=0, & b\leq\varepsilon\left(j,k,l,m\right)\,;\\\neq0, & \text{otherwise}\,,\end{cases}\tag{25}$$

for any fixed $j\in\mathcal{S}_{rd}$, $k\in\mathcal{S}_{sd}$, $l\in\mathcal{Q}_e$ and $m\in\mathcal{Q}_c$.

Definition 3: A policy is called a monotonic-type policy in the battery states, if

$$\pi\left(s=(j,k,b-1,l,m)\right)\leq\pi\left(s=(j,k,b,l,m)\right)\,,\tag{26}$$

for any fixed $j\in\mathcal{S}_{rd}$, $k\in\mathcal{S}_{sd}$, $l\in\mathcal{Q}_e$ and $m\in\mathcal{Q}_c$.

When the allowable relay action is binary, i.e., $N_a=2$, one can easily prove that the optimal relay transmission policy follows an elegant threshold structure along with the direction of the battery states, where the relay node helps forward the signals only when its instantaneous battery state is above a threshold under given channel, energy harvesting and decoding states, and the following theorem is given.

Theorem 2: For $N_a=2$, the optimal relay transmission policy is a threshold-type policy (or equivalently, a monotonic-type policy in this special case).

Proof: This can be proved by showing that $V_n^{(1)}\left(s=(j,k,b,l,m)\right)-V_n^{(0)}\left(s=(j,k,b,l,m)\right)$ is a non-decreasing function in $b\in\mathcal{Q}_b$ via the induction method, and the details can be referred to a similar proof for Theorem 2 in [12]. ■

While the relay action is not limited to a binary case, the optimal policy could exhibit a monotonic structure in the battery states. A common method to assess the existence of such a monotonic-type structure is to check whether the function $V_n^{(a)}\left(s=(j,k,b,l,m)\right)$ is supermodular in a and b or not, i.e., $V_n^{(a)}\left(s=(j,k,b,l,m)\right)-V_n^{(a-1)}\left(s=(j,k,b,l,m)\right)\geq V_n^{(a)}\left(s=(j,k,b-1,l,m)\right)-V_n^{(a-1)}\left(s=(j,k,b-1,l,m)\right)$. In fact, the existence of the structure heavily relies on the reward functions and the state transition probabilities, and it is hard to directly verify the supermodularity of the function $V_n^{(a)}\left(s=(j,k,b,l,m)\right)$. Instead, a sufficient condition in terms of the energy quantum harvesting probability, $P\left(E=w\right|Q_e=l\right)$, is provided in [12] for ensuring the supermodularity in point-to-point communications. The result can be straightforwardly extended from the point-to-point communications to the considered cooperative communications in this paper.

IV. Performance Analysis

A. Exact SER Expressions

To evaluate the performance of the optimal relay transmission policy, the exact SER of the energy harvesting cooperative communications is analyzed by calculating the stationary state probabilities of the MDP. Consider an optimal policy $\pi^\star(s)$, and denote $\mathbf{p}$ as the corresponding stationary state probability vector of the MDP whose $\left(m\left(N_bN_{rd}N_{sd}N_e\right)+l\left(N_bN_{rd}N_{sd}\right)+k\left(N_bN_{rd}\right)+jN_b+b\right)^{th}$ entry, $p_{j,k,b,l,m}$, stands for the stationary probability of the state $(H_{rd},H_{sd},Q_b,Q_e,Q_c)=(j,k,b,l,m)$. In addition, let $\mathbf{M}_{j,k,l,m}$ be an $N_b\times N_b$ battery state transition probability matrix at the state $(H_{rd},H_{sd},Q_e,Q_c)=(j,k,l,m)$ with respect to the policy $\pi^\star(s)$, and the matrix is specified as

$$\begin{aligned}&\left[\mathbf{M}_{j,k,l,m}\right]_{b',b} \qquad (27)\\&=\begin{cases}P\left(E=b'-b+\pi^\star(s)\middle|\,Q_e=l\right), & b-\pi^\star(s)\le b'\le N_b-2\,;\\ 0,\ \ b'\le b-\pi^\star(s)-1\,; & \\ 1-\sum_{n=b-\pi^\star(s)}^{N_b-2}P\left(E=n-b+\pi^\star(s)\middle|\,Q_e=l\right), & b'=N_b-1\,,\end{cases}\end{aligned}$$

where $\left[\mathbf{M}_{j,k,l,m}\right]_{b',b}$ is the transition probability from the battery state b to b', corresponding to the optimal policy. Thus, the stationary probabilities are computed by solving the balance equation as follows:

$$\begin{bmatrix}\mathbf{M}\\ \mathbf{1}^T\end{bmatrix}\mathbf{p}=\begin{bmatrix}\mathbf{p}\\ 1\end{bmatrix}, \qquad (28)$$

where $\mathbf{1}$ is an all-one column vector, and $\mathbf{M}$ is an entire MDP state transition probability matrix, the $\left(m'\left(N_{rd}N_{sd}N_e\right)+l'\left(N_{rd}N_{sd}\right)+k'\left(N_{rd}\right)+j',m\left(N_{rd}N_{sd}N_e\right)+l\left(N_{rd}N_{sd}\right)+k\left(N_{rd}\right)+j\right)^{th}$ sub-matrix of which is given as

$$\begin{aligned}&P\left(H_{rd}=j'\middle|\,H_{rd}=j\right)\cdot P\left(H_{sd}=k'\middle|\,H_{sd}=k\right) \qquad (29)\\&\cdot P\left(Q_e=l'\middle|\,Q_e=l\right)\cdot P\left(Q_c=m'\middle|\,Q_c=m\right)\cdot\mathbf{M}_{j,k,l,m}\,.\end{aligned}$$

From (13) and (28), the exact SER is expressed as

$$\begin{aligned}P_{M,exact}=&\sum_{j=0}^{N_{rd}-1}\sum_{k=0}^{N_{sd}-1}\sum_{b=0}^{N_b-1}\sum_{l=0}^{N_e-1}\sum_{m=0}^{1}p_{j,k,b,l,m}\\&\cdot P^{(\pi^\star(s))}\left(\tilde{x}\neq x\middle|\,s=(j,k,b,l,m)\right)\,. \qquad (30)\end{aligned}$$

B. Asymptotic Approximations and Bounds for SER Expressions

Here we first analyze the SER expression at asymptotically high SNR. For simplicity of notation, we denote the stationary state probability of the zeroth channel state for the channel link x in (8) as μ_x, where x could be "rd" or "sd". Let ς_a be the occurrence probability for the action a at the zeroth states of the SD and RD channels, and it is defined as $\varsigma_a=\sum_{m=0}^{1}\sum_{s\in\Omega_{m,a}}p_s$, where $\mathbf{\Omega}_{m,a}=\{s=(0,0,b,l,m)|\,\pi^\star(s)=a,b\in\mathcal{Q}_b,l\in\mathcal{Q}_e\}$. The following theorem is given.

Theorem 3: The asymptotic SER is expressed as

$$\begin{aligned}P_{M,asym.}\approx&\varsigma_0\frac{K_0^{(0)}}{\mu_{sd}c_M\eta_{sd}}\left(\frac{P_s}{N_0}\right)^{-1} \qquad (31)\\&+\sum_{a\neq0}\frac{\varsigma_a}{a}\frac{K_1^{(0,0)}}{\mu_{sd}\mu_{rd}c_M^2\eta_{sd}\eta_{rd}}\left(\frac{P_s}{N_0}\right)^{-1}\left(\frac{G}{N_0}\right)^{-1},\end{aligned}$$

where $K_0^{(0)}=\frac{M-1}{2M}+\frac{\sin(2\pi/M)}{4\pi}$ and $K_1^{(0,0)}=\frac{3(M-1)}{8M}+\frac{\sin(2\pi/M)}{4\pi}-\frac{\sin(4\pi/M)}{32\pi}$.

