ENERGY HARVESTING FOR RELAY-ASSISTED COMMUNICATIONS

Arin Minasian and Raviraj S. Adve

University of Toronto
Dept. of Elec. & Comp. Eng.,
Toronto, Ontario, Canada
{aminasian,rsadve}@comm.utoronto.ca

Shahram ShahbazPanahi

Univ. of Ontario Institute of Technology
Dept. of Elec,. Software and Comp. Eng.,
Oshawa, Ontario, Canada,
shahram.shahbazpanahi@uoit.ca

ABSTRACT

In this paper, we examine the problem of throughput maximization in an energy-harvesting two-hop amplify-and-forward relay network. This problem is investigated over a finite time horizon and in an online setting, where the causal knowledge of the harvested energy and that of fading are available. We use Markov decision process (MDP) formulation to present a mathematically tractable solution to the throughput maximization problem. In this solution, optimal power-use policy is obtained using backward induction algorithm of the corresponding discrete dynamic programming problem. We also present properties of the optimal policy for an important special case, where the power control at transmitters is limited to on-off switching. These properties facilitate the implementation of the MDP based solution. Our numerical simulations show that the proposed method outperforms existing solutions to this problem.

1. INTRODUCTION

The proliferation of wireless devices and the spread of wireless communication networks render the task of energy supply an ever-growing challenge. Reliable and sustainable energy sources should be deployed to guarantee effective performance of wireless networks. In this regard, energy harvesting technologies are emerging as promising solutions. Such technologies enable wireless devices to benefit from sustainable and theoretically unlimited energy sources that are present in their surrounding environment. The random nature of ambient energy sources and the need for power efficiency necessitate design of novel and efficient power-use policies.

Designing efficient energy-harvesting communication systems is a relatively new research topic. In [1], the information theoretic capacity of an additive white Gaussian noise (AWGN) channel with an energy harvesting transmitter is derived. The problem of transmission time minimization is studied in [2], where an energy-harvesting setting is considered for a point-to-point communication system . In [3], the authors investigate the problem of short-term throughput maximization for a wireless link with a rechargeable transmitter

In [4˘6], for two-hop communication systems with energy harvesting nodes, the problem of throughput maximization is studied in a single-relay setting, where it is assumed that the non-causal knowledge of the energy harvesting profiles of transmitting nodes is available prior to start of the transmission. In [7], the problem of throughput maximization is studied for the Gaussian relay channel with energy harvesting constraints. The authors of [7] assume that the relay operates in half-duplex mode and that it uses the decode-and-forward (DF) relaying protocol. The work in [7] is extended to the case of buffer aided link-adaptive relaying systems in [8]. In [9], the authors examine joint relay selection and power allocation for throughput maximization in an amplify-and-forward (AF) relay network in an off-line setting. The authors of [10] study a problem similar to that considered in [9] for a single-relay case and propose an alternative solution to tackle the throughput maximization problem. In [11], using Markov decision process (MDP), the problem of maximizing the rate of information transfer is investigated for a single-link communication system.

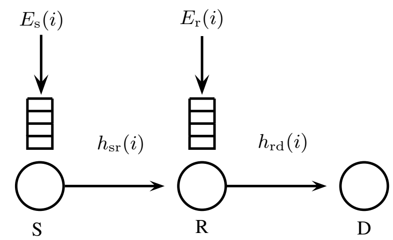

Fig. 1: A single-relay two-hop network with energy harvesting nodes.

The main focus of the aforementioned studies in energy-harvesting cooperative communication systems is to investigate the system in an off-line setting. To the best of our knowledge no mathematically tractable solution has been introduced to tackle this problem in the on-line setting, where only the causal knowledge of harvested energy and that of fading channel are available. Due to the practical importance of the on-line setting case, we herein propose an MDP-based approach to find the power-optimal transmission policies and derive certain properties of the optimal solution in such a setting.

This paper is organized as follows. The system model is presented in Section 2. Section 3 includes the MDP formulation for the throughput maximization problem in the online setting and presents certain properties of the optimal transmission policy in the special case of on-off power control. Simulation results are presented in Section 4. Concluding remarks are presented in Section 5.

