

ARQ with Adaptive Feedback for Energy Harvesting Receivers

Yuyi Mao[†], Jun Zhang[†], and K. B. Letaief^{†*}, *Fellow, IEEE*

[†]Dept. of ECE, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

^{*}Hamad bin Khalifa University, Doha, Qatar

Email: {ymaoac, eejzhang, eekhaled}@ust.hk

Abstract—Automatic repeat request (ARQ) is widely used in modern communication systems to improve transmission reliability. In conventional ARQ protocols developed for systems with energy-unconstrained receivers, an acknowledgement/negative-acknowledgement (ACK/NACK) message is fed back when decoding succeeds/fails. Such kind of non-adaptive feedback consumes significant amount of energy, and thus will limit the performance of systems with energy harvesting (EH) receivers. In order to overcome this limitation and to utilize the harvested energy more efficiently, we propose a novel ARQ protocol for EH receivers, where the ACK feedback can be adapted based upon the receiver's EH state. Two conventional ARQ protocols are also considered. By adopting the packet drop probability (PDP) as the performance metric, we formulate the throughput constrained PDP minimization problem for a communication link with a non-EH transmitter and an EH receiver. Optimal reception policies including the sampling, decoding and feedback strategies, are developed for different ARQ protocols. Simulation results will show that the proposed ARQ protocol not only outperforms the conventional ARQs in terms of PDP, but can also achieve a higher throughput.

Index Terms—Green communications, energy harvesting, ARQ, reception policy, feedback strategy, Markov decision process.

I. INTRODUCTION

Energy harvesting (EH) has recently emerged as a promising technique to provide renewable energy sources for wireless systems [1]. As EH devices can capture ambient recyclable energy, such as solar and wind energy, the lifespan of communication systems can be prolonged significantly. However, due to their intermittent and sporadic nature, conventional communication protocols may not be capable of securing full benefits of the harvested energy [2]. Consequently, protocols tailored for EH communication systems have been developed in recent years [3]–[8]. Nevertheless, most of the existing works focus on systems with EH transmitters, and few results are available for EH receivers.

In some scenarios, e.g. communicating over short distances, high data rates can be achieved with a relatively small transmit power, and the energy consumption associated with the complex detection and decoding operations at the receiver becomes dominant [9]. This motivates the recent investigation of communication systems with EH receivers. The optimal packet sampling and decoding policies for an EH receiver to maximize packet throughput were developed in [10]. By jointly considering the EH transmitter and receiver, i.e., a dual

EH link, a general utility maximization problem was solved in [11]. In order to improve transmission reliability, automatic repeat request (ARQ) based packet retransmission has been introduced for systems with EH receivers [12]–[14]. In [12], a dual EH link was modeled as a discrete-time Markov chain, and the packet drop probabilities (PDPs) under basic ARQ and hybrid ARQ with chase combining (HARQ-CC) were analyzed, respectively. A suboptimal power allocation policy was proposed for dual EH links to minimize the PDP under an acknowledgement (ACK) based ARQ in [13]. Moreover, to avoid wasting of the harvested energy, a selective-sampling scheme that allows the EH receiver to sample part of a packet was proposed in [14].

So far, the ARQ protocols adopted for systems with EH receivers, e.g. [12]–[14], are inherited from the conventional ones developed for energy-unconstrained receivers [15], where an ACK message should be fed back when decoding succeeds, and a negative-acknowledgement (NACK) message should be sent back to the transmitter when decoding fails. Such kind of aggressive and non-adaptive feedback strategy consumes significant amount of energy, which may lead to energy outage for sampling and decoding, and thus will limit the performance of systems with energy-scarce EH receivers.

To overcome this limitation, a novel ARQ protocol with adaptive feedback will be proposed in this paper, which enables ACK feedback management and helps utilize the harvested energy more efficiently. As a comparison, two conventional ARQ protocols will also be studied, one without feedback and one with non-adaptive feedback. By adopting the PDP as the performance metric, we formulate the throughput constrained PDP minimization problem for a point-to-point wireless link with a non-EH transmitter and an EH receiver. The optimal reception policies, including the sampling, decoding and feedback strategies, are investigated. For the ARQ protocol without feedback, a myopic policy is shown to be optimal. For ARQs with feedback, an optimal iterative algorithm is developed with the aid of the *Dinkelbach approach* and *Markov decision process* (MDP) techniques. Simulation results demonstrate the effectiveness of the obtained optimal reception policies. It is shown that the proposed ARQ protocol not only outperforms the conventional ARQs in terms of PDP, but it is also competent to achieve a higher throughput.

The organization of this paper is as follows. We introduce the system model and the proposed ARQ protocol in Section II, while the throughput constrained PDP minimization problem is formulated in Section III and the optimal reception

policies for different ARQ protocols are developed in Section IV. Simulation results are presented in Section V, and we will conclude this paper in Section VI.

II. SYSTEM MODEL

We consider a communication link with a non-EH transmitter and an EH receiver. Time is slotted, and the time slot length, system bandwidth and background noise variance are normalized. We assume the channel experiences independent and identically distributed (i.i.d.) block fading among different time slots, and denote the channel power gain at the t th time slot as $|h_t|^2$, with the cumulative distribution function (CDF) given by $F_H(x)$.