Proof: When the SNR values for the RD and SD channel links are sufficiently large, or equivalently $P_s\eta_{sd}/N_0\gg1$ and $G\eta_{rd}/N_0\gg1$, we can write

$$\frac{c_MP_s\eta_{sd}}{N_0\sin^2\theta}+1\approx\frac{c_MP_s\eta_{sd}}{N_0\sin^2\theta}\,; \qquad (32)$$

$$a\frac{c_MG\eta_{rd}}{N_0\sin^2\theta}+1\approx a\frac{c_MG\eta_{rd}}{N_0\sin^2\theta},\ \text{for } a\neq0\,. \qquad (33)$$

By applying (32) and (33), the conditional SER in (14) with cooperation is approximated as

$$\begin{aligned}&P^{(a)}\left(\tilde{x}\neq x\middle|\,(H_{rd},H_{sd},Q_b,Q_e)=(j,k,b,l)\,,\hat{x}=x\right) \qquad (34)\\&\approx\begin{cases}\frac{1}{\left(\exp\left(-\frac{\Gamma_{sd}^{(k)}}{\eta_{sd}}\right)-\exp\left(-\frac{\Gamma_{sd}^{(k+1)}}{\eta_{sd}}\right)\right)}\frac{N_0K_0^{(k)}}{c_MP_s\eta_{sd}}, & a=0\,;\\ \frac{1}{\left(\exp\left(-\frac{\Gamma_{sd}^{(k)}}{\eta_{sd}}\right)-\exp\left(-\frac{\Gamma_{sd}^{(k+1)}}{\eta_{sd}}\right)\right)}\\ \cdot\frac{1}{\left(\exp\left(-\frac{\Gamma_{rd}^{(j)}}{\eta_{rd}}\right)-\exp\left(-\frac{\Gamma_{rd}^{(j+1)}}{\eta_{rd}}\right)\right)}\frac{N_0^2K_1^{(k,j)}}{ac_M^2P_sG\eta_{sd}\eta_{rd}}, & a\neq0\,,\end{cases}\end{aligned}$$

where $K_0^{(k)}$ and $K_1^{(k,j)}$ are defined as

$$\begin{aligned}K_0^{(k)}=&\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\sin^2\theta\left(\exp\left(-\frac{c_MP_s\Gamma_{sd}^{(k)}}{N_0\sin^2\theta}\right)\right.\\&\left.-\exp\left(-\frac{c_MP_s\Gamma_{sd}^{(k+1)}}{N_0\sin^2\theta}\right)\right)d\theta\,; \qquad (35)\end{aligned}$$

$$\begin{aligned}K_1^{(k,j)}=&\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\sin^4\theta\left(\exp\left(-\frac{c_MP_s\Gamma_{sd}^{(k)}}{N_0\sin^2\theta}\right)\right.\\&\left.-\exp\left(-\frac{c_MP_s\Gamma_{sd}^{(k+1)}}{N_0\sin^2\theta}\right)\right)\\&\cdot\left(\exp\left(-a\frac{c_MG\Gamma_{rd}^{(j)}}{N_0\sin^2\theta}\right)-\exp\left(-a\frac{c_MG\Gamma_{rd}^{(j+1)}}{N_0\sin^2\theta}\right)\right)d\theta\,. \qquad (36)\end{aligned}$$

We can further simplify the factor $K_0^{(k)}$ as follows:

$$\begin{aligned}K_0^{(k)}=&\frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}}\sin^2\theta\exp\left(-\frac{c_MP_s\Gamma_{sd}^{(k)}}{N_0\sin^2\theta}\right)\\&\cdot\left(1-\exp\left(-\frac{c_MP_s\left(\Gamma_{sd}^{(k+1)}-\Gamma_{sd}^{(k)}\right)}{N_0\sin^2\theta}\right)\right)d\theta\,,\end{aligned}$$

$$\approx \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \sin^2\theta \exp\left(-\frac{c_M P_s \Gamma_{sd}^{(k)}}{N_0 \sin^2\theta}\right) d\theta\,,$$

$$\approx \begin{cases} \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \sin^2\theta d\theta = \frac{M-1}{2M} + \frac{\sin(2\pi/M)}{4\pi}, & k=0\,; \\ 0, \ \ \text{otherwise}\,, \end{cases} \quad (37)$$

where the first and second approximations work well, if $P_s/N_0 \gg 1$ and the channel is quantized rationally, i.e., $\Gamma_{sd}^{(k+1)} - \Gamma_{sd}^{(k)} \geq \varepsilon$ for a sufficiently large ε. Furthermore, by assuming that $G/N_0 \gg 1$, the factor $K_1^{(k,j)}$ can be approximated in a similar way:

$$K_1^{(k,j)} \approx \begin{cases} \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \sin^4\theta d\theta = \frac{3(M-1)}{8M} + \frac{\sin(2\pi/M)}{4\pi} \\ \qquad - \frac{\sin(4\pi/M)}{32\pi}, \ \ k=0, \ \ j=0\,; \\ 0, \ \ \text{otherwise}\,. \end{cases} \quad (38)$$

Likewise, an asymptotic approximation to the conditional SER of the SD link can be derived as

$$P\left(\tilde{x} \neq x \middle| H_{sd}=k\right) \approx \frac{1}{\left(\exp\left(-\frac{\Gamma_{sd}^{(k)}}{\eta_{sd}}\right) - \exp\left(-\frac{\Gamma_{sd}^{(k+1)}}{\eta_{sd}}\right)\right)} \frac{N_0 K_0^{(k)}}{c_M P_s \eta_{sd}}\,. \quad (39)$$

From (13), (34) and (39), it is concluded that the asymptotic SER for (30) only depends upon how the relay node performs its actions, to forward or not to forward, when the SD and RD channels are both stayed at the zeroth states. As a result, we get the asymptotic SER in (31). ■

It is noted that the asymptotic SER in Theorem 3 is tight at reasonably high SNR. In fact, it can be verified from the above analysis that the asymptotic SER is also an upper bound for the exact SER in (30), if P_s/N_0 and G/N_0 are sufficiently large. The aforementioned analysis clearly indicates that the asymptotic SER is affected by the occurrence probability ς_a. Below we state a theorem regarding the property of the occurrence probability.

Theorem 4: The sum of the occurrence probabilities ς_a is equal to $\sum_{a=0}^{N_a-1}\varsigma_a = \mu_{sd}\mu_{rd}$. Furthermore, the inactive probability $\varsigma_0 = \mu_{sd}\mu_{rd}P\left(Q_c=0\right) + \rho \geq \mu_{sd}\mu_{rd}P\left(Q_c=0\right)$ and the active probability $\sum_{a=1}^{N_a-1}\varsigma_a = \mu_{sd}\mu_{rd}P\left(Q_c=1\right) - \rho \leq \mu_{sd}\mu_{rd}P\left(Q_c=1\right)$, where $\rho = \sum_{s\in\Omega_{1,0}} p_s \geq 0$ is the stationary state probability at the zeroth channel states for which the decoding is successful but the relay keeps silent.

