Dynamic Resource Allocation for Multiple-Antenna Wireless Power Transfer

Gang Yang, *Student Member, IEEE*, Chin Keong Ho, *Member, IEEE*, and Yong Liang Guan, *Member, IEEE*

Abstract—We consider a point-to-point multiple-input-single-output (MISO) system where a receiver harvests energy from a transmitter. To achieve high-efficiency wireless power transfer (WPT), the transmitter performs energy beamforming by using an instantaneous channel state information (CSI). The CSI is estimated at the receiver by training via a preamble and fed back to the transmitter. In this paper, we address the key challenge of balancing the time resource used for channel estimation and WPT to maximize the harvested energy and also investigate the allocation of energy resource used for WPT. First, we consider the general scenario where the preamble length is allowed to vary dynamically depending on channel conditions. The optimal preamble length is obtained online by solving a dynamic programming (DP) problem. The DP problem is proved to reduce to an optimal stopping problem. The optimal policy is then shown to depend only on the channel estimate power. Next, we consider the scenario in which the preamble length is fixed by an offline optimization. Furthermore, we derive the optimal power allocation schemes. For the dynamic-length-preamble scenario, the power is allocated according to both the optimal preamble length and the channel estimate power, while for the fixed-length-preamble scenario, the power is allocated according to only the channel estimate power. By numerical simulations, our results show that with optimal power allocation, the energy harvested by using the optimized fixed-length preamble is close to that harvested by using a dynamic-length preamble, hence allowing a low-complexity yet close-to-optimal WPT system to be implemented in practice.

Index Terms—Wireless power transfer, energy beamforming, resource allocation, dynamic channel estimation, dynamic programming, power allocation.

I. INTRODUCTION

RECENTLY, wireless power transfer (WPT) is gaining more and more attention from both academia and industry. Although the traditional near-field inductive coupling WPT and resonant coupling WPT have high efficiency, they can only transfer power over a short distance [1]. Furthermore, these WPT methods may have limited support for power multicast or mobility of the power receiver, hence restricting their

application potential. In comparison, far-field WPT via radiated electromagnetic (EM) waves can transfer power to multiple static or moving receivers over a longer distance (typically several to tens of metres), and thus enables various engineering applications [2]. In particular, the far-field WPT is promising to address energy and lifetime bottlenecks for power-limited devices in wireless networks [3]–[5]. For example, in an energy harvesting sensor network, sensors can harvest energy from the ambient or dedicated EM waves to power themselves for data transmission by various schemes, such as wireless compressive sensing proposed in [6].

Since EM waves decay quickly over distance, they have to be concentrated into a narrow beam via multiple antennas to achieve efficient power transfer. This is referred to as *energy beamforming* [7], which was first considered for simultaneous wireless information and power transfer (SWIPT) in multiuser downlink in [7]. Assuming perfect channel state information (CSI) at the transmitter, [8] investigated the joint optimization of transmit power control, information and power transfer scheduling; [9] studied resource allocation algorithms for SWIPT in broadband wireless systems.

With the assumption of perfect CSI, the uplink wireless information transfer (WIT) powered by downlink WPT was considered in [10], [11]. A harvest-then-transmit protocol was proposed in [10], where all users first harvest the wireless energy in the downlink and then send independent information in the uplink by time-division-multiple-access (TDMA). The sum throughput was maximized by jointly optimizing the time allocation for the downlink WPT and uplink WIT. [11] proposed a wireless-powered communication network with a full-duplex access point (AP) and multiple users. The AP implements full-duplex operation through two antennas: one for broadcasting wireless energy to users in the downlink and one for simultaneously receiving uplink information from users via TDMA. Under an energy causality constraint, the authors investigated the problems of maximizing the sum throughput and minimizing the total time required for each user sending given amount of data back to the AP.