A. Energy Model at the Receiver

The EH process is modeled as successive energy packet arrivals, i.e., at the beginning of the t th time slot, $E_{H,t}$ units of energy arrives at the EH receiver. Without loss of generality, we assume $\{E_{H,t}\}$ are i.i.d. discrete random variables, which take values from set $\mathcal{E} \triangleq \{e_n\}$ (e_n 's are the possible amounts of the harvested energy), and distribute according to $\mathbb{P}_{\mathcal{E}}(e)$, $e \in \mathcal{E}$. A battery with capacity B_{\max} is deployed to store the harvested energy, and the battery energy level b_t at the beginning of the t th time slot evolves according to

$$b_{t+1} = \min\{b_t - E_{c,t} + E_{H,t+1}, B_{\max}\}, t = 1, 2, \dots \quad (1)$$

where $b_1 = E_{H,1}$, and $E_{c,t}$ denotes the energy consumed by the receiver in the t th time slot.

The energy spent on sampling and decoding are comparable, and both should be included in the receiver's energy consumption model [10]. We denote the energy consumption for each sampling and decoding operation as E_s and E_d , respectively. The energy consumed for a reliable ACK feedback, denoted by E_f , is also taken into account. The delay incurred by decoding and feedback are ignored for simplicity, and energies are assumed to be integer multiples of a basic energy quantum, E , without loss of generality.

B. Transmission Protocol

The transmitter transmits packets with a constant rate R and a fixed transmit power p_{tx} . A maximum number of transmission attempts for each packet, denoted as K ($K > 1$), is set [8], [12]–[15]. In other words, the transmitter keeps transmitting the same packet either until all transmission attempts are used or the ACK signal is fed back. We denote $k \in \mathcal{K} \triangleq \{0, 1, \dots, K-1\}$ as the transmission index if a packet is in its k th retransmission, i.e., the $(k+1)$ th transmission. When the maximum transmission attempt is used or the ACK is fed back, the transmitter will start to transmit a new packet in the next time slot, which indicates that a packet will be dropped without being received successfully within K transmission attempts. In particular, the packet drop probability (PDP) is adopted as the performance metric in this paper.

C. Reception Policy

At each time slot, the receiver decides whether to sample the received signal, i.e., $a_{s,t} \in \{0, 1\}$, where $a_{s,t} = 1$ means it will perform sampling in the t th time slot, and vice versa. If a packet is sampled, both the pilot symbols for channel training and the information bits are obtained. We assume perfect channel estimation and thus the value of $|h_t|^2$ is available at the receiver. The receiver decodes based on the sample of the packet obtained in the current time slot, similar as in the basic ARQ protocol [12], [13], [15], and it knows if a packet can be successfully decoded based on the knowledge of $|h_t|^2$. Specifically, if $|h_t|^2 \geq (2^R - 1) p_{\text{tx}}^{-1} \triangleq |h_{\text{th}}|^2$, the receiver will correctly decode the packet by consuming E_d units of energy. Otherwise, the sampled packet will be discarded and no energy will be used. When there is ACK feedback, the receiver should further determine the feedback strategy $\{a_{f,t}\}$, where $a_{f,t} = 1$ means an ACK message will be fed back in time slot t for a successfully decoded packet, and $a_{f,t} = 0$ indicates nothing will be fed back. We refer the sampling and feedback strategies as the *reception policy*, which is denoted as $\{\mathbf{a}_t\} \triangleq \{\langle a_{s,t}, a_{f,t} \rangle\}$, and the decoding policy is included implicitly.

D. ARQ Protocols

Before proposing the ARQ protocol with adaptive feedback, we first introduce two baseline ARQ protocols, including ARQ without feedback and ARQ with non-adaptive feedback:

ARQ without feedback: In this protocol, ACK feedback is not allowed, i.e., $a_{f,t} = 0$, and only $a_{s,t}$ needs to be designed ($\mathbf{a}_t \in \{\langle 0, 0 \rangle, \langle 1, 0 \rangle\}$). Thus, the transmitter starts to transmit a new packet every K time slots, regardless of the receiver's operation. This protocol can save the feedback energy, but may reduce the system throughput.

ARQ with non-adaptive feedback: In this protocol, ACK feedback is mandatory in the time slots where packets are successfully decoded, i.e., $\mathbf{a}_t \in \{\langle 0, 0 \rangle, \langle 1, 1 \rangle\}$. Such an ACK feedback strategy is commonly-used in conventional ARQ protocols, e.g. [8], [12]–[15] and references therein.

ARQ with adaptive feedback: In the proposed ARQ protocol with adaptive feedback, the ACKs can be delayed or eliminated for the successfully decoded packets, which is more flexible compared to ARQ with non-adaptive feedback. Hence, $\mathbf{a}_t \in \{\langle 0, 0 \rangle, \langle 0, 1 \rangle, \langle 1, 0 \rangle, \langle 1, 1 \rangle\}$, where $\mathbf{a}_t = \langle 0, 1 \rangle$ means the ACK is delayed for a packet decoded successfully in a previous time slot and is fed back in the t th time slot. Actually, the ARQs without feedback and with non-adaptive feedback can be regarded as special cases of the proposed ARQ protocol.

Remark 1: With conventional energy-unconstrained receivers, delaying or eliminating the ACKs brings no benefit, i.e., ARQ with adaptive feedback reduces to ARQ with non-adaptive feedback. However, with EH receivers, the proposed ARQ protocol offers the option of feedback management, and thus helps improve the efficiency of utilizing the harvested energy. As will be seen in the coming sections, with the optimal reception policies, the proposed ARQ protocol improves both the PDP and throughput performance compared to the two baseline ARQ protocols.

III. PROBLEM FORMULATION

Due to the intermittent and sporadic nature of EH, the reception policies should be carefully designed in order to maximize the system performance. In this section, a throughput constrained PDP minimization problem will be formulated, and its optimal solution will be developed in Section IV.