Proof: From the balance equation in (28), the stationary state probability satisfies

$$p_{j',k',b',l',m'} = \sum_{j=0}^{N_{rd}-1}\sum_{k=0}^{N_{sd}-1}\sum_{b=0}^{N_b-1}\sum_{l=0}^{N_e-1}\sum_{m=0}^{1} P_{\pi^\star(s)}\left(s=(j',k',b',l',m')\middle| s=(j,k,b,l,m)\right)\cdot p_{j,k,b,l,m}\,. \quad (40)$$

Define $\widehat{p}_{j,k,m} = \sum_{b=0}^{N_b-1}\sum_{l=0}^{N_e-1} p_{j,k,b,l,m}$. Using (20) and taking summation over the indices b' and l' on the both sides of the equality in (40), we can get

$$\widehat{p}_{j',k',m'} = \sum_{j=0}^{N_{rd}-1}\sum_{k=0}^{N_{sd}-1}\sum_{m=0}^{1} P\left(H_{rd}=j'\middle| H_{rd}=j\right) \cdot P\left(H_{sd}=k'\middle| H_{sd}=k\right) P\left(Q_c=m'\middle| Q_c=m\right)\widehat{p}_{j,k,m}\,, \quad (41)$$

because $\sum_{l'=0}^{N_e-1} P\left(Q_e=l'\middle| Q_e=l\right)=1$ and $\sum_{b'=0}^{N_b-1} P_{\pi^\star(s)}\left(Q_b=b'\middle| Q_b=b, Q_e=l\right)=1$. Then, it is concluded from (41) that $\widehat{p}_{j,k,m}$ is the stationary state probability for the state $(H_{rd},H_{sd},Q_c)=(j,k,m)$, and it implies that $\widehat{p}_{0,0,m} = \mu_{sd}\mu_{rd}P\left(Q_c=m\right)$ because the transitions of the channel and decoding states are independent of each other. We can therefore obtain $\sum_{a=0}^{N_a-1}\varsigma_a = \widehat{p}_{0,0,0} + \widehat{p}_{0,0,1} = \mu_{sd}\mu_{rd}$. Since the action $a=0$ is the only action for the relay when the decoding fails, it can be further shown that $\varsigma_0 = \widehat{p}_{0,0,0} + \rho \geq \widehat{p}_{0,0,0}$ and $\sum_{a=1}^{N_a-1}\varsigma_a = \widehat{p}_{0,0,1} - \rho \leq \widehat{p}_{0,0,1}$. ■

In fact, the stationary state probability ρ plays an important role in the achievable performance, and for a given policy, the probability mainly depends on the energy harvesting capability of the relay node. To get more insight into the performance behavior, an upper bound and a lower bound for the asymptotic SER are provided in the following theorem.

Theorem 5: Assume $G = c_r P_s$. For sufficiently high SNR, the asymptotic SER is upper and lower bounded by

$$\Phi\left(N_a-1\right) \leq P_{M,asym.} \leq \Phi\left(1\right)\,, \quad (42)$$

where the function $\Phi\left(z\right)$ is defined as

$$\Phi\left(z\right) = \frac{\rho K_0^{(0)}}{c_M \mu_{sd}\eta_{sd}}\left(\frac{P_s}{N_0}\right)^{-1} + \left(\frac{\mu_{rd}\gamma K_0^{(0)}}{c_M \eta_{sr}\eta_{sd}} + \frac{\left(\mu_{rd}\mu_{sd}-\rho\right)K_1^{(0,0)}}{z c_M^2 c_r \mu_{rd}\mu_{sd}\eta_{rd}\eta_{sd}}\right)\left(\frac{P_s}{N_0}\right)^{-2} - \frac{\gamma K_1^{(0,0)}}{z c_M^2 c_r \eta_{sr}\eta_{rd}\eta_{sd}}\left(\frac{P_s}{N_0}\right)^{-3}\,. \quad (43)$$

Proof: First, the weighted sum of the active probability $\sum_{a\neq 0}\frac{\varsigma_a}{a}$ in (31) is bounded by $\frac{1}{N_a-1}\sum_{a\neq 0}\varsigma_a \leq \sum_{a\neq 0}\frac{\varsigma_a}{a} \leq \sum_{a\neq 0}\varsigma_a$. Moreover, by using $e^x \approx 1+x$ for $x \ll 1$, when $\frac{P_s\eta_{sr}}{N_0} \gg \gamma$, the successful decoding probability in (11) can be approximated as $P\left(Q_c=1\right) = 1 - \frac{\gamma N_0}{P_s\eta_{sr}}$. Let $G = c_r P_s$. By applying the above results and Theorem 4 into (31), we can obtain the upper bound and the lower bound for the asymptotic SER as shown in (42). ■

It is noted from Theorem 5 that the equality in (42) holds only when two kinds of actions, either keeping silent or transmitting with a constant power level G, are accessible to the relay node. In this case, the bounds are tight and $P_{M,asym.} = \Phi\left(1\right)$. Another condition for which the lower and upper bounds get closer at extremely high SNR values of $\frac{P_s}{N_0}$ is when the stationary state probability ρ is not equal to zero. In this case, the function $\Phi(z)$ is dominated by the first term in (43).

C. Diversity Order and Energy Harvesting Gain

The main idea behind cooperative communications is to form a virtual multiple-input multiple-output (MIMO) system via separated single-antenna nodes, and it is interesting to investigate the diversity order in such an energy-limited relay network when the SNRs of the three channel links go to infinity. In addition to the diversity order, the energy harvesting gain is another important metric to characterize the performance of energy harvesting cooperative networks.

Theorem 6: If $\rho > 0$, the diversity order and the energy harvesting gain of the cooperative communications are respectively given as $d = 1$ and $g_E = \frac{c_M \mu_{sd} \eta_{sd}}{\rho K_0^{(0)}}$. When $\rho = 0$, a full diversity $d = 2$ is achieved and the energy harvesting gain is bounded by $\sqrt{\frac{c_M^2 c_r \eta_{sr} \eta_{rd} \eta_{sd}}{c_M c_r \mu_{rd} \eta_{rd} \gamma K_0^{(0)} + \eta_{sr} K_1^{(0,0)}}} \le g_E \le \sqrt{\frac{(N_a-1) c_M^2 c_r \eta_{sr} \eta_{rd} \eta_{sd}}{(N_a-1) c_M c_r \mu_{rd} \eta_{rd} \gamma K_0^{(0)} + \eta_{sr} K_1^{(0,0)}}}$.

Proof: From Theorem 5, if $\rho > 0$, the SER for sufficiently high SNR is dominated by the first term in (43), and it can be approximated as $P_{M,asym.} \approx \frac{\rho K_0^{(0)}}{c_M \mu_{sd} \eta_{sd}} \left(\frac{P_s}{N_0}\right)^{-1}$ when $\frac{P_s}{N_0} \to \infty$. Therefore, the diversity order is one and the energy harvesting gain is $g_E = \frac{c_M \mu_{sd} \eta_{sd}}{\rho K_0^{(0)}}$. As $\rho = 0$, the SER is dominated by the second term in (43) when $\frac{P_s}{N_0} \to \infty$, resulting in a diversity order of two, and the energy harvesting gain is lower and upper bounded by $\sqrt{\frac{c_M^2 c_r \eta_{sr} \eta_{rd} \eta_{sd}}{c_M c_r \mu_{rd} \eta_{rd} \gamma K_0^{(0)} + \eta_{sr} K_1^{(0,0)}}} \le g_E \le \sqrt{\frac{(N_a-1) c_M^2 c_r \eta_{sr} \eta_{rd} \eta_{sd}}{(N_a-1) c_M c_r \mu_{rd} \eta_{rd} \gamma K_0^{(0)} + \eta_{sr} K_1^{(0,0)}}}$. ■

The aforementioned discussions raise an interesting question on the prospects for achieving the full diversity order of two. In fact, whether the full diversity order is achievable directly depends on the energy quantum harvesting probability $P\left(E = w \,|\, Q_e = l\right)$ and the optimal policy at asymptotically high SNR. Before introducing the condition of the energy quantum supporting way for achieving $d = 2$, we specify the optimal cooperative transmission policy at asymptotically high SNR with the following theorem.

Theorem 7: When $\frac{G}{N_0} \to \infty$, the optimal cooperative transmission policy $\pi^\star(s)$ at asymptotically high SNR is a threshold-type policy, which is given by

$$\varepsilon\left(j,k,l,m\right) = \begin{cases} 0, & k=0, \quad m=1\,; \\ N_b - 1, & \text{otherwise}\,. \end{cases} \tag{44}$$

Proof: From (34) and (37)-(39), when $\frac{G}{N_0} \to \infty$, the reward function (13) at asymptotically high SNR becomes

$$R^{(a)}\left(s=(j,k,b,l,m)\right) = \begin{cases} 1 - \frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \\ \qquad\qquad a=0, \quad k=0\,; \\ 1, \quad \text{otherwise}\,. \end{cases} \tag{45}$$

The proof is then divided into three parts as follows.