The knowledge of CSI is an essential prerequisite for both energy beamforming and information decoding. For instance, [8] showed that the rate-energy tradeoff in SWIPT systems degrades as the CSI accuracy decreases. Typically, the receiver needs to perform channel estimation and feed back CSI to the transmitter before power transfer. In practice, perfect CSI at the transmitter is not available due to various factors such as time-varying channel, inaccurate channel estimation, quantization error and feedback error. When the channel uncertainty is considered as deterministic and norm bounded,

robust beamforming design was studied in [12] for a multiple-input-single-output (MISO) system with SWIPT, in [13] for a two-way relay system with SWIPT, and in [14] for secure communication in a multiuser SWIPT system. In [12], the harvested energy was maximized for the worst-channel realization, while guaranteeing that the information rate is above a threshold for all possible channel realizations. However, the actual worst case may occur with a very low probability. Hence, this worst-case approach may be overly conservative and therefore, leads to unnecessary performance degradation. When the CSI errors are instead considered as Gaussian random variables, the energy beamforming was studied in [15] for a SWIPT multicast system, and in [16] for a single-user MISO system. [15] proposed a stochastic beamforming scheme to achieve more balanced outage-constrained achievable rates among multiple information receivers. [16] derived the optimal time duration which maximizes an upper-bound rate of uplink WIT powered by downlink WPT. Recently, the energy beamforming with one-bit feedback was also studied in [17] for multiuser multiple-input-multiple-output (MIMO) WPT systems. It proposed a channel learning approach, which requires each user to send back one bit to the transmitter to indicate the change of the harvested energy between the present and previous feedback intervals.

Energy beamforming based on more accurate CSI contributes to higher efficiency of power transfer. The receiver, however incurs significant time (overhead) to obtain the accurate CSI. Longer time duration for channel estimation leads to more accurate CSI available at the transmitter, but also shortens the WPT duration, which may lead to less harvested energy. To maximize the harvested energy, there is thus a design freedom, namely the time spent for estimating the channel. Moreover, to improve the overall system energy efficiency, the amount of energy used for WPT should be optimized, for example, less energy is used for severely-fading channels. Nevertheless, to the best of our knowledge, there does not exist any work that takes into account the preamble overhead and energy allocation for WPT via energy beamforming.

We consider a frame-based MISO system in which the transmitter performs energy beamforming using imperfect CSI fed back from the receiver. The frame is divided into four phases as shown in Fig. 1: the channel estimation (CE) phase, the feedback phase, the wireless power transfer (WPT) phase, as well as the general energy utilization (EU) phase. In this paper, we focus on efficient wireless power transfer; and the particular use of the harvested energy, such as for uplink WIT or sensing [10], [16], [18] is not considered in this work. The feedback is assumed to be error-free and take negligible time, and is thus ignored in the analysis. The time duration for the EU phase is fixed. Unlike previous work on robust beamforming in [12]–[14], we maximize the harvested energy by balancing the time durations between the CE phase and the WPT phase, as well as allocating transmit power for the WPT phase.

To maximize the harvested energy, we consider two scenarios, where we employ dynamic-length preamble or fixed-length preamble. Given a channel estimate, we first derive the optimal energy beamformer, which applies to both scenarios. Then, we adjust the time duration for the CE phase.
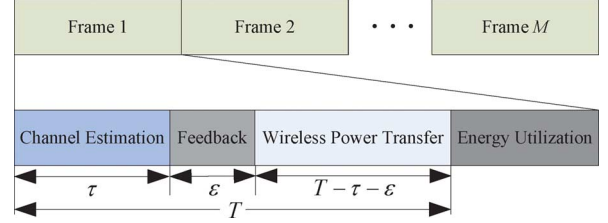

Fig. 1.   Frame Structure.

For the first scenario, the preamble length is allowed to vary dynamically depending on channel conditions. The optimal online preamble length is obtained by solving a dynamic programming (DP) problem. The DP problem is proved to reduce to an optimal stopping problem. The optimal policy is then shown to depend only on the channel estimate power. That is, if the channel estimate power is less than a time-dependent threshold, the receiver continues to perform CE, otherwise the receiver stops CE and requests wireless power. For the second scenario in which the preamble length is fixed for all frames, we optimize the preamble length offline. Moreover, we adjust the power allocated for WPT in each frame, for both scenarios. For the scenario of dynamic-length preamble, the power for WPT is allocated according to both the optimal preamble length and the channel estimate power; while for the scenario of fixed-length preamble, the power for WPT is allocated according to only the channel estimate power. Numerical results are finally given to validate our analysis.

