Diversity Analysis for Two-Way Multi-Relay Networks with Stochastic Energy Harvesting

Wei Li[*†], Meng-Lin Ku[‡], Yan Chen[*], and K. J. Ray Liu[*]

[*]Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
[†]Department of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an, 710049, China
[‡]Department of Communication Engineering, National Central University, Taiwan
Email: wli52140@umd.edu, mlku@ce.ncu.edu.tw, {yan, kjrliu}@umd.edu

***Abstract*—The pairwise error probability (PEP) performance is investigated for the energy harvesting (EH) two-way (TW) multi-relay network with stochastic EH models. In this network, two relay nodes exploit the amply-and-forward (AF) protocol and space-time network coding (STNC) to help two source nodes exchange information simultaneously. We assume only one relay node is solar-powered and equipped with a finite-capacity battery. The optimal transmission policy for the EH relay is proposed to minimize the long-term PEP by considering the causal EH information, battery energy and stochastic channel status. Thus, the design framework is formulated as a Markov decision process (MDP). We reveal that the full diversity order is achievable only if the EH relay's battery empty probability is equal to zero. Further, the PEP performance can be improved when the battery empty probability becomes smaller. Simulations validate the theoretical analysis and show that the proposed optimal policy outperforms other myopic policies.**

I. Introduction

Energy harvesting (EH) communication has attracted significant attention due to its effectiveness to tackle energy supply problems in wireless networks without fixed power grid. Recently, there has been a growing interest in investigating two-way (TW) relay networks with EH nodes, which can harvest energy from ambient energy sources. Unlike the traditional TW relay networks, not only the TW relay fading channels, but also the stochastic and uncertain energy harvested from environments, should be seriously considered in the transmission scheduling of the EH TW relay networks.

Power allocation for maximizing short-term sum rates in EH TW relay networks was studied in [1]-[3] under deterministic EH models with the assumption of non-causal EH information. The authors in [1] considered an EH relay with a data buffer to cache data so that a more flexible scheduling policy can be applied. Moreover, a generalized iterative directional water-filling method was developed in [2]. In [3], an optimal scheme was proposed by considering the uncertainty of channel state information. In addition, an optimal transmission policy for maximizing the long-term sum rates of EH TW relay networks was developed in [4] by exploiting stochastic EH models, in which the EH information is causal and unknown to transmitters. Besides, the wireless energy transfer in TW relay networks was investigated in the literature. The authors in [5] studied the optimal beamforming for multiple relay nodes to maximize the throughput of the TW relay network, wherein the source nodes can harvest energy from the signals of the relay nodes. Similarly, the authors in [6] concentrated on the transceiver design for the EH TW relay network, wherein a multi-antenna relay can transfer the energy and information to two source nodes simultaneously. The wireless energy transfer from the source nodes to the relay in the EH TW relay network was also considered in [7]. However, so far, the performance of the EH TW relay network with multiple relay nodes has not been investigated for stochastic EH models.

Motivated by the above discussions, we investigate the pairwise error probability (PEP) performance of the EH TW multi-relay network utilizing the data-driven stochastic EH model in [8]. In this network, two source nodes exchange information simultaneously with the help of two relay nodes utilizing the amplify-and-forward (AF) protocol and space-time network coding (STNC) to fulfill cooperative communications [9]. Since there is no direct link between the two source nodes, one relay node with fixed power supply is exploited to guarantee the reliable communication. Besides, we assume the other relay node with a finite-sized battery is solar-powered, and aim at studying the impact of the stochastic EH on the PEP performance. First, an optimal transmission policy for the EH relay node is developed to minimize the long-term PEP through a Markov decision process (MDP) framework, by considering the stochastic harvested energy and channel states. Then, we analyze the diversity performance, and an interesting result is revealed, i.e., the full diversity order can be achieved only if the EH relay's battery empty probability is equal to zero; otherwise, the diversity order is only one.

This work was supported by the China Scholarship Council (CSC), the National Natural Science Foundation of China (NSFC) under Grant No.61401348, and the Postdoctoral Science Foundation of China under Grant No.2014M550493.