(i) If $m = 0$, the threshold is given by $\varepsilon\left(j,k,l,m\right) = N_b - 1$ because the decoding is incorrect and the relay can only keep silent.

(ii) If $k \neq 0$ and $m = 1$, it can be shown from (18), (21) and (45) that for fixed j and l, we get

$$\begin{aligned} & V_{n+1}^{(0)}\left(s=(j,k,b,l,m)\right) - V_{n+1}^{(a)}\left(s=(j,k,b,l,m)\right) = \\ & \lambda \cdot \left(\mathbb{E}_{j,k,l,m}\left[V_n\left(s=(j',k',\min\left(b+w,N_b-1\right),l',m'\right)\right)\right. \\ & \qquad \left.\left. - V_n\left(s=(j',k',\min\left(b-a+w,N_b-1\right),l',m'\right)\right)\right]\right) \\ & \ge 0\,, \end{aligned} \tag{46}$$

for $a \neq 0$, and the inequality is valid due to the result obtained in Theorem 1. Thus, the action $a = 0$ is optimal in this case.

(iii) We define $\Xi_n\left(s=(j,k,b,l,m)\right) = V_n^{(1)}\left(s=(j,k,b,l,m)\right) - V_n^{(0)}\left(s=(j,k,b,l,m)\right)$ and $\Delta_n^{(xy)}\left(s=(j,k,b,l,m)\right) = V_n^{(x)}\left(s=(j,k,b-1,l,m)\right) - V_n^{(y)}\left(s=(j,k,b,l,m)\right)$. For $k = 0$ and $m = 1$, it suffices to prove $\varepsilon\left(j,k,l,m\right) = 0$ by showing that $\Xi_n\left(s=(j,k,b,l,m)\right) > 0$ for all $b \neq 0$ and any n in the following. We first claim that at each value iteration n,

$$\Delta_n^{(00)}\left(s=(j',k',b',l',m')\right) > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \quad \text{for any } j', k', b', l' \text{ and } m'\,; \tag{47}$$

$$\Xi_n\left(s=(j,k,b,l,m)\right) > 0, \;\; b \neq 0\,, \tag{48}$$

which can be proved by induction in the following:

(a) Without loss of generality, we initialize $V_0\left(s=(j',k',b',l',m')\right) = 0$. For $n = 1$, by using (18) and (45), we get $\Delta_1^{(00)}\left(s=(j',k',b',l',m')\right) = 0 > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}$. In addition, we have $\Xi_1\left(s=(j,k,b,l,m)\right) = \frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}} > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}$.

(b) Assume $n = i$ holds. It then immediately implies from (48) that $V_i^{(1)}\left(s=(j,k,b,l,m)\right) > V_i^{(0)}\left(s=(j,k,b,l,m)\right)$, for $b \neq 0$. On the other hand, it can be derived from (45) and Theorem 1 that $V_i^{(1)}\left(s=(j,k,b,l,m)\right) \ge V_i^{(a)}\left(s=(j,k,b,l,m)\right)$, for $b \neq 0$ and $a \ge 2$, since $R^{(a)}\left(s=(j,k,b,l,m)\right) = 1$ when $a \neq 0$. In conclusion, we get $V_i\left(s=(j,k,b,l,m)\right) = \max_{a \in \mathcal{A}_{m,b}} \{ V_i^{(a)}\left(s=(j,k,b,l,m)\right)\} = V_i^{(1)}\left(s=(j,k,b,l,m)\right)$, for $b \neq 0$.

(c) For $n = i+1$, it is obtained from (18), (21) and (45) that

$$\begin{aligned} & \Delta_{i+1}^{(00)}\left(s=(j',k',b',l',m')\right) = \lambda \cdot (\mathbb{E}_{j',k',l',m'}[\\ & \quad V_i(s=(\tilde{j},\tilde{k},\min(b'-1+w,N_b-1),\tilde{l},\tilde{m})) \\ & \quad - V_i(s=(\tilde{j},\tilde{k},\min(b'+w,N_b-1),\tilde{l},\tilde{m}))])\,. \end{aligned} \tag{49}$$

From the optimal thresholds in the cases of (i) and (ii) and the discussions in (b), it yields

$$\begin{aligned} & V_i(s=(\tilde{j},\tilde{k},\tilde{b}-1,\tilde{l},\tilde{m})) - V_i(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) \\ & = \begin{cases} \Delta_i^{(00)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \\ \qquad\qquad \tilde{m}=0\,; \\ \Delta_i^{(00)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \\ \qquad\qquad \tilde{m}=1, \tilde{k} \neq 0\,; \\ \Delta_i^{(01)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) = -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \\ \qquad\qquad \tilde{m}=1, \tilde{k}=0, \tilde{b}=1\,; \\ \Delta_i^{(11)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}, \\ \qquad\qquad \tilde{m}=1, \tilde{k}=0, \tilde{b} \ge 2\,, \end{cases} \end{aligned} \tag{50}$$

where the first inequality and the second inequality are due to (47) at $n = i$, the third equality is calculated from $\Delta_i^{(01)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) = R^{(0)}(s=(\tilde{j},\tilde{k},\tilde{b}-1,\tilde{l},\tilde{m})) - R^{(1)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m}))$, and the last inequality comes from $\Delta_i^{(11)}(s=(\tilde{j},\tilde{k},\tilde{b},\tilde{l},\tilde{m})) = \Delta_i^{(00)}(s = (\tilde{j},\tilde{k},\tilde{b}-1,\tilde{l},\tilde{m}))$ and the claim of (47) at $n = i$. Using (49) and (50), we then have $\Delta_{i+1}^{(00)}(s=(j',k',b',l',m')) > -\frac{N_0 K_0^{(0)}}{\mu_{sd} c_M P_s \eta_{sd}}$. Furthermore, by definition, it leads to

$$\begin{aligned}
&\Xi_{i+1}(s=(j,k,b,l,m)) \quad (51)\\
&= V_{i+1}^{(1)}(s=(j,k,b,l,m)) - V_{i+1}^{(0)}(s=(j,k,b,l,m))\\
&= R^{(0)}(s=(j,k,b-1,l,m))\\
&\quad - R^{(0)}(s=(j,k,b-1,l,m))\\
&\quad + V_{i+1}^{(1)}(s=(j,k,b,l,m)) - V_{i+1}^{(0)}(s=(j,k,b,l,m))\\
&= R^{(1)}(s=(j,k,b,l,m)) - R^{(0)}(s=(j,k,b-1,l,m))\\
&\quad + \Delta_{i+1}^{(00)}(s=(j,k,b,l,m)) > 0\,,
\end{aligned}$$

for all $b \neq 0$. As a result, the optimal threshold in this case is given by $\varepsilon(j,k,l,m) = 0$.

■

This theorem gives an important insight into understanding the optimal cooperative transmission strategy for the energy harvesting relay node at asymptotically high SNR regimes. Under this circumstance, the relay node with non-empty battery spends its harvested energy only when the decoding is successful and the source node experiences the worst channel condition in its direct link. For a special case of $N_a = 2$, the optimal policy at asymptotically high SNR is simply given by

$$\pi^\star(s=(j,k,b,l,m)) = \begin{cases} 1, & k=0,\\ & \quad b\geq 1,\ \ m=1\,;\\ 0, & \text{otherwise}\,. \end{cases} \quad (52)$$

With the optimal cooperative transmission policy in Theorem 7, a theorem regarding the energy quantum supporting way for achieving the full diversity gain is provided in the following.

Theorem 8: The energy harvesting cooperative communications can achieve a diversity order of two, if and only if the energy quantum outage probability $P(E=0|\,Q_e=l) = 0$, for $l = 0,\ldots,N_e-1$.