2. SYSTEM MODEL

The system under consideration is an energy-harvesting two-hop relay network illustrated in Fig. 1. This network consists of an energy-harvesting source (S), an energy-harvesting relay (R) and a destination (D). The source-relay (S-R) and the relay-destination (R-D) channels are assumed to be statistically independent. The transmission occurs over a time block comprising of N transmission intervals, each of length $2T$. Each transmission interval includes two

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

time slots. In the first time slot, the source transmits its message to the relay. In the second time slot, the relay amplifies and re-transmits the signal received from the source. We assume that the rate of transmission of the source and that of the relay can be changed, adaptively, by changing the transmission power. We will use the AWGN channel capacity expression in [12] as a measure for the rate of the transmission.

At the beginning of the i-th transmission interval, the source and the relay harvest $E_{\rm s}(i)$ and $E_{\rm r}(i)$ units of energy. We assume that the power usage for purposes other than transmission is negligible compared to the transmission power. At the beginning of the i-th interval, the energy stored in the battery of the source and that of the relay is denoted as $B_{\rm s}(i)$ and $B_{\rm r}(i)$, respectively. The battery of the source (relay) has a limited capacity of $B_{\rm s}^{\max}$ ($B_{\rm r}^{\max}$) units of energy. The data buffer size of the relay is assumed to be infinite so there will be no data loss due to limited data storage capacity in the relay.

During the i-th transmission interval, using power $p_{\rm s}(i)$, the source transmits the signal $x_{\rm s}(i)$ which is given by

$$x_{\rm s}(i) = \sqrt{p_{\rm s}(i)}m(i). \tag{1}$$

Here, $m(i)$ is the i-th transmitted message with $\mathbb{E}\{|m(i)|^2\} = 1$, where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. The signal $y_{\rm r}(i)$ received at the relay and the corresponding SNR can, respectively, be expressed as

$$y_{\rm r}(i) = h_{\rm sr}(i)\sqrt{p_{\rm s}(i)}m(i) + n_{\rm sr}(i) \tag{2}$$

$$\mathrm{SNR}_{\rm r}(i) = \frac{|h_{\rm sr}(i)|^2 p_{\rm s}(i)}{\sigma_{\rm sr}^2} \tag{3}$$

where $n_{\rm sr}(i)$ is the additive Gaussian noise with zero mean and variance $\sigma_{\rm sr}^2$ and $h_{\rm sr}(i)$ is the channel coefficient of the S-R link. In the second time slot of the i-th interval, the relay transmits signal $x_{\rm r}(i)$ with power $p_{\rm r}(i)$ which is given by

$$x_{\rm r}(i) = \sqrt{\frac{p_{\rm r}(i)}{p_{\rm s}(i)|h_{\rm sr}(i)|^2 + \sigma_{\rm sr}^2}}\, y_{\rm r}(i). \tag{4}$$

The signal received at the destination, $y_{\rm d}(i)$, and the corresponding SNR are, respectively, expressed as

$$\begin{aligned} y_{\rm d}(i) =& \alpha(i)\sqrt{p_{\rm s}(i)}h_{\rm rd}(i)h_{\rm sr}(i)m(i) \\ &+ \alpha(i)h_{\rm rd}(i)n_{\rm sr}(i) + n_{\rm rd}(i) \end{aligned} \tag{5}$$

$$\mathrm{SNR}_{\rm d}(i) = \frac{\gamma_{\rm sr}(i)p_{\rm s}(i)\gamma_{\rm rd}(i)p_{\rm r}(i)}{\gamma_{\rm sr}(i)p_{\rm s}(i) + \gamma_{\rm rd}(i)p_{\rm r}(i) + 1} \tag{6}$$

where $h_{\rm rd}(i)$ is the channel coefficient of the relay-destination (R-D) link, $\gamma_{\rm sr}(i) \triangleq |h_{\rm sr}(i)|^2/\sigma_{\rm sr}^2$ and $\gamma_{\rm rd}(i) \triangleq |h_{\rm rd}(i)|^2/\sigma_{\rm rd}^2$. The throughput of the system in the i-th transmission interval is given by

$$R(p_{\rm s}(i), p_{\rm r}(i)) = \frac{1}{2}\log\left(1 + \frac{\gamma_{\rm sr}(i)p_{\rm s}(i)\gamma_{\rm rd}(i)p_{\rm r}(i)}{\gamma_{\rm sr}(i)p_{\rm s}(i) + \gamma_{\rm rd}(i)p_{\rm r}(i) + 1}\right) \tag{7}$$

where the factor 1/2 is due to the half-duplex operation mode in the relay. This concludes our system model.