II. Energy Harvesting Two-Way Multi-Relay Networks with Network Coding

An EH TW relay network is considered in Fig. 1, where two wireless source nodes, A and B, exchange information

978-1-4799-7591-4/15/$31.00 ©2015 IEEE

simultaneously via two intermediate relay nodes, R_1 and R_2. It is assumed that each node is half-duplex and equipped with a single antenna. A, B and R_2 are traditional wireless nodes with fixed transmission power, while R_1 is equipped with a finite-sized battery and can only harvest energy from the solar for data transmission. We also assume that there is no direct link between the two source nodes, and the wireless channels are reciprocal, quasi-static and Rayleigh flat fading. All channel coefficients are independent and identically distributed (i.i.d.) random variables with complex Gaussian distribution $\mathcal{CN}(0,1)$. R_1 has the perfect knowledge of the channel state information (CSI) related to itself. Moreover, the amplify-and-forward (AF) protocol and STNC vectors are utilized by R_1 and R_2 to forward signals. The whole transmission period includes a multiple access (MA) phase and a broadcast (BC) phase, and each phase consists of $L=2$ slots. The detailed transmission period is described as follows.

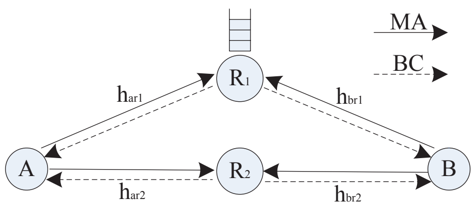

Fig. 1. The EH TW multi-relay network

A. MA Phase. The source nodes A and B simultaneously transmit a set of $L=2$ symbols to the relays in L slots. The received signal vector at $R_l\,(l\in\{1,2\})$, is given by

$$\mathbf{y}_{sr_l} = h_{ar_l}\sqrt{P}\mathbf{s}_a + h_{br_l}\sqrt{P}\mathbf{s}_b + \mathbf{n}_{sr_l}, \tag{1}$$

where $\mathbf{s}_a$ and $\mathbf{s}_b$ are the signal vectors transmitted by the nodes A and B, respectively, P represents their transmission power, h_{ar_l} and h_{br_l} denote the channel coefficients between R_l and the two source nodes, $\mathbf{n}_{sr_l}$ is the additive white Gaussian noise (AWGN) vector with i.i.d. entries following $\mathcal{CN}(0,N_0)$ at R_l. Subsequently, each relay takes linear transformation on the received signal vector, and the transmitted signal of R_l is expressed as $x_{r_l} = \beta_l\theta_l^{\mathrm{T}}\mathbf{y}_{sr_l}$, where $\theta_l = [\theta_{l1},\theta_{l2}]^{\mathrm{T}}$ is the STNC vector with unit norm for R_l, and meets the full diversity criterion [9], i.e., $|\theta_l(\mathbf{s}-\tilde{\mathbf{s}})| \neq 0, \forall \mathbf{s}, \forall\tilde{\mathbf{s}}(\neq \mathbf{s})$. The signal amplifying factor β_l is defined as $\beta_l = \sqrt{LP_{R_l}/\left(P\gamma_{ar_l}+P\gamma_{br_l}+N_0\right)}$, where $\gamma_{ar_l} = |h_{ar_l}|^2$, $\gamma_{br_l} = |h_{br_l}|^2$, P_{R_l} denotes the transmission power of R_l.

B. BC Phase. The two relay nodes take turn to forward signal in two slots, i.e., each slot is solely occupied by one relay. Each user node can subtract its self-interference, and thus the received signal at the source A can be expressed as

$$\tilde{y}_{r_la} = h_{ar_l}h_{br_l}\beta_l\sqrt{P}\theta_l^{\mathrm{T}}\mathbf{s}_b + h_{ar_l}\beta_l\theta_l^{\mathrm{T}}\mathbf{n}_{sr_l} + n_{r_la}, \tag{2}$$

where n_{r_la} is the AWGN with distribution $\mathcal{CN}(0,N_0)$ at the node A. Observing the received signal $\{\tilde{y}_{r_la}\}_{l=1}^{2}$ in the two slots, The node A can exploit the maximum likelihood (ML) criterion to jointly decode $\mathbf{s}_b$.