Proof: From Theorem 6, it suffices to prove this theorem by showing that $\rho = 0$ if and only if $P(E=0|\,Q_e=l)=0$, for $l=0,\ldots,N_e-1$. By using (9), (10) and (20), the balance equation in (40) for the stationary state probability p_s at the state $s=(0,0,b',l',1)$ can be rewritten as

$$\begin{aligned}
p_{0,0,b',l',1} = &\sum_{j=0}^{1}\sum_{k=0}^{1}\sum_{l=0}^{N_e-1}\sum_{m=0}^{1} P(H_{rd}=0|\,H_{rd}=j)\\
&\cdot P(H_{sd}=0|\,H_{sd}=k)\,P(Q_e=l'|\,Q_e=l)\\
&\cdot P(Q_c=1|\,Q_c=m)\\
&\cdot \sum_{b=0}^{N_b-1} P_{\pi^\star(s)}(Q_b=b'|\,Q_b=b,Q_e=l)\,p_{j,k,b,l,m}\,. \quad (53)
\end{aligned}$$

For the *sufficiency* part, since $\pi^\star(s=(j,k,b,l,m)) \leq b$ and $P(E=0|\,Q_e=l) = 0$, for $l = 0,\ldots,N_e-1$, it is obtained from (12) that

$$\begin{aligned}
&P_{\pi^\star(s)}(Q_b=0|\,Q_b=b,Q_e=l)\\
&= P(E=-b+\pi^\star(s)|\,Q_e=l) = 0\,, \quad (54)
\end{aligned}$$

where the harvested energy quantum must be non-negative, i.e., $P(E<0|\,Q_e=l) = 0$. By substituting (54) into (53), we can get $p_{0,0,0,l',1} = 0$. According to the optimal policy at asymptotically high SNR in Theorem 7 and the definition of ρ in Theorem 4, it then implies $\rho = \sum_{s\in\Omega_{1,0}} p_s = \sum_{l'=0}^{N_e-1} p_{0,0,0,l',1} = 0$, where the second equality is attributed to the definition of $\mathbf{\Omega}_{m,a} = \{s=(0,0,b,l,m)|\,\pi^\star(s)=a, b\in\mathcal{Q}_b, l\in\mathcal{Q}_e\}$.

For the *necessity* part, the condition $\rho = 0$ with the optimal policy at asymptotically high SNR in Theorem 7 requires $p_{0,0,0,l',1} = 0$, for $l' = 0,\ldots,N_e-1$. From (53), the requirement of $p_{0,0,0,l',1} = 0$ implicitly indicates that $P_{\pi^\star(s)}(Q_b=0|\,Q_b=1,Q_e=l) = 0$ because the stationary state probability $p_{j,k,1,l,m}$ does not necessarily equal zero. Since $\pi^\star(s=(0,0,1,l,1)) = 1$ at sufficiently high SNR, we can get $P_{\pi^\star(s)}(Q_b=0|\,Q_b=1,Q_e=l) = P(E=0|\,Q_e=l) = 0$. ■

From Theorem 7 and Theorem 8, we know that any policy that obeys the threshold structure in (44) can achieve the same full diversity order but have different energy harvesting gains, if the energy quantum outage probability $P(E=0|\,Q_e=l) = 0$, for $l = 0,\ldots,N_e-1$. Under this circumstance, the inactive and active probabilities are simply given as $\varsigma_0 = \mu_{sd}\mu_{rd}P(Q_c=0)$ and $\sum_{a=1}^{N_a-1}\varsigma_a = \mu_{sd}\mu_{rd}P(Q_c=1)$ because $\rho=0$. By using Theorem 3 and $P(Q_c=1) = 1-\frac{\gamma N_0}{P_s\eta_{sr}}$ in the proof of Theorem 5, the asymptotic SER for the full-diversity achieving policy in (44) is therefore given as

$$\begin{aligned}
P_{M,asym.} = &\left(\frac{\mu_{rd}\gamma K_0^{(0)}}{c_M\eta_{sr}\eta_{sd}} + \left(\sum_{a=1}^{N_a-1}\frac{\varsigma_a}{a}\right)\right.\\
&\left.\cdot\frac{K_1^{(0,0)}}{c_M^2 c_r\mu_{rd}\mu_{sd}\eta_{rd}\eta_{sd}}\right)\left(\frac{P_s}{N_0}\right)^{-2}\\
\leq &\left(\frac{\mu_{rd}\gamma K_0^{(0)}}{c_M\eta_{sr}\eta_{sd}} + \frac{K_1^{(0,0)}}{c_M^2 c_r\eta_{rd}\eta_{sd}}\right)\left(\frac{P_s}{N_0}\right)^{-2}\\
&-\frac{\gamma K_1^{(0,0)}}{c_M^2 c_r\eta_{sr}\eta_{rd}\eta_{sd}}\left(\frac{P_s}{N_0}\right)^{-3}\,, \quad (55)
\end{aligned}$$

where the upper bound is obtained from (42) by setting ρ in $\Phi(1)$ as zero. It is worth noting that the above equality holds when the optimal policy for $N_a = 2$ in (52) is applied, and the corresponding energy harvesting gain is given by $\sqrt{\frac{c_M^2 c_r\eta_{sr}\eta_{rd}\eta_{sd}}{c_M c_r\mu_{rd}\eta_{rd}\gamma K_0^{(0)}+\eta_{sr}K_1^{(0,0)}}}$. From (55), to maximize the energy harvesting gain, one can appropriately design the non-zero power actions for the states $s = (j,0,b,l,1)$, for $b \geq 1$, by alternatively minimizing $\sum_{a=1}^{N_a-1}\frac{\varsigma_a}{a}$ subject to a sum probability constraint $\sum_{a=1}^{N_a-1}\varsigma_a = \mu_{sd}\mu_{rd}P(Q_c=1)$, which strikes a balance between the occurrence probability ς_a and the power scaling effect $\frac{1}{a}$. While the optimal policy for attaining the maximum diversity and energy harvesting gains can be acquired by the value iteration algorithm of the MDP, this analytical result suggests an aggressive way to spend the harvested energy for those states with non-zero power actions to obtain the better energy harvesting gain, since the power scaling effect usually dominates the occurrence probability.

TABLE I
ENERGY HARVESTING STATE TRANSITION PROBABILITY AND ENERGY QUANTUM HARVESTING PROBABILITY.

(a) State transition probability $P\left(Q_e = l' \middle| Q_e = l\right)$.

	$l'=0$	$l'=1$	$l'=2$	$l'=3$
$l=0$	0.979	0.015	0.006	0
$l=1$	0.005	0.988	0.007	0
$l=2$	0.006	0.009	0.975	0.010
$l=3$	0	0	0.007	0.993

(b) Energy quantum harvesting probability $P\left(E = w \middle| Q_e = l\right)$ ($\Omega = 8$ cm^2).

	$E=0$	$E=1$	$E=2$	$E=3$	$E=4$	$E=5$	$E=6$	$E\geq 7$
$Q_e=0$	0.087	0.455	0.384	0.058	0.001	3×10^{-6}	1×10^{-9}	0.015
$Q_e=1$	4×10^{-4}	0.015	0.144	0.402	0.343	0.089	0.006	1×10^{-4}
$Q_e=2$	2×10^{-5}	5×10^{-4}	0.006	0.039	0.141	0.276	0.295	0.241
$Q_e=3$	5×10^{-28}	2×10^{-20}	7×10^{-16}	1×10^{-10}	1×10^{-6}	0.001	0.061	0.937

(c) Energy quantum harvesting probability $P\left(E = w \middle| Q_e = l\right)$ ($\Omega = 4$ cm^2).