The paper is organized as follows. In Section II, we describe the system model, and give the problem formulations. We study the optimal energy beamformer in Section III. In Section IV, we allow the preamble length to vary with frames, and use dynamic programming to find the optimal preamble length. In Section V, we fix the preamble length for all frames, and derive the optimal preamble length offline. Section VI derives the optimal power allocation schemes. Section VII gives the numerical results. Section VIII concludes this paper.

II. System Model

We consider a frame-based wireless power transfer system, consisting of a wireless power (WP) transmitter with m antennas, a single-antenna receiver that is also known as a WP receiver, a downlink channel for wireless power transfer from the WP transmitter to the WP receiver, as well as a feedback channel to send CSI (and data) from the WP receiver to the WP transmitter. Hence, the WP transmitter and WP receiver also serve as the information receiver and information transmitter, respectively. We assume that the WPT system operates with a frequency-division-duplexing protocol.

As in Fig. 1, each frame consists of four phases. To focus on efficient WPT, the time duration for the fourth EU phase is fixed and not considered in this paper. We assume the time duration for the CE, feedback and WPT phases in one frame is fixed as T symbol periods, which is normalized to be one second for convenience. In the first CE phase, the WP transmitter sends preambles, and the WP receiver performs channel estimation in an interval of τ symbol periods. In the second phase, the WP receiver feeds the CSI back to the WP transmitter within ε symbol

periods. In the third WPT phase, the WP transmitter delivers power via beamforming. The WP receiver harvests energy from the radio-frequency (RF) signals.

We assume there is a lossless link for CSI feedback[1]. For simplicity, we assume the feedback time $\varepsilon = 0$. The downlink MISO channel $\mathbf{h}$ is assumed to undergo quasi-static flat Rayleigh fading in each frame, i.e., $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{R})$, where $\mathbf{0}_m$ is the all-zero column vector, and $\mathbf{R} \triangleq \mathbb{E}[\mathbf{h}\mathbf{h}^H]$ denotes the channel covariance matrix. We assume that $\mathbf{R}$ is a full-rank matrix and has equal diagonal elements. The channel is referred as uncorrelated if $\mathbf{R} = \xi\mathbf{I}_m$, where $\xi \in (0, 1]$ is the path loss, and $\mathbf{I}_m$ is the identity matrix. The channel may vary independently from frame to frame.

A. Wireless Energy Beamforming

We assume the time duration for CE and WPT can be divided into N time slots, each of which consists of m successive symbol periods[2], i.e., $T = mN$. The preambles that consist of $k = \frac{\tau}{m}$ time slots are used to obtain the channel estimate, denoted as $\hat{\mathbf{h}}_k$. In this paper, both k and τ are discrete variables, while m is a constant integer. In the WPT phase, the received baseband signal in the n-th symbol period is written as

$$y_n = \tilde{\mathbf{x}}_n^H \mathbf{h} + z_n, \tag{1}$$

where $\tilde{\mathbf{x}}_n$ is the $m \times 1$ transmitted signal vector, and $z_n \sim \mathcal{CN}(0, \sigma_z^2)$ is the additive white Gaussian noise. For convenience of analysis, σ_z^2 is assumed to be the noise power normalized to the variance ξ of channel coefficients. The channel coefficients are accordingly considered to have unit variance in the following analysis. Given channel estimate $\hat{\mathbf{h}}_k$, we denote the $m \times 1$ beamforming vector as $\mathbf{w}(\hat{\mathbf{h}}_k)$. Then, we have $\tilde{\mathbf{x}}_n = \mathbf{w}(\hat{\mathbf{h}}_k)s_n$, where s_n is a scalar that depends on the allowable transmit power. The subscript n is ignored in the sequel.