Let $\Delta\mathbf{s}_b = \mathbf{s}_b - \tilde{\mathbf{s}}_b \neq 0$, $\alpha_l = \left|\theta_l^{\mathrm{T}}\Delta\mathbf{s}_b\right|^2\,(l\in\{1,2\})$, and

$$W_{R_l} = \frac{\gamma_{ar_l}\gamma_{br_l}P_{R_l}P\alpha_l}{4\left[\left(P+P_{R_l}\right)\gamma_{ar_l}+P\gamma_{br_l}+N_0\right]N_0}, l\in\{1,2\}. \tag{3}$$

Then the instantaneous PEP of the source node A conditioned on $\{\gamma_{ar_l}\}_{l=1}^2$ and $\{\gamma_{br_l}\}_{l=1}^2$ can be calculated by using the Q-function $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$ as follows [9]

$$\begin{aligned}\tilde{\Pr}\{\mathbf{s}_b\to\tilde{\mathbf{s}}_b\} &= Q\left(\sqrt{2W_{R_1}+2W_{R_2}}\right) \\ &= \frac{1}{\pi}\int_0^{\pi/2}\exp\left(-\frac{W_{R_1}+W_{R_2}}{\sin^2\theta}\right)d\theta \\ &\approx P_{e,R_1}\times P_{e,R_2},\end{aligned} \tag{4}$$

where $P_{e,R_1} = \exp\left(-W_{R_1}\right)$ and $P_{e,R_2} = \frac{1}{\pi}\int_0^{\pi/2}\exp\left(-\frac{W_{R_2}}{\sin^2\theta}\right)d\theta$ represent the capacities of contribution to the PEP by R_1 and R_2, respectively.

III. Markov Decision Process Formulation

Our goal is to find the optimal transmission power of the EH relay R_1, in order to minimize the long-term PEP considering the causal solar power condition, the battery energy and the channel status. In such a case, the design framework is formulated as an MDP model, which is mainly composed of the state space, the action space, the state transition probability and the reward function. Thus, let $\mathcal{S} = \mathcal{S}_E\times\mathcal{S}_{AR_1}\times\mathcal{S}_{BR_1}\times\mathcal{S}_B$ be a four-tuple state space, where $\times$ denotes the Cartesian product, $\mathcal{S}_E = \{0,1,\cdots,N_e-1\}$ represents the solar power state subspace, $\mathcal{S}_{AR_1} = \{0,1,\cdots,N_c-1\}$ and $\mathcal{S}_{BR_1} = \{0,1,\cdots,N_c-1\}$ denote the two channel state subspaces for A-R_1 link and B-R_1 link, respectively, and $\mathcal{S}_B = \{0,1,\cdots,N_b-1\}$ is the battery state subspace for R_1. Meanwhile, let $S = (S_E,S_{AR_1},S_{BR_1},S_B)\in\mathcal{S}$ represent the stochastic state of the MDP. Define $\mathcal{A} = \{0,1,\cdots,N_p-1\}$ as the action space, and the action $a\in\mathcal{A}$ represents the transmission power of R_1.

A. Solar Power State. An N_e-state stochastic EH model in [8] is used to capture the dynamic of the harvested solar power conditions. This data-driven solar power model is a Gaussian mixture hidden Markov chain, and different states possess different solar irradiance intensity. The harvested solar power per unit area, P_h, is a continuous random variable with Gaussian distribution $\mathcal{N}(\mu_e,\rho_e)$ for the solar power state $e\in\mathcal{S}_E$. Thus, the harvested solar energy $E_h = P_hT\Omega\eta$ during one policy period T is also continuous, where Ω is the solar panel area size, and η denotes the energy conversion efficiency. However, in practice, the solar energy storage and usage model is discrete and designed based on the numbers of energy quanta. Thus the harvested energy is first quantized in units of one basic energy quantum E_U and then stored in the battery for the forthcoming data transmission.

B. Battery State. The battery storage capacity is finite and evenly quantized into multiple levels. The battery state denotes the number of available energy quanta in the battery.