3. THROUGHPUT MAXIMIZATION USING MDP FORMULATION

In this section, using causal knowledge of fading and that of the harvested energy, we aim to find the optimal transmission policy at the source and at the relay such that the average total throughput of the network is maximized. Although, the optimal solution can be obtained using dynamic programming, such a solution is in general computationally prohibitive to implement since the number of possible states of system at each transmission interval is infinite. To overcome this issues, we use discrete dynamic programming by casting the optimization problem as a Markov decision process (MDP).

We hereafter assume that during each transmission interval, the source has full knowledge of battery levels at both nodes and that of the channel states of both S-R and R-D links. The MDP formulation consists of the following components:

1) *State Space*: The state space $\mathcal{S}$ is defined as

$$\mathcal{S} \triangleq \mathcal{B}_{\rm s} \times \mathcal{B}_{\rm r} \times \mathcal{G}_{\rm sr} \times \mathcal{G}_{\rm rd} \tag{8}$$

where $\mathcal{B}_{\rm s} \triangleq \{0, B_{\rm s}^{\rm max}/n, \cdots, B_{\rm s}^{\rm max}\}$ is the set of possible battery levels at the source and $\mathcal{B}_{\rm r} \triangleq \{0, B_{\rm r}^{\rm max}/n, \cdots, B_{\rm r}^{\rm max}\}$ is the set of battery levels at the relay. The parameter n determines the number of possible battery levels at the source and at the relay. Moreover, $\mathcal{G}_{\rm sr} \triangleq \{g_{\rm sr}^1, g_{\rm sr}^2, \cdots, g_{\rm sr}^m\}$ and $\mathcal{G}_{\rm rd} \triangleq \{g_{\rm rd}^1, g_{\rm rd}^2, \cdots, g_{\rm rd}^m\}$ are the sets of states corresponding to an m-state first-order Markov chain representing the S-R and the R-D links, respectively. The first-order Markov chain is an accurate model for slow fading channels [13]. The size of the resulting state space $\mathcal{S}$ is $(n+1)^2m^2$.

At the i-th transmission interval, the state of the system is given by

$$\mathbf{s}_i = (B_{\rm s}(i), B_{\rm r}(i), G_{\rm sr}(i), G_{\rm rd}(i)) \tag{9}$$

where $B_{\rm s}(i)$ and $B_{\rm r}(i)$ represent the battery levels of the source and that of the relay in i-th transmission interval, respectively, while $G_{\rm sr}(i)$ and $G_{\rm rd}(i)$ are the states of the S-R and R-D channels, respectively.

2) *Action space*: The set of allowable actions associated with the state $\mathbf{s}$, is denoted by $\mathcal{A}^{\mathbf{s}}$ and is defined as

$$\begin{aligned} \mathcal{A}^{\mathbf{s}} \triangleq \{&(0,0), (0, B_{\rm r}^{\rm max}/nT), (B_{\rm s}^{\rm max}/nT, 0), \\ &(B_{\rm s}^{\rm max}/nT, B_{\rm r}^{\rm max}/nT), \cdots, (b_{\rm s}/T, b_{\rm r}/T)\}. \end{aligned} \tag{10}$$

Since choosing the action $\mathbf{a} = (a_{\rm s}, a_{\rm r})$ in the i-th transmission interval is equivalent to setting $p_{\rm s}(i) = a_{\rm s}$ and $p_{\rm r}(i) = a_{\rm r}$, the set $\mathcal{A}^{\mathbf{s}_i}$ is determined by the battery levels of the source and the relay. The action space is defined as

$$\mathcal{A} = \bigcup_{\mathbf{s}\in\mathcal{S}} \mathcal{A}^{\mathbf{s}} \tag{11}$$