Due to the law of energy conservation with efficiency ρ, the harvested RF-band energy[3] in one baseband symbol period, denoted by E_0, at the WP receiver is assumed to be proportional to that of the received baseband signal, i.e.,

$$E_0 = \rho\mathbb{E}_{\mathbf{h},\tilde{\mathbf{x}}}[|\tilde{\mathbf{x}}^H\mathbf{h}|^2] = \rho\mathbb{E}_{\mathbf{h},\hat{\mathbf{h}}_k}[|\mathbf{w}^H(\hat{\mathbf{h}}_k)\mathbf{h}|^2]. \tag{2}$$

We assumed in (2) that the energy due to the ambient noise cannot be harvested. For convenience, we also assume $\rho = 1$ in this paper.

B. Problem Formulation

The WP receiver aims to harvest energy as much as possible in the WPT phase. Intuitively, longer preambles can increase the accuracy of channel estimation, and thus increase the efficiency of power transfer, but at the cost of reduced time left for the WPT phase. We also note that the power of the received signal depends on the fading condition in one frame. Hence, to

maximize the harvested energy, we first consider two scenarios with constant preamble power, where we optimize the preamble length dynamically for each frame or optimize a fixed preamble length offline. Then we optimize the transmit power for WPT in each frame, via the power allocation.

1) WPT With Dynamic-Length Preamble: We consider the scenario where the preamble length is allowed to vary dynamically, i.e., the receiver can decide to perform CE or request WP at any time slot based on its current channel estimate. We denote the beginning of the $(k + 1)$-th time slot as time instant k, where $k = 0, 1, \ldots, N - 1$. At time instant $k = 0$, the receiver decides to perform CE or request WP in the first slot. If it decides to request WP at $k = 0$, the transmitter performs WPT in the first slot without beamforming. Otherwise, the transmitter sends preambles in the first slot, and the receiver obtains the channel estimate $\hat{\mathbf{h}}_1$ at the end of the first slot. For the subsequent time instant k, if the receiver decides to request WP, it feeds back the channel estimate $\hat{\mathbf{h}}_k$ to the transmitter. Then the transmitter performs WPT using optimal beamformer $\mathbf{w}^\star(\hat{\mathbf{h}}_k)$ in the next slot. If the receiver decides to continue CE at instant k, the transmitter sends preambles in the $(k + 1)$-th time slot. The optimal beamformer $\mathbf{w}^\star(\hat{\mathbf{h}}_k)$ will be found in Section III.

In Section IV, we first formulate a dynamic programming (DP) problem to maximize the harvested energy assuming that constant transmit power is used for WPT. We define therein the control space $\mathcal{C}$, decision variable u_k and the system state $\mathbf{x}_k$. We define a policy as a sequence of functions $\mu_k(\mathbf{x}_k)$ which maps each system state into a decision at time instant $k = 0, 1, \ldots, N - 1$. The set of all possible policies is denoted as Π. Let $g_k(\mathbf{x}_k, u_k)$ be the energy harvested in slot k with state $\mathbf{x}_k$ and decision u_k. To maximize the expected harvested energy in all slots, we thus have the following optimization problem

$$(\text{P}_1) \max_{\pi \in \Pi} \mathbb{E}\left[\sum_{k=0}^{N-1} g_k(\mathbf{x}_k, u_k)\right]. \tag{3}$$

The expectation is performed over all random variables, specifically the channel $\mathbf{h}$ and the channel estimates $\{\hat{\mathbf{h}}_k\}$ which become available only after the decision of CE or WP is made. The optimal policy $\pi^\star$ for (P_1) is obtained in Section IV-C.

In particular, we assume an explicit feedback protocol in which the receiver feeds back the decision on whether to continue sending preambles or to stop sending preambles. Although the explicit feedback protocol is difficult to implement in practice, it allows us to investigate the best possible performance under a general setting. We shall see later in Section IV that it is in fact optimal to feed back once to stop preamble transmission, resulting into an efficient implicit feedback protocol, see Remark 1.