C. Channel States. The instantaneous channel gains of the links related to R_1, γ_{ar_1} and γ_{br_1}, are quantized into N_c levels using several thresholds, given by $\Gamma = \{0=\Gamma_0, \Gamma_1, \cdots, \Gamma_{N_c}=\infty\}$. The i^{th} channel state represents the channel gain belongs to the interval $[\Gamma_i, \Gamma_{i+1})$. Moreover, the channel state transition is modeled as an N_c-state Markov chain in [10].

D. EH Relay Action. If the action $a \in \mathcal{A}$ is chosen, the transmission power of R_1 is set as $P_{R_1} = aP_U$, where P_U is a constant transmission power corresponding to one energy quantum E_U in the battery during the half transmission period, i.e., $E_U = \frac{1}{2}P_U T$.

E. MDP State Transition. Since we utilize the *harvest-store-use* protocol [11], the harvested energy in the current policy period is first stored in the battery, and then consumed for data transmission in the next policy period. Thus, the battery state will transit from the state b to the state $b' = \min(b-a+q,\, N_b-1)$, when the action a is taken and the number of harvested energy quanta is q in the current period. Meanwhile, it implies that $a \le b$ due to the battery constraint. Besides, the battery state transition probability at the e^{th} solar power state, $P_a(S_B=b'|S_B=b,S_E=e)$, is given in [8].

Since the transition probabilities of the solar power states and channel states are independent, the system state transition probability from $s=(e,f,g,b)$ to $s'=(e',f',g',b')$ for the action a can be expressed as

$$\begin{aligned} P_a(s'|s) =& P(S_E=e'|S_E=e)\cdot P(S_{AR_1}=f'|S_{AR_1}=f) \\ &\cdot P(S_{BR_1}=g'|S_{BR_1}=g)\cdot P_a(S_B=b'|S_B=b,S_E=e), \end{aligned} \quad (5)$$

where the solar power state transition and the channel state transition are defined in [8] and [10], respectively,

F. Reward Function. According to (4), the PEP is determined by both P_{e,R_1} and P_{e,R_2}; however, only the term P_{e,R_1} is related with the transmission power P_{R_1} in the MDP. Thus, the reward function is defined as the conditional P_{e,R_1} associated with the state $s=(e,f,g,b)$ and action a:

$$\begin{aligned} R_a(s) =& \frac{\int_{\Gamma_g}^{\Gamma_{g+1}}\int_{\Gamma_f}^{\Gamma_{f+1}} e^{-\gamma_1}e^{-\gamma_2}\cdot\exp(-W_{R_1})\,d\gamma_1 d\gamma_2}{P(S_{AR_1}=f)\cdot P(S_{BR_1}=g)} \\ \triangleq& P_{e,R_1}(a,f,g). \end{aligned} \quad (6)$$

For simplicity, the transmission power of the source nodes is P=P_U. From (3), W_{R_1} can be rewritten as

$$W_{R_1} = \gamma_1\gamma_2 a\eta^2\alpha_1/[4((a+1)\eta\gamma_1+\eta\gamma_2+1)], \quad (7)$$

where $\eta = \frac{P_U}{N_0}$, $\gamma_1 = \gamma_{ar_1}$, and $\gamma_2 = \gamma_{br_1}$.

G. Optimization of Relay Transmission Policy. Define $\pi(s): \mathcal{S} \to \mathcal{A}$ as the power action. The goal of the MDP is to find the optimal $\pi(s)$ in the state s to minimize the expected discounted long-term reward, given by

$$V_\pi(s_0) = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty}\lambda^k R_{\pi(s_k)}(s_k)\right], s_k \in \mathcal{S}, \pi(s_k) \in \mathcal{A}, \quad (8)$$

where $0 \le \lambda < 1$ is a discount that guarantees the convergence. The optimal policy can be found by the Bellman equation, which can be solved by the value iteration approach [12].

In the following, a non-conservative property of the optimal relay transmission policy is provided.

Proposition 1: For any fixed system state $s=(e,f,g,b \ge 1) \in \mathcal{S}$ with the non-empty battery, in asymptotically high signal-to-noise power ratio (SNR), i.e., $\eta \gg 1$, the optimal power action a^* is non-zero.

The reason is explained as follows. In asymptotically high SNRs, the immediate reward for non-zero actions approaches to zero. As compared with the zero action that conserves the battery energy to improve the future reward, the non-zero action leads to better long-term reward in high SNRs.