We denote the sample space of the random variable corresponding to the harvested energy at the source (relay) by $\mathcal{E}_{\rm s}$ ($\mathcal{E}_{\rm r}$). We assume that $\mathcal{E}_{\rm s}$ and $\mathcal{E}_{\rm r}$ are discrete and finite, that is

$$\mathcal{E}_{\rm s} \triangleq \{0, e_{\rm s}^{\min}, \cdots, ke_{\rm s}^{\min}\}, \qquad \mathcal{E}_{\rm r} \triangleq \{0, e_{\rm r}^{\min}, \cdots, ke_{\rm r}^{\min}\}$$

where $e_{\rm s}^{\min}$ and $e_{\rm r}^{\min}$ are the smallest non-zero energy packets that can be harvested at the source and at the relay, respectively. Throughout the transmission, the energy stored in the source battery and that stored in the relay battery are changing as

$$\begin{aligned} B_{\rm s}(i) =& f_{\rm s}(B_{\rm s}(i-1), p_{\rm s}(i-1), E_{\rm s}(i)) \\ &\triangleq \min\{B_{\rm s}(i-1) - Tp_{\rm s}(i-1) + E_{\rm s}(i), B_{\rm s}^{\rm max}\} \end{aligned} \tag{12}$$

$$\begin{aligned} B_{\rm r}(i) =& f_{\rm r}(B_{\rm r}(i-1), p_{\rm r}(i-1), E_{\rm r}(i)) \\ &\triangleq \min\{B_{\rm r}(i-1) - Tp_{\rm r}(i-1) + E_{\rm r}(i), B_{\rm r}^{\rm max}\} \end{aligned} \tag{13}$$

where $B_{\rm s}(0)$ and $B_{\rm r}(0)$ correspond to the initial energy stored in the source and the relay batteries, respectively. The battery levels and the allowable actions are chosen such that for any action $\mathbf{a} = (a_s, a_r) \in \mathcal{A}, \forall e_s \in \mathcal{E}_s, \forall e_r \in \mathcal{E}_r, \forall b_s \in \mathcal{B}_s$ and $\forall b_r \in \mathcal{B}_r$, we have

$$\min\{b_s - Ta_s + e_s, B_s^{\max}\} \in \mathcal{B}_s, \text{ for } i = 1, \cdots, N$$
$$\min\{b_r - Ta_r + e_r, B_r^{\max}\} \in \mathcal{B}_r, \text{ for } i = 1, \cdots, N.$$

3) *Reward*: The reward function in the i-th transmission interval is the throughput of the system which is a function of the current state of the system state $\mathbf{s}_i$ and the action $\mathbf{a}_i$. It is given in (14).

4) *Transition Probability*: The transition probability, denoted by $p_i(\mathbf{s}'|\mathbf{s}, \mathbf{a})$, is the probability that in the $(i+1)$-th transmission interval, the system will end up in state $\mathbf{s}' = (b'_s, b'_r, g'_{sr}, g'_{rd}) \in \mathcal{S}$, given that in the i-th interval, the system is in state $\mathbf{s} = (b_s, b_r, g_{sr}, g_{rd}) \in \mathcal{S}$ and that the action $\mathbf{a} = (a_s, a_r)$ is taken. This probability is given by

$$\begin{aligned} p_i(\mathbf{s}'|\mathbf{s}, \mathbf{a}) =& \Pr\{\mathbf{s}_{i+1} = \mathbf{s}'|\mathbf{s}_i = \mathbf{s}, \mathbf{a}_i = \mathbf{a}\} \\ =& \Pr\{B_s(i+1) = b'_s|B_s(i) = b_s, p_s(i) = a_s\} \\ &\times \Pr\{B_r(i+1) = b'_r|B_r(i) = b_r, p_r(i) = a_r\} \\ &\times \Pr\{G_{sr}(i+1) = g'_{sr}|G_{sr}(i) = g_{sr}\} \\ &\times \Pr\{G_{rd}(i+1) = g'_{rd}|G_{rd}(i) = g_{rd}\} \end{aligned} \quad (15)$$

where the last two terms are the transition probabilities of the first-order Markov chain which models the fading channel. Using the total probability theorem and by conditioning on the harvested energy, the transition probability in (15) can be re-written as

$$\begin{aligned} p_i(\mathbf{s}'|\mathbf{s}, \mathbf{a}) =& \sum_{e_s \in \mathcal{E}} \Pr\{E_s(i+1) = e_s\} I(f_s(b_s, a_s, e_s) = b'_s) \\ &\times \sum_{e_r \in \mathcal{E}} \Pr\{E_r(i+1) = e_r\} I(f_r(b_r, a_r, e_r) = b'_r) \\ &\times \Pr\{G_{sr}(i+1) = g'_{sr}|G_{sr}(i) = g_{sr}\} \\ &\times \Pr\{G_{rd}(i+1) = g'_{rd}|G_{rd}(i) = g_{rd}\}. \end{aligned} \quad (16)$$

where $I(\cdot)$ is the indicator function and is equal to one if its argument is true and zero otherwise.