Then, we derive the optimal power allocation in Section VI-A1, which allocates transmit power for WPT according to both optimal preamble length adapted by using the optimal policy $\pi^\star$, and the channel estimate power.

2) WPT With Fixed-Length Preamble: To reduce implementation complexity, we consider the scenario in which the preamble length is fixed as k time slots in all frames, but can be optimized offline. Then, the WPT phase in each frame consists of $(T - \tau)$ symbol periods, where $\tau = km$.

---

[1] In practice, the cyclic-redundancy-check scheme with retransmission can be used to ensure an error-free feedback.

[2] For simplicity, we assume the number of successive symbol periods is m, although it is not necessary in practice.

[3] Note that (2) is the harvested energy for the scenario in which the path loss is normalized to one and the time duration of T symbol periods is normalized to be one second.

In Section V, we first maximize the harvested energy assuming that constant transmit power is used for WPT, by optimizing both the preamble length k and the beamforming vector $\mathbf{w}(\hat{\mathbf{h}}_k)$. Specifically, we have the following optimization problem

$$(\mathrm{P}_2) \quad \max_{\substack{\{\mathbf{w}(\hat{\mathbf{h}}_k),\ \forall \hat{\mathbf{h}}_k\},\\ 0 \le k \le N-1}} \quad (T - km)\mathbb{E}_{\mathbf{h},\hat{\mathbf{h}}_k}[|\mathbf{w}^H(\hat{\mathbf{h}}_k)\mathbf{h}|^2] \quad (4a)$$

$$\text{s. t.} \quad \|\mathbf{w}(\hat{\mathbf{h}}_k)\| = 1, \quad (4b)$$

where $\|\cdot\|$ is the l_2-norm. We will find the optimal solution $\mathbf{w}^\star$ (as a function of $\hat{\mathbf{h}}_k$) and $k^\star$ in Sections III, and V, respectively. Then, we derive the optimal power allocation scheme in Section VI-B, which allocates transmit power for WPT according to only the channel estimate power.

III. OPTIMAL ENERGY BEAMFORMING

In this section, we obtain the optimal beamforming vector $\mathbf{w}^\star(\hat{\mathbf{h}})$, which shall be used to find the solutions to problem $(\mathrm{P}_1), (\mathrm{P}_2)$ in Sections IV, V, respectively.

A. Partial or Full Feedback

In practice, it is difficult for the transmitter to obtain full CSI due to the limited feedback capacity. This motivates us to investigate the impact of different amount of feedback on energy beamforming and thus the harvested energy. We let the receiver selectively feed back only q $(1 \le q \le m)$ largest channel coefficients to the transmitter, so as to reduce the feedback amount. If $q = m$, it reduces to the conventional full CSI feedback. Let $\hat{h}_{i_r}$ denote the channel coefficient with the r-th largest channel gain. The receiver quantizes the vector $\hat{\mathbf{h}}_q \triangleq [\hat{h}_{i_1}, \hat{h}_{i_2}, \ldots, \hat{h}_{i_q}]^T$ and the corresponding index set $\mathcal{I} \subseteq \{1, 2, \ldots, q\}$ to denote the selected antennas, and feeds back the coded digital bits to the transmitter. An additional $\log(C_q^m)$ bits are required for the feedback of the index set $\mathcal{I}$. The parameter q is defined as a metric, namely the *feedback dimension*, to quantify the cost/amount of feedback. The transmitter uses only q antennas with index in $\mathcal{I}$ to perform energy beamforming.