IV. Diversity Analysis

In this section, the expected PEP and diversity order of the proposed optimal transmission policy in the EH TW multi-relay network are analyzed. For a fixed policy $\pi(s)$, the steady state probability $p_\pi(s)$ can be computed by using the state transition probability in (5) and the balance equation in [8]. Thus, the expected reward can be calculated by taking expectation over the reward function using $p_\pi(s)$ as follows

$$\bar{R}_\pi = \mathbb{E}_\pi\{P_{e,R_1}\} = \sum_{s\in\mathcal{S}} p_\pi(s)\times R_{a=\pi(s)}(s). \quad (9)$$

From (4), the expected PEP can be computed as

$$\Pr\{\mathbf{s}_b \to \tilde{\mathbf{s}}_b\} \approx \mathbb{E}_\pi\{P_{e,R_1}\}\times\mathbb{E}\{P_{e,R_2}\}, \quad (10)$$

where $\mathbb{E}\{P_{e,R_2}\}$ is the expectation result with respect to the channel gains, γ_{ar_2} and γ_{br_2}. When $\eta \gg 1$, it can be obtained that $\mathbb{E}\{P_{e,R_2}\} \propto \eta^{-1}/\alpha_2$, by following the similar steps of Lemma 1 in [13].

Next, the following proposition is introduced to compute the asymptotic approximation of the reward function.

Proposition 2: In asymptotically high SNRs, i.e., $\eta \gg 1$, the reward function $R_a(s)\,(a>0)$ can be approximated as

$$P_{e,R_1}(a,f,g) \approx \begin{cases} \frac{8\eta^{-1}}{a(1-e^{-\Gamma_1})\alpha_1}, & \min(f,g)=0; \\ 0, & \min(f,g)\ge 1, \end{cases} \quad (11)$$

Proof: From (7), when $\eta \gg 1$, we have $W_{R_1} \approx \frac{a\eta\alpha_1\gamma_1\gamma_2}{4(\gamma_1+\gamma_2)}$. By using the harmonic mean inequality $\frac{1}{2}\min(x,y) \le \frac{xy}{x+y} \le \min(x,y)$, it is obtained that $W_{R_1} \approx \frac{a\eta\alpha_1}{8}\min(\gamma_1,\gamma_2)$. For the channel states where $S_{AR_1} = f < S_{BR_1} = g$, the asymptotic approximation of the reward function in (6) is given by

$$\begin{aligned} R_a(s) \approx& \frac{\int_{\Gamma_g}^{\Gamma_{g+1}}\int_{\Gamma_f}^{\Gamma_{f+1}} e^{-\gamma_1}\cdot e^{-\gamma_2}\cdot\exp(-\frac{a\alpha_1\eta}{8}\gamma_1)d\gamma_1 d\gamma_2}{P(S_{AR_1}=f)\cdot P(S_{BR_1}=g)} \\ \approx& \begin{cases} \frac{8\eta^{-1}}{a(1-e^{-\Gamma_1})\alpha_1}, & f=0; \\ 0, & f\ge 1. \end{cases} \end{aligned} \quad (12)$$

Similar derivation can be obtained when $f > g$. As as result, we can get (11). ∎

Finally, we provide the following theorem to for the diversity performance of the EH TW multi-relay network.

Theorem 1: In the proposed EH TW network-coded multi-relay network with the optimal policy π^*, if the EH relay's

TABLE I
SIMULATION PARAMETERS

Modulation type	QPSK
Basic transmission power (P_U)	10mW
Policy management period (T)	300s
Energy conversion efficiency (η)	20%
Channel simulation model	Jakes' model
Normalized Doppler frequency (f_D)	0.05
Channel quantization thresholds (Γ)	$\{0, 0.3, 0.6, 1.0, 2.0, 3.0, \infty\}$
Discount factor (λ)	0.99

battery empty probability is equal to zero, the full diversity order $d = 2$ can be achieved. Otherwise, the diversity order is only $d = 1$.