The PDP is defined as the ratio of the number of dropped packets to the number of transmitted data packets [13] as

$$p_{\text{drop}} \triangleq \lim_{T \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{D_t} \right]}{\mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{N_t} \right]} \stackrel{(\dagger)}{=} \frac{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{D_t} \right]}{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{N_t} \right]}, \quad (2)$$

where $\mathbf{1}_X = 1$ if event X happens, and D_t (N_t) denotes the event of packet drop (starting a new packet) at the t th time slot.¹ Accordingly, the throughput, i.e., the average number of successfully received packets per time slot, is defined as

$$\mathcal{T} \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{S_t} \right], \quad (3)$$

where S_t denotes the event of decoding a packet correctly at time slot t . It can be verified that the following identity holds:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \mathbf{1}_{N_t} \right] = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T (\mathbf{1}_{D_t} + \mathbf{1}_{S_t}) \right]. \quad (4)$$

This is because the transmitter starts to transmit a new packet either if the ACK signal is received or the K th transmission attempt is used.

Denote the system state at the beginning of the t th time slot as \mathbf{s}_t , which can be represented by a triplet, i.e., $\mathbf{s}_t \triangleq \langle b_t, k_t, i_t \rangle$, where $b_t \in \{0, E, \dots, ME\}$ ($ME \triangleq B_{\max}$) is the battery energy level, $k_t \in \mathcal{K}$ is the transmission index, and $i_t \in \{0, 1\}$ denotes the reception state. In particular, $i_t = 0$ indicates the current transmitting packet has not been decoded correctly, and vice versa. A reception policy is mathematically characterized by a mapping from the system state to the action, i.e., $\Psi: \mathbf{s} \rightarrow \mathbf{a}$. We denote $d(\mathbf{s}_t, \mathbf{a}_t)$, $n(\mathbf{s}_t, \mathbf{a}_t)$ and $s(\mathbf{s}_t, \mathbf{a}_t)$ as the expected values of $\mathbf{1}_{D_t}$, $\mathbf{1}_{N_t}$ and $\mathbf{1}_{S_t}$ when the system is in state \mathbf{s}_t while action \mathbf{a}_t is taken. Thus, the throughput constrained PDP minimization problem [15] is formulated as

$$\mathcal{P}_1: \min_{\Psi} \frac{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T d(\mathbf{s}_t, \mathbf{a}_t) \right]}{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T n(\mathbf{s}_t, \mathbf{a}_t) \right]} \quad (5)$$

$$\text{s.t. } \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T s(\mathbf{s}_t, \mathbf{a}_t) \right] \geq \mathcal{T}_{\text{th}} \quad (6)$$

$$\mathbf{a}_t \in \mathcal{A}_{\mathbf{s}_t}, \quad (7)$$

where the expectations are with respect to the sample-path of \mathbf{s}_t and \mathbf{a}_t induced by Ψ . (6) is the throughput constraint while (7) stands for the feasible action set, $\mathcal{A}_{\mathbf{s}_t}$, when the system is in state \mathbf{s}_t . Note that $\mathcal{A}_{\mathbf{s}_t}$ is related to the adopted ARQ protocol, and it is further specified as $\mathcal{A}_{\mathbf{s}_t}^{\text{wo}}$, $\mathcal{A}_{\mathbf{s}_t}^{\text{nA}}$ and $\mathcal{A}_{\mathbf{s}_t}^{\text{A}}$ for the ARQs without feedback, with non-adaptive feedback and with

adaptive feedback, respectively. Generally, \mathcal{P}_1 is a stochastic optimization problem with a fractional objective function, and we will investigate its optimal solutions for different ARQ protocols in the next section.

IV. OPTIMAL RECEPTION POLICIES

In this section, we will investigate the optimal reception policies for the three ARQ protocols. For ARQ without feedback, a myopic policy will be shown to be optimal. For ARQs with (non-adaptive or adaptive) feedback, an optimal iterative algorithm will be proposed.

A. ARQ without ACK Feedback

For ARQ without feedback, the transmitter starts to transmit a new packet every K time slots, i.e., $\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T n(\mathbf{s}_t, \mathbf{a}_t) \right] = 1/K$. Thus, we can minimize the numerator in (5) via optimizing the sampling strategy $\{a_{s,t}\}$. Since only $a_{s,t}$ needs to be designed, we use \mathbf{a}_t and $a_{s,t}$ interchangeably in this subsection. According to (4), it is equivalent to maximize the throughput. Therefore, we consider the following throughput maximization problem

$$\mathcal{P}_2: \max_{\Psi} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T s(\mathbf{s}_t, \mathbf{a}_t) \right] \text{ s.t. } \mathbf{a}_t \in \mathcal{A}_{\mathbf{s}_t}^{\text{wo}}, \quad (8)$$

and the solution for \mathcal{P}_1 can be obtained based on Lemma 1.

Lemma 1: Denote the optimal solution and the optimal value of \mathcal{P}_2 as $\Psi_{\mathcal{P}_2}^*$ and $\mathcal{T}_{\mathcal{P}_2}^*$, respectively. If $\mathcal{T}_{\mathcal{P}_2}^* < \mathcal{T}_{\text{th}}$, \mathcal{P}_1 is infeasible. Otherwise, $\Psi_{\mathcal{P}_2}^*$ is the optimal solution for \mathcal{P}_1 .