B. Optimal Energy Beamforming

The energy beamforming is performed by using imperfect CSI at the transmitter. We first obtain the distribution of the channel $\mathbf{h}$ conditioned on a general *unbiased* channel estimate $\hat{\mathbf{h}}$. We consider the q-dimensional feedback of CSI. Define the estimation error $\mathbf{e}_q \triangleq \hat{\mathbf{h}}_q - \mathbf{h}_q$. Let $\mathbf{R}_q \triangleq \mathbb{E}[\hat{\mathbf{h}}_q \hat{\mathbf{h}}_q^H]$ and $\mathbf{R}_{\mathrm{e},q} \triangleq \mathbb{E}[\mathbf{e}_q \mathbf{e}_q^H]$ be the q-dimensional counterparts of channel covariance matrix $\mathbf{R}$ and the error covariance matrix $\mathbf{R}_\mathrm{e}$, respectively. From (16) in [19], we have the following lemma.

Lemma 1: Let $\hat{\mathbf{h}}_q = \mathbf{h}_q + \mathbf{e}_q$. Assume $\mathbf{h}_q \sim \mathcal{CN}(\mathbf{0}_q, \mathbf{R}_q)$, the error vector $\mathbf{e}_q \sim \mathcal{CN}(\mathbf{0}_q, \mathbf{R}_{\mathrm{e},q})$, and $\mathbf{h}_q$ and $\mathbf{e}_q$ are jointly Gaussian distributed. Given $\hat{\mathbf{h}}_q$, the vector $\mathbf{h}_q$ follows a complex Gaussian distribution, i.e.,

$$\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q \sim \mathcal{CN}\left(\mathbf{m}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}, \boldsymbol{\Sigma}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}\right) \quad (5)$$

where $\mathbf{m}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q} = (\mathbf{R}_{\mathrm{e},q}\mathbf{R}_q^{-1} + \mathbf{I}_q)^{-1}\hat{\mathbf{h}}_q$, and $\boldsymbol{\Sigma}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q} = (\mathbf{R}_q^{-1} + \mathbf{R}_{\mathrm{e},q}^{-1})^{-1}$.

From Lemma 1, the conditional correlation matrix is

$$\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q} = \boldsymbol{\Sigma}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q} + \mathbf{m}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}\mathbf{m}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}^H. \quad (6)$$

Denote the singular value decomposition (SVD) of $\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}$ by $\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q} = \mathbf{U}_q \boldsymbol{\Gamma}_q \mathbf{U}_q^H$, where $\mathbf{U}_q = [\mathbf{u}_{1,q}\ \mathbf{u}_{2,q}\ \cdots\ \mathbf{u}_{q,q}], \boldsymbol{\Gamma}_q = \mathrm{diag}\{\gamma_1\ \gamma_2\ \cdots\ \gamma_q\}$, and $\gamma_1 \ge \gamma_2 \ge \cdots \ge \gamma_q$. We further have Lemma 2.

Lemma 2: Assume $\mathbf{h}_q \sim \mathcal{CN}(\mathbf{0}_q, \mathbf{R}_q), \mathbf{e}_q \sim \mathcal{CN}(\mathbf{0}_q, \mathbf{R}_{\mathrm{e},q})$, and $\mathbf{h}_q$ and $\mathbf{e}_q$ are jointly Gaussian. Given $\hat{\mathbf{h}}_q$, the optimal beamforming vector that maximizes the normalized harvested energy, is given by

$$\mathbf{w}_q^\star(\hat{\mathbf{h}}_q) = \mathbf{u}_{1,q}. \quad (7)$$

Proof: The harvested energy in one symbol period in (2) can be rewritten as

$$E_0 = \mathbb{E}_{\hat{\mathbf{h}}_q}\left[\mathbb{E}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}\left[\mathbf{w}_q^H(\hat{\mathbf{h}}_q)\mathbf{h}_q\mathbf{h}_q^H\mathbf{w}_q(\hat{\mathbf{h}}_q)\right]\right]$$

$$\overset{(a)}{=} \mathbb{E}_{\hat{\mathbf{h}}_q}\left[\mathbf{w}_q^H(\hat{\mathbf{h}}_q)\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}\mathbf{w}_q(\hat{\mathbf{h}}_q)\right], \quad (8)$$

where (a) is from the fact that conditioned on the channel estimate $\hat{\mathbf{h}}_q$, the beamformer $\mathbf{w} = \mathbf{w}(\hat{\mathbf{h}}_q)$ is fixed and treated as a constant. Clearly, the E_0 is maximized when the beamformer is the largest eigenmode of $\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}$. ∎