A decision rule $\mathbf{d}_i(\mathbf{s}) : \mathcal{S} \to \mathcal{A}^{\mathbf{s}}$ is a function that determines the action to be taken when the system is in state $\mathbf{s}$ at the i-th transmission interval. A policy $\boldsymbol{\pi} = \{\mathbf{d}_1(\mathbf{s}_1), \cdots, \mathbf{d}_N(\mathbf{s}_N)\}$ is a sequence of decision rules for all transmission intervals. If the policy $\boldsymbol{\pi} = \{\mathbf{d}_1(\mathbf{s}_1), \cdots, \mathbf{d}_N(\mathbf{s}_N)\}$ is used, the reward-to-go function at the i-th interval (i.e., the summation of the expected reward from the i-th interval to the last interval) is given as

$$u_i^{\boldsymbol{\pi}}(\mathbf{s}_i) = r_i(\mathbf{s}_i, \mathbf{d}_i(\mathbf{s}_i)) + \sum_{\mathbf{s}' \in \mathcal{S}} p_i(\mathbf{s}'|\mathbf{s}_i, \mathbf{d}_i(\mathbf{s}_i)) u_{i+1}^{\boldsymbol{\pi}}(\mathbf{s}'), \quad \text{for } i = 1, \cdots, N-1 \quad (17)$$

$$u_N^{\boldsymbol{\pi}}(\mathbf{s}_N) = r_N(\mathbf{s}_N, \mathbf{d}_N(\mathbf{s}_N)). \quad (18)$$

Our goal is to find the optimal policy which maximizes the expected total reward of the system over N transmission intervals, i.e., $u_1^{\boldsymbol{\pi}}(\mathbf{s}_1)$. Using (17) and (18), the optimality equations (also known as the Bellman equation) can be written as [14]

$$u_i^*(\mathbf{s}_i) = \max_{\mathbf{a} \in \mathcal{A}^{\mathbf{s}_i}} \left\{ r(\mathbf{s}_i, \mathbf{a}) + \sum_{\mathbf{s}' \in \mathcal{S}} p_i(\mathbf{s}'|\mathbf{s}_i, \mathbf{a}) u_{i+1}^*(\mathbf{s}') \right\}, \quad \text{for } i = 1, \cdots, N-1 \quad (19)$$

$$u_N^*(\mathbf{s}_N) = r_N(\mathbf{s}_N, (\frac{B_s(N)}{T}, \frac{B_r(N)}{T})). \quad (20)$$

Note that (20) implies that in the last transmission interval, the optimal action is to use up all the available energy in the battery of the source and that of the relay.

Since for every state in state space, the action set $\mathcal{A}^{\mathbf{s}}$ is finite and the reward function is bounded, a deterministic Markovian policy satisfying (19) and (20) exists. To find this optimal policy, we use the backward induction algorithm [14] presented in Algorithm 1.

Algorithm 1 The Backward Induction Algorithm

1. Set $i = N$ and

$$u_N^*(\mathbf{s}_N) = r_N(\mathbf{s}_N, (\frac{B_s(N)}{T}, \frac{B_r(N)}{T})) \ \ \forall \mathbf{s}_N \in \mathcal{S}.$$

2. Set $i = i - 1$ and compute $u_i^*(\mathbf{s}_i)$ and $\mathbf{d}_i^*(\mathbf{s}_i)$ for $\forall \mathbf{s}_i \in \mathcal{S}$ as

$$u_i^*(\mathbf{s}_i) = \max_{\mathbf{a} \in \mathcal{A}^{\mathbf{s}_i}} \left\{ r(\mathbf{s}_i, \mathbf{a}) + \sum_{\mathbf{s}' \in \mathcal{S}} p_i(\mathbf{s}'|\mathbf{s}_i, \mathbf{a}) u_{i+1}^*(\mathbf{s}') \right\} \quad (21)$$

$$\mathbf{d}_i^*(\mathbf{s}_i) = \arg\max_{\mathbf{a} \in \mathcal{A}^{\mathbf{s}_i}} \left\{ r(\mathbf{s}_i, \mathbf{a}) + \sum_{\mathbf{s}' \in \mathcal{S}} p_i(\mathbf{s}'|\mathbf{s}_i, \mathbf{a}) u_{i+1}^*(\mathbf{s}') \right\}. \quad (22)$$