Proof: When $\mathcal{T}_{\mathcal{P}_2}^* < \mathcal{T}_{\text{th}}$, \mathcal{P}_1 is infeasible since the maximum throughput is less than \mathcal{T}_{th} . If (6) is added, \mathcal{P}_2 is equivalent to \mathcal{P}_1 . Thus, when $\mathcal{T}_{\mathcal{P}_2}^* \geq \mathcal{T}_{\text{th}}$, (6) holds under $\Psi_{\mathcal{P}_2}^*$, i.e., $\Psi_{\mathcal{P}_2}^*$ is also optimal for \mathcal{P}_1 . ■

In order to solve \mathcal{P}_2 , we first specify the feasible action set $\mathcal{A}_{\mathbf{s}_t}^{\text{wo}}$: If $i_t = 0$ and $b_t \geq E_s + E_d$, we have $\mathcal{A}_{\mathbf{s}_t}^{\text{wo}} = \{0, 1\}$. Otherwise, we have $\mathcal{A}_{\mathbf{s}_t}^{\text{wo}} = \{0\}$, which indicates that the receiver can sample only when the current transmitting packet has not been decoded successfully, and meanwhile, the available energy is sufficient for both the sampling and decoding operations².

When $a_s = 0$, no energy is consumed and the reception state remains unchanged (unless it is in the K th transmission attempt due to the start of a new packet in the next time slot), i.e., no packet is decoded in the current time slot and $s(\mathbf{s}, a_s = 0) = 0$. When $a_s = 1$, the receiver performs sampling, which consumes E_s units of energy. With probability $1 - F_H(|h_{\text{th}}|^2) \triangleq p_c$, the packet will be successfully decoded (the reception state becomes 1 in the next time slot unless it is in the K th transmission) and E_d units of energy will be consumed by the decoder; while with probability $\bar{p}_c = 1 - p_c$, the channel will be in outage and the decoder will stay inactive, i.e., the reception state remains 0. Thus, $s(\mathbf{s}, a_s = 1) = p_c$.

¹(†) holds under the condition that both the limits of the numerator and denominator exist, which are assumed in this work.

²In principle, the receiver can perform sampling after a packet has been decoded correctly. However, this has no contribution to the PDP/throughput performance. It can also sample when $E_s \leq b_t < E_s + E_d$. However, without the decoding capability, the sampling energy is wasted since the receiver decodes based on the sample of a packet obtained in the current time slot. Without loss of optimality, these two actions have been precluded in $\mathcal{A}_{\mathbf{s}_t}^{\text{wo}}$.

Theoretically, \mathcal{P}_2 is an MDP problem and can be solved by standard algorithms, which, however, will have high complexity. We propose a myopic policy for \mathcal{P}_2 , for which the receiver performs sampling when the available energy is sufficient for sampling and decoding, while the current transmitting packet has not been received correctly, i.e.,

$$a_{s,t} = \mathbf{1}_{i_t=0, b_t \geq E_s + E_d}, \quad (9)$$

which is optimal for \mathcal{P}_2 as shown in the following proposition.

Proposition 1: The myopic policy in (9) is optimal for \mathcal{P}_2 .

Proof: Please refer to the Appendix. ■

Under mild conditions, the system state constitutes an ergodic Markov chain under the myopic policy. Thus, the steady state distribution $\{\pi_s\}$ can be obtained by solving the balance equation, and the maximum throughput is given by $\mathcal{T}_{\mathcal{P}_2}^* = \sum_{b \geq E_s + E_d, i=0} \pi_s p_c$. Together with Lemma 1, we can either identify the infeasibility or the optimal solution for \mathcal{P}_1 .

B. ARQ with ACK Feedback

The reception policies design for the ARQ protocols with feedback, either non-adaptive or adaptive, is much more challenging due to the fractional objective function, as the value of $\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T n(s_t, \mathbf{a}_t) \right]$ is unknown. To develop the optimal policies, we first transform \mathcal{P}_1 into \mathcal{P}_3 with a weighted linear objective function, and reveal their relationship in Lemma 2.

Lemma 2: If \mathcal{P}_1 is feasible, the minimum PDP, p_{drop}^* , is achieved if and only if the optimal value of the following weighted minimization problem \mathcal{P}_3 is zero when $q = p_{\text{drop}}^*$:

$$\begin{aligned} \mathcal{P}_3 : \min_{\Psi} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T (d(s_t, \mathbf{a}_t) - q \cdot n(s_t, \mathbf{a}_t)) \right] \\ \text{s.t. (6), (7).} \end{aligned}$$

Moreover, the optimal solution for \mathcal{P}_3 with parameter p_{drop}^* , $\Psi_{\mathcal{P}_3}^*$, is also optimal for \mathcal{P}_1 .

Proof: The proof is similar to that for Theorem 1 in [16], which is omitted due to space limitation. ■

As $\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T n(s_t, \mathbf{a}_t) \right] \in [1/K, 1]$, the objective function in \mathcal{P}_3 is strictly monotonic decreasing with q , which suggests an iterative algorithm to obtain p_{drop}^* and $\Psi_{\mathcal{P}_1}^*$ (known as the Dinkelbach approach [17]), as summarized in Algorithm 1. Note that if \mathcal{P}_3 can be solved optimally, Algorithm 1 is guaranteed to converge to the optimal solution, which can be proved by following a similar approach as in [16], [17]. Next, we will develop the optimal solution for \mathcal{P}_3 and focus on the ARQ protocol with adaptive feedback. As a special case, the solution for ARQ with non-adaptive feedback can be obtained with minor modifications.