Using the optimal beamformer in (7), the normalized harvested energy E_0 is the mean of the largest eigenvalue γ_1 of the matrix $\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}$ in (6). The total harvested energy in all remaining slots is thus given by

$$E = (T - \tau)\,\mathbb{E}_{\hat{\mathbf{h}}_q}\left[\gamma_1\left(\mathbf{R}_{\mathbf{h}_q \,|\, \hat{\mathbf{h}}_q}\right)\right]. \quad (9)$$

We have assumed the channel estimate is unbiased, which can be obtained by the widely-used least-square (LS) channel estimator [20][4]. We keep the transmit power for preambles a constant, which implies the effective receive signal-to-noise ratio (SNR) for CE is only proportional to the preamble length τ that is to be optimized. We next derive the optimal beamformer.

C. Optimal Beamformer for LS Channel Estimation

We first describe the optimal design of preambles. We set the length of preambles as $\tau = km$, where $k = 0, \ldots, N-1$. When the total power for sending preambles is fixed as k, it is shown in [22] that the LS estimation performance can be optimized by using the preamble matrix as

$$\mathbf{X} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \ldots \quad \mathbf{X}_k]^T, \quad (10)$$

where $\mathbf{X}_i = \frac{1}{\sqrt{m}}\mathbf{I}_m$, for $i = 1, 2, \ldots, k$. From [20], we obtain the LS estimate as follows

$$\hat{\mathbf{h}} = \mathbf{X}^{-1}\mathbf{y} = \mathbf{h} + \frac{\sqrt{m}}{k}\sum_{i=1}^{k}\mathbf{z}_i, \quad (11)$$

---

[4]It turns out that the optimal beamformer in Lemma 2 also applies to other channel estimators such as an linear minimum mean-square-error estimator (LMMSE, see Section III-D in [21])

where the length-m noise vector $\mathbf{z}_i \sim \mathcal{CN}(\mathbf{0}_m, \sigma_z^2 \mathbf{I}_m)$. Clearly, the estimation error $\mathbf{e}$ is distributed as $\mathcal{CN}(\mathbf{0}_m, \sigma_e^2 \mathbf{I}_m)$, where $\sigma_e^2 = \frac{\sigma_z^2}{\beta}$ and $\beta = \frac{k}{m} = \frac{\tau}{m^2}$.

1) Correlated Channel: From Lemma 1, we state that given $\hat{\mathbf{h}}_q$, the channel vector $\mathbf{h}_q$ is distributed as $\mathcal{CN}((\sigma_e^2 \mathbf{R}_q^{-1} + \mathbf{I}_q)^{-1}\hat{\mathbf{h}}_q, (\mathbf{R}_q^{-1} + \frac{1}{\sigma_e^2}\mathbf{I}_q)^{-1})$. The conditional correlation matrix $\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q}$ yields

$$\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q} = \left(\mathbf{R}_q^{-1} + \frac{1}{\sigma_e^2}\mathbf{I}_q \right)^{-1}$$
$$+ \left(\sigma_e^2 \mathbf{R}_q^{-1} + \mathbf{I}_q \right)^{-1} \hat{\mathbf{h}}_q \hat{\mathbf{h}}_q^H \left(\sigma_e^2 \mathbf{R}_q^{-1} + \mathbf{I}_q \right)^{-1}. \quad (12)$$

From (7), the optimal beamforming vector is the largest eigenmode of $\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q}$ in (12).