	$E=0$	$E=1$	$E=2$	$E=3$	$E=4$	$E=5$	$E=6$	$E\geq 7$
$Q_e=0$	0.315	0.640	0.030	2×10^{-6}	2×10^{-14}	2×10^{-26}	1×10^{-42}	0.015
$Q_e=1$	0.008	0.352	0.588	0.051	7×10^{-5}	5×10^{-10}	1×10^{-17}	2×10^{-5}
$Q_e=2$	3×10^{-4}	0.026	0.299	0.521	0.148	0.006	3×10^{-5}	2×10^{-6}
$Q_e=3$	1×10^{-20}	5×10^{-11}	5×10^{-4}	0.279	0.688	0.033	7×10^{-7}	3×10^{-16}

In addition, Theorem 8 raises an interesting question of how to reach the energy quantum supporting condition of $P\left(E = 0 \middle| Q_e = l\right) = 0$, for $l = 0, \ldots, N_e - 1$, in practice. We recall from [12] that the energy quantum outage probability $P\left(E = 0 \middle| Q_e = l\right)$ in the real solar-data-driven energy harvesting model is given as

$$\begin{aligned} &P\left(E = 0 \middle| Q_e = l\right) \\ &= \left(1 - \frac{\bar{\mu}_l}{E_U}\right) g_1\left(\bar{\mu}_l, \bar{\rho}_l\right) - g_2\left(\bar{\mu}_l, \bar{\rho}_l\right) , \end{aligned} \tag{56}$$

where $\bar{\mu}_l$ and $\bar{\rho}_l$ are two underlying parameters which represent the mean and the variance of the random harvested energy at the l^{th} energy harvesting state, respectively, and the functions $g_1\left(\bar{\mu}_l, \bar{\rho}_l\right)$ and $g_2\left(\bar{\mu}_l, \bar{\rho}_l\right)$ are defined as [12]

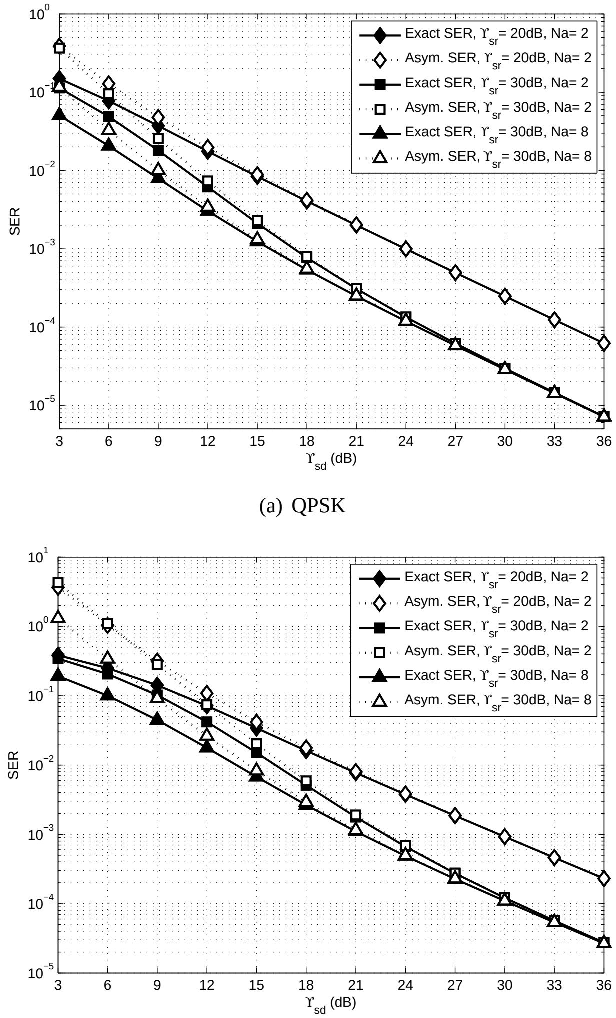
$$\begin{aligned} g_1\left(\bar{\mu}_l, \bar{\rho}_l\right) = \frac{1}{2} &\left(\operatorname{erfc}\left(\frac{-\bar{\mu}_l}{\sqrt{2\bar{\rho}_l}}\right) \right. \\ &\left. - \operatorname{erfc}\left(\frac{1}{\sqrt{2\bar{\rho}_l}}\left(E_U - \bar{\mu}_l\right)\right)\right) ; \end{aligned} \tag{57}$$

$$\begin{aligned} g_2\left(\bar{\mu}_l, \bar{\rho}_l\right) = \sqrt{\frac{\bar{\rho}_l}{2\pi E_U^2}} &\left(\exp\left(-\frac{\bar{\mu}_l^2}{2\bar{\rho}_l}\right) \right. \\ &\left. - \exp\left(-\frac{1}{2\bar{\rho}_l}\left(E_U - \bar{\mu}_l\right)^2\right)\right) . \end{aligned} \tag{58}$$

Notice that the larger the value $\bar{\mu}_l$, the better the energy harvesting condition, and the energy harvesting capability can be enhanced by increasing the values of some system parameters such as solar panel area, energy harvesting time duration and energy harvesting conversion efficiency, etc. Accordingly, we can find from (56)-(58) that the probability $P\left(E = 0 \middle| Q_e = l\right)$ approaches to zero when $E_U \ll \bar{\mu}_l$, and this can be achieved by either reducing the energy quantum size or improving the energy harvesting capability.

V. Numerical Results and Discussions

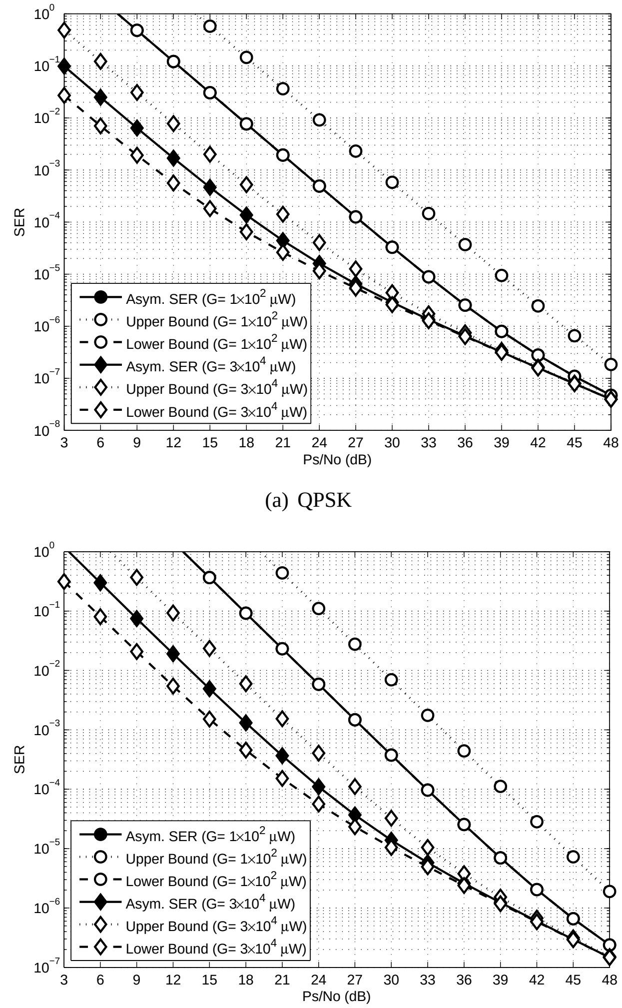
In this section, some numerical examples are demonstrated to substantiate the analytical derivations and theorems for the energy harvesting cooperative communications. The simulation parameters are set as follows. In the cooperative communications, the policy management period T_M is given as 300 sec, and the transmission power levels of the source node and the relay node (if active) are the same and equal to $P_s = G = 4 \times 10^4$ μW. Accordingly, the size of one energy quantum can be computed as $E_U = \frac{1}{2} G T_M = 6 \times 10^3$ mJ. The numbers of channel and battery states are set to three and eight, respectively. The channel quantization thresholds for the SD and the RD links are both randomly chosen as $\Gamma_{sd} = \Gamma_{rd} = \{0, 2.0, 3.0, \infty\}$, and it is assumed that the channels vary slowly with the normalized Doppler frequency $f_D = 5 \times 10^{-2}$. The decoding capability for the relay node is given by $\gamma = 15$ dB, which corresponds to a successful decoding probability of 0.95 at $\frac{P_s}{N_0} = 28$ dB. The discount factor in the MDP is set close to unity, given by $\lambda = 0.99$. The adopted modulation schemes are quadrature-phase-shift keying (QPSK) and 8PSK. The solar-data-driven energy harvesting model in [12] is utilized to capture the influence of parameter settings on the system performance such as the solar panel size Ω, the energy quantum size E_U, etc. The number of energy harvesting states is set as four, and the data record of the solar irradiance measured at the solar site in Elizabeth City State University in June from 2008 to 2010 is adopted for training the energy harvesting model in our simulation [28]. With the solar panel size $\Omega = 8$ cm^2 or $\Omega = 4$ cm^2 and the conversion efficiency for energy harvesting $\vartheta = 20$ %, the training results, including the energy harvesting state transition probability and the energy quantum harvesting probability, are listed in Table I. We note from Table I(b) and Table I(c) that the relay node has more opportunities to harvest a higher number of energy quanta when the panel size is expanded from 4 cm^2 to 8 cm^2. The above settings are used throughout the simulation, except as otherwise stated. When the value iteration algorithm is executed, the integrations in the reward functions (13)-(15) are carried out via a Riemann sum method.