In order to solve \mathcal{P}_3 , we first specify the feasible action set $\mathcal{A}_{s_t}^A$: When $i_t = 0, E_s + E_d \leq b_t < E_s + E_d + E_f$, we have $\mathcal{A}_{s_t}^A = \{(0, 0), \langle 1, 0 \rangle\}$, which corresponds to the scenario that the packet has not been decoded correctly, while the battery energy is only sufficient for sampling and decoding, i.e., the ACK cannot be fed back due to energy shortage; when $i_t = 0, b_t \geq E_s + E_d + E_f$, we have $\mathcal{A}_{s_t}^A = \{(0, 0), \langle 1, 0 \rangle, \langle 1, 1 \rangle\}$;

Algorithm 1 Optimal Iterative Algorithm for \mathcal{P}_1 for ARQ with Adaptive Feedback.

```

1: Initialize the maximum number of iterations  $I_{\max}$  and the
   maximum tolerance  $\Delta > 0$ , set  $q = 1$  and  $n = 0$ ;
2: repeat
3:   if  $\mathcal{P}_3$  is feasible then
4:     Solve  $\mathcal{P}_3$  for a given  $q$  and obtain the optimal solution
        $\Psi_{\mathcal{P}_3}^*$  and the optimal value  $\mathcal{W}_q^*$ ;
5:   else
6:     Break;
7:   end if
8:   if  $\mathcal{W}_q^* \geq -\Delta$  then
9:     Convergence = true and return  $\Psi_{\mathcal{P}_1}^* = \Psi_{\mathcal{P}_3}^*$ ;
10:  else
11:    Set  $q = \frac{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^{\Psi_{\mathcal{P}_3}^*} [\sum_{t=1}^T d(s_t, \mathbf{a}_t)]}{\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^{\Psi_{\mathcal{P}_3}^*} [\sum_{t=1}^T n(s_t, \mathbf{a}_t)]}$  and  $n = n + 1$ ;
12:    Convergence = false;
13:  end if
14: until Convergence = true or  $n = I_{\max}$ 

```

when $i_t = 1, b_t \geq E_f$, we have $\mathcal{A}_{s_t}^A = \{(0, 0), \langle 0, 1 \rangle\}$ since the packet has been well received; otherwise, $\mathcal{A}_{s_t}^A = \{(0, 0)\}$.

Denote the state transition probability as $\mathbb{P}(s'|s, \mathbf{a})$, which is the probability the system will be in state s' in the next time slot, given the current state is s and action \mathbf{a} is taken. We derive the expressions of $\mathbb{P}(s'|s, \mathbf{a})$, $d(s, \mathbf{a})$, $n(s, \mathbf{a})$ and $s(s, \mathbf{a})$ in the following lemma.

Lemma 3: For the ARQ protocol with adaptive feedback, $n(s, \mathbf{a}) = \mathbf{1}_{k=0}$, $s(s, \mathbf{a}) = p_c \mathbf{1}_{i=0, a_s=1}$ and $d(s, \mathbf{a}) = \mathbf{1}_{i=0, k=K-1} (\bar{p}_c \mathbf{1}_{a_s=1} + \mathbf{1}_{a_s=0})$. And $\mathbb{P}(s'|s, \mathbf{a})$ is given as

$$\begin{aligned} \mathbb{P}(s'|s, \mathbf{a} = \langle 0, 0 \rangle) &= \begin{cases} \mathbf{1}_{k'=(k+1) \bmod K}, & i = i', k < K-1 \text{ or} \\ \sum_{\forall e} \mathbf{1}_{b'=\phi(b, 0, e)} \mathbb{P}_{\mathcal{E}}(e), & i' = 0, k = K-1 \\ 0, & \text{else,} \end{cases} \\ \mathbb{P}(s'|s, \mathbf{a} = \langle 1, 0 \rangle) &= \begin{cases} \sum_{\forall e} [\mathbf{1}_{i'=0, b'=\phi(b, E_s, e)} \bar{p}_c + & i = 0, b \geq E_s + E_d, \\ \mathbf{1}_{i'=1, b'=\phi(b, E_s + E_d, e)} p_c] \mathbb{P}_{\mathcal{E}}(e), & k < K-1, k' = k+1 \\ \mathbf{1}_{i'=0} \sum_{\forall e} [\mathbf{1}_{b'=\phi(b, E_s, e)} \bar{p}_c & i = 0, b \geq E_s + E_d, \\ + \mathbf{1}_{b'=\phi(b, E_s + E_d, e)} p_c] \mathbb{P}_{\mathcal{E}}(e), & k = K-1, k' = 0 \\ 0, & \text{else,} \end{cases} \\ \mathbb{P}(s'|s, \mathbf{a} = \langle 1, 1 \rangle) &= \begin{cases} \sum_{\forall e} [\mathbf{1}_{k'=0, b'=\phi(b, E_s + E_d + E_f, e)} p_c + & i = i' = 0, \\ \mathbf{1}_{k'=(k+1) \bmod K, b'=\phi(b, E_s, e)} \bar{p}_c] \mathbb{P}_{\mathcal{E}}(e), & b \geq E_s + E_d + E_f \\ 0, & \text{else,} \end{cases} \\ \mathbb{P}(s'|s, \mathbf{a} = \langle 0, 1 \rangle) &= \begin{cases} \sum_{\forall e} \mathbf{1}_{k'=0, b'=\phi(b, E_f, e)} \mathbb{P}_{\mathcal{E}}(e), & i = 1, i' = 0, \\ 0, & k > 0, b \geq E_f, \\ & \text{else,} \end{cases} \end{aligned}$$

where $\phi(b, u, e) \triangleq \min\{b - u + e, B_{\max}\}$.

Proof: By definition, $n(s, \mathbf{a}) = \mathbf{1}_{k=0}$ holds. As discussed in Section IV-A, when $a_s = 1$, the packet will be decoded correctly with probability p_c , while $a_s = 0$, no packet will be decoded, which lead to the expressions of $s(s, \mathbf{a})$ and $d(s, \mathbf{a})$.