Proof: The expected reward in (9) for the optimal policy π^* can be rewritten by considering the battery states as follows

$$\begin{aligned}\bar{R}_{\pi^*} = \sum\nolimits_{s\in\mathcal{S}} &\left[p_{\pi^*}\left(e,f,g,b=0\right)\times R_{a^*}\left(e,f,g,b=0\right)\right.\\ &\left.+p_{\pi^*}\left(e,f,g,b\geq 1\right)\times R_{a^*}\left(e,f,g,b\geq 1\right)\right]. \end{aligned} \quad (13)$$

Since we utilize the harvest-store-use protocol, the relay R_1 keeps silent when the battery is empty. Meanwhile, according to Proposition 1, the optimal action a^* is non-zero in high SNRs when the battery is non-empty. Further, by applying Proposition 2, the asymptotic approximation of the expected reward in high SNRs for the optimal policy π^* can be calculated as

$$\bar{R}_{\pi^*} \approx P_{\pi^*}\{b=0\} + \frac{\sum_{s\in\mathcal{S}_0} p_{\pi^*}\left(s\right)}{a^*\left(1-e^{-\Gamma_1}\right)\alpha_1}\eta^{-1}, \quad (14)$$

where $P_{\pi^*}\{b=0\} = \sum_{s\in\mathcal{S}} p_{\pi^*}\left(e,f,g,b=0\right)$ represents the battery empty probability, and $\mathcal{S}_0 = \{s = (e, f, g, b), \min(f, g) = 0, b \geq 1, s \in \mathcal{S}\}$ denotes the system state set in which the battery is non-empty and at least one channel state is zero.

From (10) and (14), in high SNRs we have

$$\Pr\{\mathbf{s}_b \to \tilde{\mathbf{s}}_b\} \propto \frac{P_{\pi^*}\{b=0\}}{\alpha_2}\eta^{-1} + \frac{\sum_{s\in\mathcal{S}_0} p_{\pi^*}\left(s\right)}{a^*\left(1-e^{-\Gamma_1}\right)\alpha_1\alpha_2}\eta^{-2},$$

where α_1 and α_2 must be positive due to the full diversity criterion of the STNC.

In summary, if $P_{\pi^*}\{b=0\} > 0$, the expected PEP in high SNRs is dominated by the first term, resulting in a diversity order of one. Otherwise, a full diversity order of two is achieved. ■

V. SIMULATION RESULTS

The average PEP of the proposed optimal policy based on the stochastic EH model in [8] is evaluated by Monte-Carlo simulation. The numbers of the solar power states, channel states and battery states are four, six and twelve, respectively. Other simulation parameters are listed in Table I. The STNC matrix is an $L \times L$ Vandemonde matrix, whose entries are defined as $\theta_{ln} = e^{j\pi(4l-1)(n-1)/(2L)},\ l, n = 1, 2, \cdots, L$ [9].

Fig. 2 illustrates the average PEP of our proposed optimal policy for different solar panel area sizes Ω. The EH relay's battery empty probability $P_{\pi^*}\{b=0\}$ becomes smaller when Ω gets larger. When Ω is set as 0.5, 1 and $2\mathrm{cm}^2$, the battery empty probability is approximately given as 10^{-1}, 10^{-2} and 10^{-10}, respectively. It can be seen that the diversity order is one when $\Omega = 0.5$ or $1\mathrm{cm}^2$, while the diversity order is close to two when Ω is set as $2\mathrm{cm}^2$. This phenomenon coincides with Theorem 1. Further, the average PEP can be improve by changing Ω from $0.5\mathrm{cm}^2$ to $1\mathrm{cm}^2$, even though the diversity order remains one. This is because the smaller $P_{\pi^*}\{b=0\}$ is, the better PEP performance is achieved.

Fig. 2 also compares the PEP performance between the proposed optimal policy and two myopic policies, in which the EH relay transmission power is set without concern for the stochastic system states. In Greedy Policy, the whole available energy in the battery is consumed for one policy management period. Regarding with Economical Policy, the EH relay utilizes the lowest power, i.e., the basic power P_u, to transmit signals. It can be seen that the average PEP of the optimal policy is superior to those of the two myopic policies.