(a) QPSK

(b) 8PSK

Fig. 2. Comparison of the exact SER and the asymptotic SER under various SNR values of Υ_{sr} for different modulation schemes: (a) QPSK and (b) 8PSK ($\Omega = 8$ cm^2).



(a) QPSK

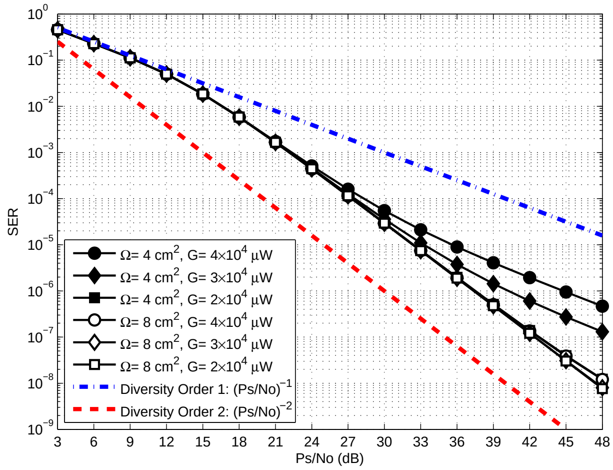
(b) 8PSK

Fig. 3. Upper and lower bounds for the asymptotic SER under various values of $\frac{P_s}{N_0}$ for different modulation schemes: (a) QPSK and (b) 8PSK ($N_b = 20$, $N_a = 20$, $\Upsilon_{sr} = 40$ dB, and $\Omega = 8$ cm^2).

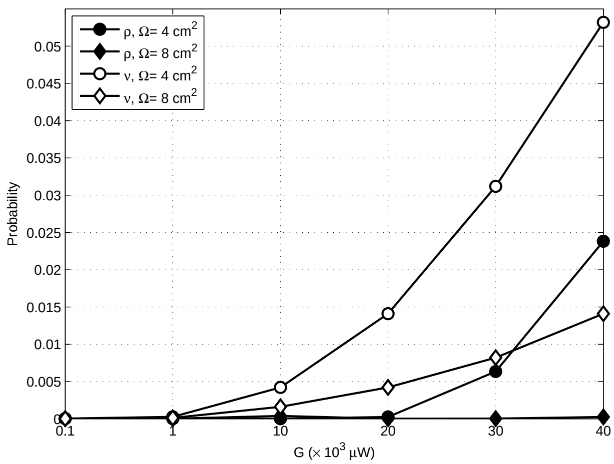
Based on Theorem 7, two myopic policies which are able to achieve the full diversity order if the energy quantum outage probability is equal to zero are included in the simulation for performance comparison. Both of the two myopic policies abide by the same threshold structure as described in (44), but with different relay actions when the relay is active. The first one is an aggressive policy, called Myopic Policy I, in which the largest available energy in the battery is consumed for relaying the signals when the relay is active. Regarding the second one, called Myopic Policy II, the relay helps transmit the signals only at the lowest transmission power level when it is active.

Fig. 2 shows the performance comparison between the exact SER and the asymptotic SER for QPSK and 8PSK modulation schemes under various SNR values of Υ_{sr}. With the obtained optimal policy, the exact SER is directly computed from (30) by applying the reward functions in (13)-(15) without any approximation, whereas the asymptotic SER is computed by using the approximate formula in (31). Just as mentioned in Theorem 3, we can observe from these two figures that the asymptotic SER is an upper bound for the exact SER, and our asymptotic results yield an excellent agreement with the exact curves in medium and high SNR regions. Hence, this expression is useful to correctly predict the characteristic of the SER performance for the energy harvesting cooperative communications in medium and high SNRs. Furthermore, the SER performance is improved as the operating SNRs Υ_{sr} and Υ_{sd} increase. We can make an interesting observation that for a fixed modulation scheme at $\Upsilon_{sr} = 30$ dB, the optimal policy with a larger number of affordable relay actions can achieve a better SER performance, but the performance curves, e.g., $N_a = 2$ and $N_a = 8$, become identical when Υ_{sd} is sufficiently high, no matter how many number of non-zero actions is available for the relay node.

Selected examples of the upper and lower bounds for the asymptotic SER versus $\frac{P_s}{N_0}$ are demonstrated in Fig. 3, where the parameters of N_b, N_a and Υ_{sr} are set as 20, 20 and 40 dB,



(a) Diversity order



(b) Stationary state probability ρ

Fig. 4. The diversity order of the asymptotic SER and the corresponding stationary state probability ρ under different values of the solar panel size Ω and the basic transmission power level G (QPSK, $N_b = 4$, $N_a = 4$, and $\eta_{sr} = \eta_{rd} = \eta_{sd} = 1$).

respectively. The adopted modulation schemes in Fig. 3(a) and Fig. 3(b) are QPSK and 8PSK, respectively. The transmission power level could be $G = 1 \times 10^2$ μW or $G = 3 \times 10^4$ μW, resulting in different energy quantum sizes. It is found that the lower and upper bounds are quite tight when $\frac{P_s}{N_0}$ ranges from mediate to high SNRs, which confirms our qualitative findings in Theorem 5. In particular, the asymptotic SER and the corresponding lower bound performance are almost overlapped for $G = 1 \times 10^2$ μW, since $\sum_{a \neq 0} \frac{\varsigma_a}{a} \approx \frac{1}{N_a - 1} \sum_{a \neq 0} \varsigma_a$ in this case. This phenomenon is accredited to the fact that when the basic transmission power G is considerably low, the probability for harvesting a huge number of energy quanta becomes very high and thus the occurrence probability ς_a with the highest power action $a = N_a - 1$ dominates the others.

Fig. 4 shows the diversity order of the asymptotic SER and the stationary state probability ρ under different values of the solar panel size Ω and the basic transmission power level G. The parameters of N_b and N_a are both set as four, and the adopted modulation scheme is QPSK. We also include

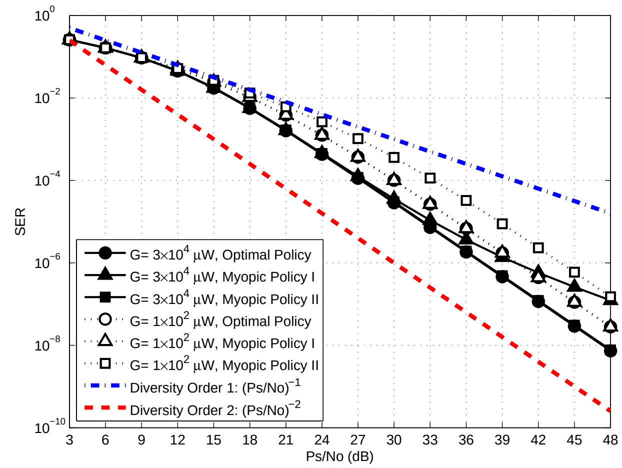


Fig. 5. The exact SER for the optimal policy and the two myopic policies under various values of $\frac{P_s}{N_0}$ (QPSK, $N_b = 8$, $N_a = 8$, $\Omega = 8$ cm^2, and $\eta_{sr} = \eta_{rd} = \eta_{sd} = 1$).