The state transition when $\mathbf{a}_t = \langle 0, 0 \rangle$ and $\langle 1, 0 \rangle$ are similar to those in ARQ without feedback. When $\mathbf{a}_t = \langle 1, 1 \rangle$, if $|h_t|^2 \geq |h_{th}|^2$, the packet can be decoded correctly and $E_s + E_d + E_f$ units of energy will be consumed which accounts for sampling, decoding and feedback, i.e., both k and i will be 0 in the next time slot. When $\mathbf{a}_t = \langle 0, 1 \rangle$, E_f units of energy will be consumed and the transmitter will transmit a new packet in the next time slot. Detailed derivation for $\mathbb{P}(\mathbf{s}'|\mathbf{s}, \mathbf{a})$ is omitted due to space limitation. ■

It is not difficult to identify that \mathcal{P}_3 is a constrained MDP (CMDP) problem. In the following, we provide a linear programming (LP) approach to obtain the optimal solution for \mathcal{P}_3 by solving the following LP problem \mathcal{P}_4 [18]:

$$\mathcal{P}_4 : \min \sum_{\mathbf{s} \in \mathcal{A}_s} \sum_{\mathbf{a} \in \mathcal{A}_s^A} c(\mathbf{s}, \mathbf{a}) x(\mathbf{s}, \mathbf{a}) \quad (10)$$

$$\text{s.t.} \quad \sum_{\mathbf{a} \in \mathcal{A}_s^A} x(\mathbf{s}', \mathbf{a}) = \sum_{\mathbf{s} \in \mathcal{A}_s} \sum_{\mathbf{a} \in \mathcal{A}_s^A} \mathbb{P}(\mathbf{s}'|\mathbf{s}, \mathbf{a}) x(\mathbf{s}, \mathbf{a}), \forall \mathbf{s}' \quad (11)$$

$$\sum_{\mathbf{s} \in \mathcal{A}_s} \sum_{\mathbf{a} \in \mathcal{A}_s^A} s(\mathbf{s}, \mathbf{a}) x(\mathbf{s}, \mathbf{a}) \geq \mathcal{T}_{th} \quad (12)$$

$$\sum_{\mathbf{s} \in \mathcal{A}_s} \sum_{\mathbf{a} \in \mathcal{A}_s^A} x(\mathbf{s}, \mathbf{a}) = 1, 0 \leq x(\mathbf{s}, \mathbf{a}) \leq \mathbf{1}_{\mathbf{a} \in \mathcal{A}_s^A}, \forall \mathbf{s}, \mathbf{a} \quad (13)$$

where $c(\mathbf{s}, \mathbf{a}) \triangleq d(\mathbf{s}, \mathbf{a}) - q \cdot n(\mathbf{s}, \mathbf{a})$, and $x(\mathbf{s}, \mathbf{a})$ is the occupation measure that gives the steady state probability that the system is in state \mathbf{s} and action \mathbf{a} is chosen. (10) is an equivalent representation of $\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T c(\mathbf{s}_t, \mathbf{a}_t) \right]$, while (11) is the balance equation. With the optimal solution to \mathcal{P}_4 , $x^*(\mathbf{s}, \mathbf{a})$, we can construct an optimal stationary and randomized policy to \mathcal{P}_3 as $\Psi_{\mathcal{P}_3}^* : \mathbf{s} \rightarrow \mathbf{a}$ with probability $x^*(\mathbf{s}, \mathbf{a}) / \sum_{\mathbf{a} \in \mathcal{A}_s^A} x^*(\mathbf{s}, \mathbf{a})$ (If $\sum_{\mathbf{a} \in \mathcal{A}_s^A} x^*(\mathbf{s}, \mathbf{a}) = 0$, \mathbf{a} can be an arbitrary feasible action and set to be $\langle 0, 0 \rangle$). Besides, the update of q in line 11 of Algorithm 1 can be rewritten as $q = \sum_{\mathbf{s}, \mathbf{a} \in \mathcal{A}_s^A} d(\mathbf{s}, \mathbf{a}) x^*(\mathbf{s}, \mathbf{a}) / \sum_{\mathbf{s}, \mathbf{a} \in \mathcal{A}_s^A} n(\mathbf{s}, \mathbf{a}) x^*(\mathbf{s}, \mathbf{a})$.

Remark 2: For systems with an energy-unconstrained receiver, the myopic policy is optimal, i.e., $\mathbf{a}_t = \langle 1, 0 \rangle$ and $\langle 1, 1 \rangle, \forall t$ for ARQs without feedback and with non-adaptive feedback, respectively. Based on the renewal theory, the PDPs achieved by ARQs without feedback and with non-adaptive feedback are the same as \bar{p}_c^K , while the throughput achieved by ARQ with non-adaptive feedback is p_c , which is greater than $(1 - \bar{p}_c^K) / K$ achieved by the one without feedback.

V. SIMULATION RESULTS

In simulations, we assume the channel is Rayleigh fading with unit variance, $R = 0.5$ bps/Hz, and $p_{tx} = 1$ W. The EH process is modeled as an i.i.d. Bernoulli process with EH probability ρ (also termed as the normalized EH rate) [5], i.e., $\mathcal{E} = \{0, \hat{e}\}$, $\mathbb{P}_{\mathcal{E}}(0) = 1 - \rho$ and $\mathbb{P}_{\mathcal{E}}(\hat{e}) = \rho$. We use $\hat{e} = 6E$ as an example, and set $B_{max} = 15E$, $E_d = E_s = 3E$, $I_{max} = 20$, and $\Delta = 10^{-6}$. For comparison, we introduce the myopic policies for the ARQ protocols with feedback: For ARQ with non-adaptive feedback, $\mathbf{a}_t = \langle 1, 1 \rangle$ when $i_t = 0, b_t \geq E_s + E_d + E_f$; while for ARQ with adaptive feedback, $\mathbf{a}_t = \langle 1, 0 \rangle$ when $i_t = 0, E_s + E_d \leq b_t < E_s + E_d + E_f$, $\mathbf{a}_t = \langle 1, 1 \rangle$ when $i_t = 0, b_t \geq E_s + E_d + E_f$, and $\mathbf{a}_t = \langle 0, 1 \rangle$ when $i_t = 1, b_t \geq E_f$.