3. Stop if $i = 1$ otherwise go to step 2.

An important special case which is worth considering is when the power control in the transmitting nodes (i.e., in the source and the relay) is limited to on-off switching. In such a setting, the action set is binary and it is given by

$$\mathcal{A} = \{\boldsymbol{a}_0 = (0, 0), \boldsymbol{a}_1 = (P_s, P_r)\} \quad (23)$$

where P_s and P_r are fixed transmission powers satisfying $P_s \leq B_s^{\max}/T$ and $P_r \leq B_r^{\max}/T$. In the sequel, we will show that in case of such a binary action space, the optimal actions have a special structure which facilitates the implementation of the backward induction of algorithm. The following theorems are presented for the case of block fading channels, that is, when channel levels in each transmission interval are statistically independent of those in other intervals.

Theorem 1: *In the case of block fading channels and a binary action set, if for some certain state* $\hat{\mathbf{s}}_i = (\hat{b}_s, \hat{b}_r, \hat{g}_{sr}, \hat{g}_{rd})$, $\mathbf{d}_i^*(\hat{\mathbf{s}}_i) = \boldsymbol{a}_1$ *then for* $\mathbf{s}_i \in \{\mathbf{s} = (b_s, b_r, g_{sr}, g_{rd}) \mid b_s = \hat{b}_s,\ b_r = \hat{b}_r,\ g_{sr} \geq \hat{g}_{sr},\ g_{rd} = \hat{g}_{rd}\}$, $\mathbf{d}_i^*(\mathbf{s}_i) = \boldsymbol{a}_1$.

Proof: The proof is omitted due to space limitations. See [15]. ■

With a slight change in the statement, a similar theorem can be obtained for the R-D channel level.

Theorem 2 : *In the case of block fading channels and a binary action set, if for some certain state* $\hat{\mathbf{s}}_i = (\hat{b}_s, \hat{b}_r, \hat{g}_{sr}, \hat{g}_{rd})$, $\mathbf{d}_i^*(\hat{\mathbf{s}}_i) = \boldsymbol{a}_1$ *then for* $\mathbf{s}_i \in \{\mathbf{s} = (b_s, b_r, g_{sr}, g_{rd}) \mid b_s = \hat{b}_s,\ b_r = \hat{b}_r,\ g_{sr} = \hat{g}_{sr},\ g_{rd} \geq \hat{g}_{rd}\}$, $\mathbf{d}_i^*(\mathbf{s}_i) = \boldsymbol{a}_1$.

Proof: The proof is omitted due to space limitations. See [15]. ■

The results of Theorems 1 and 2 can be used to simplify the backward induction algorithm. Indeed, if for a specific state $\hat{\mathbf{s}}$ in the $\mathcal{S}$, the optimal action in the i-th transmission interval is equal to $\boldsymbol{a}_1$, then without the need for any further calculations, the optimal actions for any state belonging to either of the sets $\{\mathbf{s} = (b_s, b_r, g_{sr}, g_{rd}) \mid b_s = \hat{b}_s,\ b_r = \hat{b}_r,\ g_{sr} \geq \hat{g}_{sr},\ g_{rd} = \hat{g}_{rd}\}$ or $\{\mathbf{s} = (b_s, b_r, g_{sr}, g_{rd}) \mid b_s = \hat{b}_s,\ b_r = \hat{b}_r,\ g_{sr} = \hat{g}_{sr},\ g_{rd} \geq \hat{g}_{rd}\}$ can

$$r_i(\mathbf{s}_i, \mathbf{a}_i) = \frac{1}{2} \log \left(1 + \frac{G_{\mathrm{sr}}(i) a_{\mathrm{s}}(i) G_{\mathrm{rd}}(i) a_{\mathrm{r}}(i)}{G_{\mathrm{sr}}(i) a_{\mathrm{s}}(i) + G_{\mathrm{rd}}(i) a_{\mathrm{r}}(i) + 1} \right) \text{ for } i = 1, \cdots, N \tag{14}$$

be set to $\boldsymbol{a}_1$. This significantly speeds up the implementation of the backward induction algorithm.