an average energy quantum outage probability ν in Fig. 4(b), which is averaged over the energy harvesting steady state probability, i.e., $\nu = \sum_{l=0}^{N_e-1} P\left(Q_e = l\right) P\left(E = 0 | \, Q_e = l\right)$. Since the validity of the asymptotic SER is attested to in Fig. 2, the diversity order for the exact SER can be quantified by inspecting the asymptotic SER. We can see from Fig. 4(a) that for $\Omega = 4$ cm^2 and $G = 4 \times 10^4$ μW or $G = 3 \times 10^4$ μW, the diversity order for the asymptotic SER is one. While the solar panel size is enlarged to $\Omega = 8$ cm^2 and the basic transmission power level is set below $G = 3 \times 10^4$ μW, the slope of the asymptotic SER bears a resemblance to the performance curve with a diversity order of two. The reasons behind this can be explained as follows. As shown in Fig. 4(b) and Table I, when Ω and G are sufficiently large and small, respectively, the average energy quantum outage probability ν for the relay node is almost zero, thereby resulting in an almost zero stationary state probability ρ and a diversity order of two. In contrast, the diversity order for the energy harvesting cooperative communications turns out to be one when the stationary state probability ρ is not equal to zero.

Fig. 5 demonstrates the exact SER for the optimal policy and the two myopic policies under two different basic transmission power levels G. The parameters of N_b and N_a are both set as eight, the solar panel size is given by $\Omega = 8$ cm^2, and the adopted modulation scheme is QPSK. It is shown that for $G = 3 \times 10^4$ μW, the exact SER performance of the Myopic Policy II is comparable to that of the optimal policy, since the energy is spent conservatively in the Myopic Policy II and the stationary state probability ρ for which the relay runs out of the battery is very small. On the other hand, it can be seen that the Myopic Policy I performs worst than the optimal policy when the operating SNR becomes high due to the aggressive use of the energy, yielding a relatively large stationary state probability ρ. For the case of $G = 1 \times 10^2$ μW, where the energy quantum outage probability becomes much close to zero, one can observe that the performance of the optimal policy is still superior to those of the two myopic policies, while they exhibit the same diversity order of two (reflected by the slope of the SER). This is because the optimal policy can achieve the maximum energy harvesting gain. Besides, it is found that the Myopic Policy I has an energy harvesting gain of about 4 dB in terms of $\frac{P_s}{N_0}$ over the Myopic Policy II due to the power scaling effect $\frac{1}{a}$ in (55).

VI. Conclusions

In this paper, we investigated the long-term average SER minimization problem for the DF cooperative communications with an energy harvesting relay node. By means of the MDP, we designed the optimal stochastic cooperative transmission scheme for the relay node that varies the transmission power with different energy harvesting, channel, battery, and decoding states. Closed-form expressions for the exact SER and the asymptotic SER of the proposed optimal cooperative transmission policy were analytically derived. Furthermore, we provided the SER upper and lower bounds to quantify the diversity order and the energy harvesting gain, which proves that a diversity order of two is guaranteed as long as the energy quantum outage probability is zero. We also explored a threshold-type structure for the cooperative transmission policies which are capable of achieving the full diversity, and examined the energy quantum supporting way which can lower the energy quantum outage probability in a practical energy harvesting model. The developed results and the effect of various system parameters on the SER performance were validated through extensive computer simulations. The design framework is useful toward understanding how to deploy the energy harvesting relays in cooperative networks as more efficiently as possible.

References

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

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

[3] W. Su, A. K. Sadek, and K. J. Ray Liu, “Cooperative communication protocols in wireless networks: performance analysis and optimum power allocation, ” *Wireless Personal Commun.*, vol. 44, no. 2, pp. 181-217, Jan. 2008.

[4] S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: survey and implications, ” *IEEE Commun. Surveys Tutorials*, vol. 13, no. 3, pp. 443-461, Third Quad. 2011.

[5] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Transmission with energy harvesting nodes in fading wireless channels: optimal policies, ” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732-1743, Sept. 2011.

[6] J. Yang and S. Ulukus, “Optimal packet scheduling in an energy harvesting communication system, ” *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220-230, Jan. 2012.

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

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

[9] O. Ozel and S. Ulukus, “Achieving AWGN capacity under stochastic energy harvesting, ” *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6471-6483, Oct. 2012.

[10] D. Niyato, E. Hossain, and A. Fallahi, “Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: performance analysis and optimization, ” *IEEE Trans. Mobile Comput.*, vol. 6, no. 2, pp. 221-236, Feb. 2007.

[11] M. L. Ku, Y. Chen, and K. J. R. Liu, “Data-driven stochastic scheduling for solar-powered sensor communications, ” *Proc. IEEE Glob. Conf. Signal and Info. Process. (GlobalSIP)*, 2014.

[12] M. L. Ku, Y. Chen, and K. J. R. Liu, “Data-driven stochastic models and policies for energy harvesting sensor communications, ” to appear, *IEEE J. Sel. Areas Commun.*, Special Issue on Wireless Communications Powered by Energy Harvesting and Wireless Energy Transfer.

[13] Y. Luo, J. Zhang, and K. B. Letaief, “Optimal scheduling and power allocation for two-hop energy harvesting communication systems, ” *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4729-4741, Sep. 2013.

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

[15] I. Ahmed, A. Ikhlef, R. Schober, and R. K. Mallik, “Power allocation for conventional and buffer-aided link adaptive relaying systems with energy harvesting nodes, ” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1182-1195, Mar. 2014.

[16] A. Minasian, S. Shahbazpanahi, and R. S. Adve, “Energy harvesting cooperative communication systems, ” *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6118-6131, Nov. 2014.

[17] D. Gunduz and B. Devillers, “Two-hop communication with energy harvesting, ” *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 201-204, 2011.

[18] O. Orhan and E. Erkip, “Optimal transmission policies for energy harvesting two-hop networks, ” *Proc. Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, 2012.

[19] K. Tutuncuoglu, B. Varan, and A. Yener, “Energy harvesting two-way half-duplex relay channel with decode-and-forward relaying: optimum power policies, ” *Proc. IEEE Int. Conf. Digital Signal Process. (DSP)*, pp. 1-6, 2013.

[20] W. Li, M. L. Ku, Y. Chen, and K. J. R. Liu, “On the achievable sum rate for two-way relay networks with stochastic energy harvesting, ” *Proc. IEEE Glob. Conf. Signal and Info. Process. (GlobalSIP)*, 2014.

[21] H. Li, N. Jaggi, and B. Sikdar, “Cooperative relay scheduling under partial state information in energy harvesting sensor networks, ” *Proc. IEEE Glob. Commun. Conf.*, pp. 1-5, 2010.

[22] M. Kashef and A. Ephremides, “Optimal partial relaying for energy-harvesting wireless networks, ” *IEEE/ACM Trans. on Networking*, vol. PP, no. 99, pp. 1, 2014.

[23] T. Li, P. Fan, and K. B. Letaief, “Outage porobability of energy harvesting relay-aided cooperative networks over Rayleigh fading channel, ” *arXiv preprint arXiv:1409.0286*, 2014.

[24] B. Medepally and N. B. Mehta, “Voluntary energy harvesting relays and selection in cooperative wireless networks, ” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3543-3553, Nov. 2010.

[25] H. S. Wang and N. Moayeri, “Finite-state Markov channel-a useful model for radio communication channels, ” *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163-171, Feb. 1995.

[26] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*, New York: John Wiley & Sons, 2000.

[27] M. L. Puterman, *Markov Decision Processes- Discrete Stochastic Dynamic Programming*, New York: Wiley, 1994.

[28] N. R. E. Laboratory. (2012, Feb.) Solar radiation resource information. [Online]. Available: http://www.nrel.gov/rredc/.