The PDPs achieved by different ARQs are shown in Fig. 1. In general, a higher EH rate is needed in order to meet

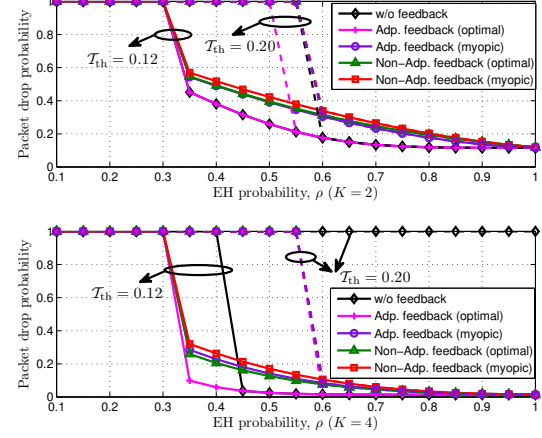


Fig. 1. PDP vs. EH probability ρ , $E_f = E$. We set the PDP to 1 if the throughput constraint can not be satisfied.

a higher throughput requirement. Also, with a larger K , i.e., the maximum number of transmission attempts, the system has the potential to achieve a smaller PDP, but the throughput requirement becomes more difficult to satisfy. Besides, when the throughput constraint is met, ARQ without feedback and ARQ with adaptive feedback achieve the same performance under the optimal policies. This is because each packet uses all the transmission attempts to avoid packet drop, and the proposed ARQ protocol reduces to the ARQ without feedback. From this set of results, we can draw new design insights for communication systems with EH receivers:

- In contrast to systems with energy-unconstrained receivers, for which the myopic policies are optimal, with EH receivers, noticeable performance gains are provided by the optimal reception policies for the ARQs with feedback compared to the myopic policies. This shows the significant difference in retransmission policy design with EH receivers.
- When the required throughput is achieved, the ARQ without feedback outperforms the one with non-adaptive feedback, which is different from systems with energy-unconstrained receivers, where both protocols achieve the same PDP as discussed in Remark 2. This is due to the *feedback energy consumption* and the *waste of transmission attempts* incurred by energy shortage at the receiver. Thus ARQ without feedback may be beneficial in certain scenarios with EH receivers.
- Compared to the ARQ protocol with non-adaptive feedback, significant performance improvement is achieved by the proposed ARQ protocol, and a smaller EH rate is needed to meet the throughput requirement. These confirm the importance of intelligent feedback strategies for EH receivers and reveal the difference from systems with energy-unconstrained receivers, where the proposed ARQ protocol reduces to the one with non-adaptive feedback as emphasized in Remark 1.

In Fig. 2, we show the achievable success probability³-

³Success probability $\triangleq 1 - \text{PDP}$.

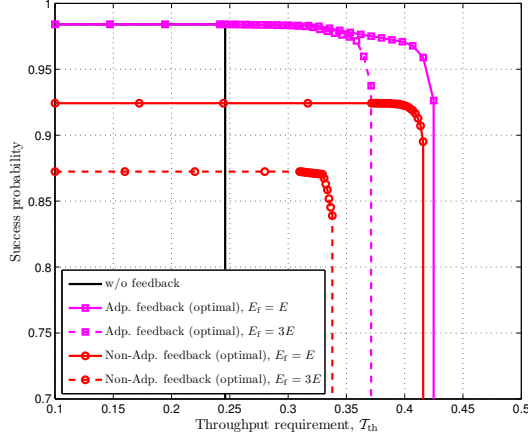


Fig. 2. Success probability vs. \mathcal{T}_{th} , $K = 4$ and $\rho = 0.6$.

throughput (S-T) regions for different ARQ protocols. Key observations can be drawn:

- The S-T region of the ARQ without feedback is a rectangle since any PDP-optimal policy is also throughput-optimal as discussed in Section IV-A.
- For ARQs with feedback, a higher value of E_f leads to a smaller S-T region, and there exists a tradeoff between the achievable success probability and throughput, i.e., a proper operating point should be chosen to balance the system's reliability and efficiency.
- With a small value of \mathcal{T}_{th} , ARQ with adaptive feedback reduces to the one without feedback, similar as the case in Fig. 1; while with a large value of \mathcal{T}_{th} , the proposed ARQ protocol not only enjoys a higher success probability, but it is also competent to meet the throughput requirements compared to ARQ with non-adaptive feedback.

VI. CONCLUSIONS

In this paper, we proposed a novel ARQ protocol with adaptive feedback for EH receivers, which offers the option of ACK feedback management and helps utilize the harvested energy more efficiently. The throughput constrained packet drop probability minimization problem was investigated, and the optimal reception policies were developed. Simulation results demonstrated the benefits of adaptive feedback in minimizing the wastage of the harvested energy and transmission attempts. In addition, this investigation has revealed new design insights for communication systems with EH receivers, and indicated the importance of a comprehensive consideration of different components of energy consumption.