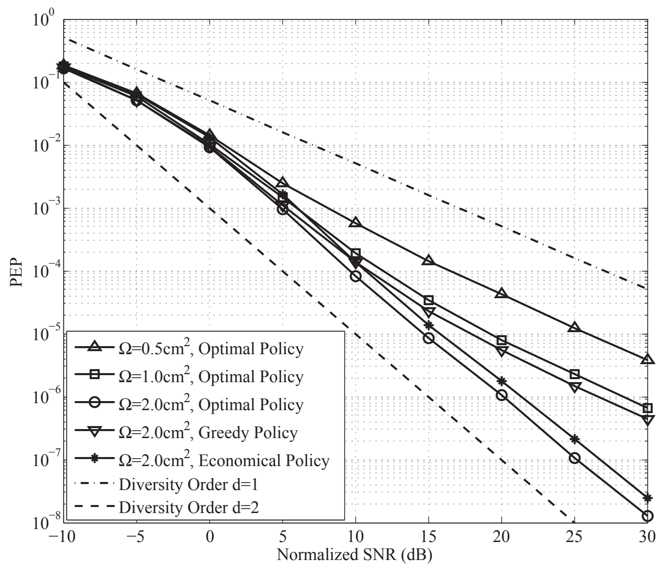

Fig. 2. PEP of the optimal policy for different solar panel area sizes Ω, as well as the performance comparison between the optimal policy and two myopic policies

VI. CONCLUSION

The PEP performance is analyzed in the EH TW network-coded multi-relay network with stochastic EH models. The optimal transmission policy for the EH relay is developed to minimize the long-term PEP by considering the stochastic EH information and channel status. We prove that the full diversity order is achieved only if the EH relay's battery empty probability is equal to zero.

References

[1] B. Varan and A. Yener, "The Energy Harvesting Two-Way Decode-and-Forward Relay Channel with Stochastic Data Arrival," in *Proc. of 2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2013, pp. 371–374.

[2] K. Tutuncuoglu, B. Varan, and A. Yener, "Optimum Transmission Policies for Energy Harvesting Two-way Relay Channels," in *Proc. of 2013 International Conference on Communication (ICC)*, Budapest, Hungary, June 2013, pp. 586–590.

[3] I. Ahmed, A. Ikhlef, D. W. K. Ng, and R. Schober, "Optimal Resource Allocation for Energy Harvesting Two-Way Relay Systems with Channel Uncertainty," in *Proc. of 2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2013, pp. 345–348.

[4] W. Li, M.-L. Ku, Y. Chen, and K. J. R. Liu, "On the Achievable Sum Rate for Two-Way Relay Networks with Stochastic Energy Harvesting," in *Proc. of 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Atlanta, GA, USA, Dec. 2014, pp. 133–137.

[5] D. Li, C. Shen, and Z. Qiu, "Two-Way Relay Beamforming for Sum-Rate Maximization and Energy Harvesting," in *Proc. of 2013 International Conference on Communication (ICC)*, Budapest, Hungary, June 2013, pp. 3115–3120.

[6] Z. Wen, S. Wang, C. Fan, and W. Xiang, "Joint Transceiver and Power Splitter Design Over Two-Way Relaying Channel With Lattice Codes and Energy Harvesting," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 2039–2042, Nov. 2014.

[7] Z. Fang, X. Yuan, and X. Wang, "Distributed Energy Beamforming for Simultaneous Wireless Information and Power Transfer in the Two-Way Relay Channel," *IEEE Signal Processing Lett.*, vol. 22, no. 6, pp. 656–660, 2015.

[8] M.-L. Ku, Y. Chen, and K. J. R. Liu, "Data-Driven Stochastic Transmission Policies for Energy Harvesting Sensor Communications," *IEEE J. Select. Areas Commun.*, vol. PP, no. 99, p. 1, Jan. 2015.

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

[10] H. S. Wang and N. Moayeri, "Finite-State Markov Channel-A Useful Model for Radio Communication Channels," *IEEE Trans. Wireless Commun.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.

[11] N. Michelusi, L. Badia, and M. Zorzi, "Optimal Transmission Policies for Energy Harvesting Devices with Limited State-of-Charge Knowledge," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3969–3982, Nov. 2014.

[12] M. Puterman, *Markov Decision Process-Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.

[13] W. Guan and K. Liu, "Diversity Analysis of Analog Network Coding with Multi-User Interferences," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 668–679, Feb. 2013.