4. SIMULATION RESULTS

In this section, we examine the numerical performance of the our MDP formulation in two scenarios of general and limited (on-off) power control, and compare the results to that of the harvesting rate (HR) assisted scheme introduced in [9]. The results of the alternating convex search algorithm (ACS) presented in [10] serve as a performance benchmark. For the general scenario $n = 10$ is chosen. Furthermore, it is assumed that $E_{\mathrm{s}}(i)$ and $E_{\mathrm{r}}(i)$ can independently take values from the ternary set $\{0, H, 2H\}$ with equal probability. The battery capacity of the source and that of the relay are set to $5H$.

The simulations are carried out for two channel models: the block fading channel and the correlated channel. In the block fading scenario, the channel coefficients are independently and identically distributed according to an exponential distribution with an average of one. The correlated channel is simulated using the inverse discrete Fourier transform approach introduced in [16]. The maximum Doppler frequency is set to $f_{\mathrm{d}} = 8$Hz and $f_{\mathrm{d}}T = 0.04$. For both scenarios, the channels are modeled as a first order Markov chain with 10 states using the equal-probability steady state-distribution proposed in [17]

For the case of block fading channels, in Fig. 2, the throughput of the system is depicted versus the average energy harvesting rate. It can be clearly seen that the performance of the MDP-based method in the general case approaches the performance of the benchmark method and is superior to the other on-line solution.

Fig. 3 demonstrates the total throughput of the system versus the average energy harvesting rate for different online solutions as well as the performance benchmark in case of a correlated channel. As can seen from this figure, the performance of our MDP formulation approaches that of the benchmark and it is significantly better than the other on-line algorithm.

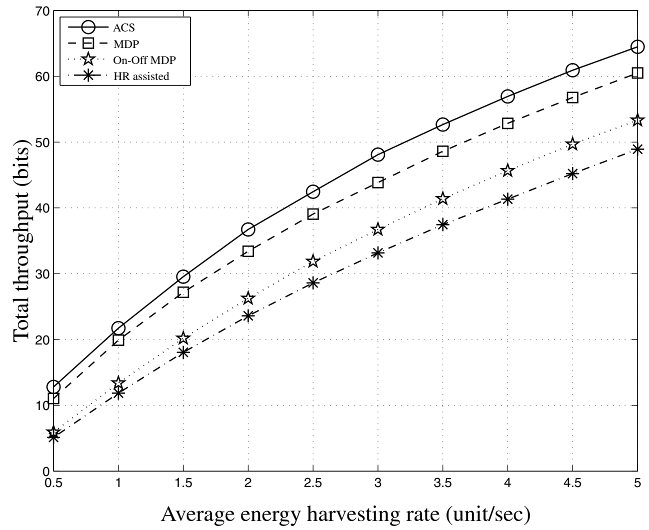

Fig. 2: Throughput curves versus average energy harvesting rate for different online algorithms and for the performance benchmark for block fading channel case.

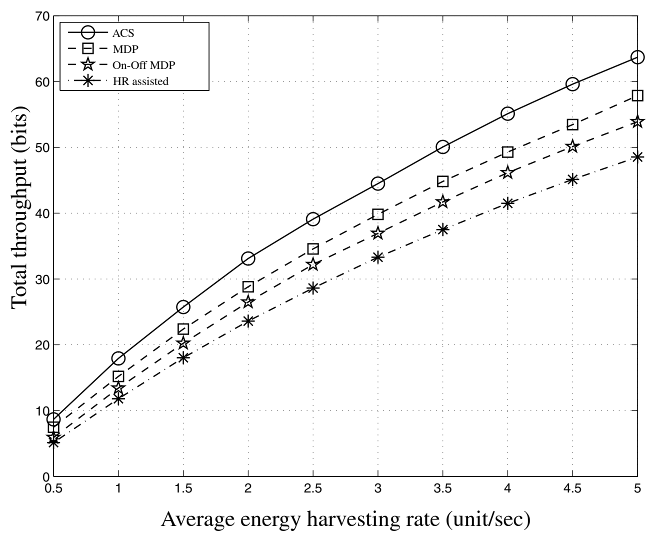

Fig. 3: Throughput curves versus average energy harvesting rate for different online algorithms and for the performance benchmark for correlated channel.