APPENDIX

Proof for Proposition 1: The optimal $a_{s,t}$ should maximize the long-term average expected throughput, i.e., $a_{s,t}^* = \arg \max_{a \in \mathcal{A}_{s,t}^{\text{wo}}} \lim_{T \rightarrow \infty} \frac{V_T^a}{T-t+1}$, where $V_T^a \triangleq s(s_t, a) + \mathbb{E} \left[\sum_{\tau=t+1}^T s(s_\tau, a_{s,\tau}^*) | s_t, a \right]$. The first term in V_T^a denotes the expected number of successfully decoded packets in time slot t , while the second term stands for the expected number of total successfully decoded packets from the $t+1$ to the T th

time slot under the optimal policy $\{a_{s,\tau}^*\}$. In the following, we will show if $a_{s,t} = 1 \in \mathcal{A}_{s,t}^{\text{wo}}$, then $V_T^1 \geq V_T^0, \forall T > t$.

Suppose $a_{s,t} = 0$ and let $t_1 = \min\{\tau \geq t+1 | \hat{a}_{s,\tau}^* = 1\}$, where $\hat{a}_{s,\tau}^*, t+1 \leq \tau \leq T$, is a sample-path of the optimal sampling policy given $a_{s,t} = 0$. For convenience, $t_1 \triangleq T+1$ if $\hat{a}_{s,\tau}^* = 0, \forall \tau \leq T$. Note that t_1 is random and takes value from $\{t+1, \dots, T+1\}$. Thus, $V_T^0 = \mathbb{E}_{t_1} [s(s_{t_1}, 1) \cdot \mathbf{1}_{t_1 \leq T} + V_{T,t_1}^0]$, where $V_{T,t_1}^0 = \mathbb{E} \left[\sum_{\tau=t_1+1}^T s(s_\tau, \hat{a}_{s,\tau}^*) | s_{t_1}, \hat{a}_{s,\tau}^*, \tilde{\tau} \leq t_1 \right] \cdot \mathbf{1}_{t_1 < T}$.

Suppose $a_{s,t} = 1$, then $a_{s,\tau} = 0, \forall t < \tau \leq t_1$ is feasible for a given t_1 . Therefore, $V_T^1 \geq s(s_t, 1) + \mathbb{E}_{t_1} [V_{T,t_1}^1]$, where $V_{T,t_1}^1 = \mathbb{E} \left[\sum_{\tau=t_1+1}^T s(s_\tau, \tilde{a}_{s,\tau}^*) | s_t, a_{s,\tau}, \tilde{\tau} \leq t_1 \right] \cdot \mathbf{1}_{t_1 < T}$, and $\tilde{a}_{s,\tau}^*, t_1+1 \leq \tau \leq T$, denotes the optimal sampling policy given $a_{s,t} = 1$ and $a_{s,\tau} = 0, t < \tau \leq t_1$.

By expanding V_{T,t_1}^0 and V_{T,t_1}^1 conditioned on the events S_{t_1} and S_t , respectively, and utilizing the fact that B_{\max} is finite, we have $s(s_t, 1) + V_{T,t_1}^1 \geq s(s_{t_1}, 1) \cdot \mathbf{1}_{t_1 \leq T} + V_{T,t_1}^0, \forall t_1$. By taking the expectation with respect to t_1 for both sides of the inequality, we can conclude $V_T^1 \geq V_T^0, \forall T > t$, i.e., $a_{s,t} = 1$ performs no worse than $a_{s,t} = 0$, which ends the proof.

REFERENCES

- [1] Y. Mao, Y. Luo, J. Zhang, and K. B. Letaief, "Energy harvesting small cell networks: Feasibility, deployment and operations," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 94-101, Jun. 2015.
- [2] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, and K. Huang, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360-381, Mar. 2015.
- [3] 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, Sep. 2011.
- [4] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 322-332, Feb. 2014.
- [5] Y. Mao, J. Zhang, S. H. Song, and K. B. Letaief, "Joint link selection and relay power allocation for energy harvesting relaying systems," in *Proc. IEEE Globecom*, Austin, TX, Dec. 2014.
- [6] Y. Mao, J. Zhang, and K. B. Letaief, "A Lyapunov optimization approach for green cellular networks with hybrid energy supplies," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2463-2477, Dec. 2015.
- [7] —, "Grid energy consumption and QoS tradeoff in hybrid energy supply wireless networks," *IEEE Trans. Wireless Commun.*, to appear.
- [8] A. Aprem, C. R. Murthy, N. B. Mehta, "Transmit power control policies for energy harvesting sensors with retransmissions," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 5, pp. 895-906, Oct. 2013.
- [9] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349-2360, Sep. 2005.
- [10] R. Yates and H. Mahdavi-Doost, "Energy harvesting receivers: Packet sampling and decoding policies," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 558-570, Mar. 2015.
- [11] K. Tutuncuoglu and Y. Yener, "Communicating with energy harvesting transmitters and receivers," in *Proc. 2012 Inf. Theory and Appl. Workshop (ITA)*, San Diego, CA, Feb. 2012.
- [12] M. K. Sharma and C. R. Murthy, "Packet drop probability analysis of ARQ and HARQ-CC with energy harvesting transmitters and receivers," in *Proc. IEEE Global Conf. Signal and Inf. Process. (GlobalSIP)*, Atlanta, GA, Dec. 2014.
- [13] S. Zhou, T. Chen, W. Chen, and Z. Niu, "Outage minimization for a fading wireless link with energy harvesting transmitter and receiver," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 496-511, Mar. 2015.
- [14] A. Yadav, M. Goonewardena, W. Ajib, and H. Elbiaze, "Novel retransmission scheme for energy harvesting transmitter and receiver," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, UK, Jun. 2015.
- [15] P. Wu and N. Jindal, "Coding versus ARQ in fading channels: How reliable should the PHY be?," *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3363-3374, Dec. 2011.

- [16] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 2572-2585, Jul. 2012.
- [17] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, pp. 492-498, Mar. 1967.
- [18] E. Altman, *Constrained Markov Decision Processes*, Chapman and Hall/CRC, 1999.