5. CONCLUSION

In this work, we tackled the problem of maximization of the total throughput of an energy-harvesting amplify-and-forward two-hop network in an on-line setting. We proposed using Markov decision process (MDP) to convert the optimization problem into a discrete dynamic programming problem which is a mathematically tractable basis to determine the optimal power-use policies. We also examined an important special case where the transmitters can either be silent (off) or transmit with a fixed power (on). Certain interesting properties of the optimal transmission scheme are derived. These properties can facilitate the implementation of the MDP-based solution. Simulation results demonstrate the superiority of our presented formulation compared to the only existing method which studies the on-line energy harvesting.

6. REFERENCES

[1] O. Ozel and S. Ulukus,   Information-theoretic analysis of an energy harvesting communication system,ƒ in *21st IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops)*, Istanbul, Sep. 2010, pp. 330˘335.

[2] J. Yang and S. Ulukus,   Optimal packet scheduling in an energy harvesting communication system,ƒ *IEEE Transactions on Communications,*, vol. 60, pp. 220˘230, Jan. 2012.

[3] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener,   Transmission with energy harvesting nodes in fading wireless channels: optimal policies,ƒ *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1732˘1743, Sep. 2011.

[4] D. Gunduz and B. Devillers,   Two-hop communication with energy harvesting,ƒ in *Proc. the 4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, San Juan, PR, Dec. 2011, pp. 201˘204.

[5] Y. Luo, Zhang J., and K. B. Letaief,   Optimal scheduling and power allocation for two-hop energy harvesting communication systems,ƒ *IEEE Transactions on Wireless Communications*, vol. 12, pp. 4729˘4741, Sep. 2013.

[6] O. Orhan and E. Erkip,   Energy harvesting two-hop networks: optimal policies for the multi-energy arrival case,ƒ in *Proc. the 35th IEEE Sarnoff Symposium (SARNOFF)*, New Brunswick, May 2012.

[7] C. Huang, Zhang R., and S. Cui,   Optimal resource allocation for Gaussian relay channel with energy harvesting constraints,ƒ in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Mar. 2012, pp. 2817˘2820.

[8] I. Ahmed, A. Ikhlef, Schober R., and R. Mallik,   Power allocation for conventional and buffer-aided link adaptive relaying systems with energy harvesting nodes,ƒ *Arxiv e-prints, http://arxiv.org/abs/1209.2192*, Sep. 2012.

[9] I. Ahmed, A. Ikhlef, Schober R., and R. Mallik,   Joint power allocation and relay selection in energy harvesting AF relay systems,ƒ *IEEE Wireless Communications Letters*, vol. 2, pp. 239˘242, Apr. 2013.

[10] A. Minasian, R. Adve, and S. ShahbazPanahi,   Optimal resource allocation in energy harvesting amplify and forward relay networks,ƒ in *to be presented in 1st IEEE Global Conference on Signal and Information Processing (GLOBALSIP)*, Austin, TX, Dec. 2013.

[11] A. Sinha and P. Chaporkar,   Optimal power allocation for a renewable energy source,ƒ in *Proc. National Conference on Communications (NCC)*, Kharagpur, Feb. 2012.

[12] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley and Sons Inc., 2nd edition, 2006.

[13] P. Sadeghi, R.A. Kennedy, P.B. Rapajic, and R. Shams,   Finite-state Markov modeling of fading channels - a survey of principles and applications,ƒ *IEEE Signal Processing Magazine*, vol. 25, pp. 57˘80, Sep. 2008.

[14] Martin L. Puterman,   *Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1994.

[15] A. Minasian, S. ShahbazPanahi, and R. S. Adve,   Energy harvesting cooperative communication systems,ƒ *Submitted to IEEE Transactions on Wireless Communications*, 2013.

[16] D.J. Young and N.C. Beaulieu,   The generation of correlated Rayleigh random variates by inverse discrete fourier transform,ƒ *IEEE Transactions on Communications,*, vol. 48, pp. 1114˘1127, Jul. 2000.

[17] Hong-Shen W. and N. Moayeri,   Finite-state Markov channel - a useful model for radio communication channels,ƒ *IEEE Transactions on Vehicular Technology,*, vol. 44, pp. 163˘171, Feb. 1995.