2) Uncorrelated Channel: From Lemma 1, given $\hat{\mathbf{h}}_q$, the channel is distributed as $\mathbf{h}_q \sim \mathcal{CN}(\frac{\hat{\mathbf{h}}_q}{1+\sigma_e^2}, \frac{\sigma_e^2}{1+\sigma_e^2}\mathbf{I}_q)$. The conditional correlation matrix is thus rewritten as

$$\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q} = \frac{\sigma_e^2}{1+\sigma_e^2}\mathbf{I}_q + \frac{\hat{\mathbf{h}}_q \hat{\mathbf{h}}_q^H}{(1+\sigma_e^2)^2}. \quad (13)$$

Note that $\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q}$ is the sum of a scaled identity matrix and a rank-one matrix. The eigenvectors can be constructed as follows: take the normalized $\hat{\mathbf{h}}_q$ as the right eigenvector corresponding to the maximal eigenvalue, and construct other mutually orthogonal eigenvectors by Gram-Schmidt algorithm. From Lemma 2, the optimal beamformer is

$$\mathbf{w}_q^\star(\hat{\mathbf{h}}_q) = \frac{\hat{\mathbf{h}}_q}{\|\hat{\mathbf{h}}_q\|}. \quad (14)$$

Associated to the optimal beamformer in (14), the largest eigenvalue of the matrix $\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q}$ is

$$\gamma_1 \left(\mathbf{R}_{\mathbf{h}_q \mid \hat{\mathbf{h}}_q} \right) = \frac{\sigma_e^2}{1+\sigma_e^2} + \frac{\|\hat{\mathbf{h}}_q\|^2}{(1+\sigma_e^2)^2}. \quad (15)$$

It is noted that the largest eigenvalue in (15) gives the expected harvested energy in one symbol period with the channel estimate $\hat{\mathbf{h}}_q$.

IV. WPT WITH DYNAMIC-LENGTH PREAMBLE

In this section, we consider the scenario where the preamble length is allowed to vary dynamically depending on the current channel estimate. To maximize the expected harvested energy, we first formulate a dynamic programming (DP) problem [23], which will be shown to reduce to an optimal stopping problem, and thus can be simplified considerably. Using the optimal DP policy, we shall derive the optimal power allocation scheme in Section VI-A.

We assume uncorrelated flat Rayleigh fading channels that are static in each frame but vary independently among frames. The extension of the model to more general case of Markovian channels is more tedious but conceptually straightforward, see, e.g., [24]. Let $\hat{\mathbf{h}}_k$ denote the channel estimate available at time instant k (i.e., the beginning of the $(k+1)$-th time slot). Assuming no priori channel knowledge is available, we initialize the channel estimate as the mean of $\mathbf{h}$, i.e., $\hat{\mathbf{h}}_0 = \mathbf{0}$. We assume that the receiver adopts an LS channel estimator and performs

full (i.e., m-dimensional) feedback[5]. The optimal beamformer in (14) for $q = m$ is thus used in this section. We employ the preamble matrix in (10). For $k = 0, 1, \ldots, N-2$, it is useful to rewrite the LS channel estimate in (11) as the following recursive equation

$$\hat{\mathbf{h}}_{k+1} = \frac{k}{k+1}\hat{\mathbf{h}}_k + \frac{\mathbf{h}}{k+1} + \frac{\sqrt{m}\mathbf{z}_{k+1}}{k+1}. \quad (16)$$

A. Statistical Properties of Channel Estimates

Before formulating the problem and obtaining the solutions, we first obtain some useful statistical properties. Lemma 3 quantifies the statistical relationship of two adjacent channel estimates, while Lemma 4 shows that the most recent channel estimate provides sufficient statistics for estimating the channel.

Lemma 3: Given $\hat{\mathbf{h}}_k$, the next channel estimate $\hat{\mathbf{h}}_{k+1}$ is distributed as $\mathcal{CN}(\bar{\mathbf{u}}_k, \bar{\sigma}_k^2 \mathbf{I}_m)$, where

$$\bar{\mathbf{u}}_k = \frac{k(k+1+m\sigma_z^2)}{(k+1)(k+m\sigma_z^2)}\hat{\mathbf{h}}_k, \quad \bar{\sigma}_k^2 = \frac{m\sigma_z^2(k+1+m\sigma_z^2)}{(k+1)^2(k+m\sigma_z^2)}.$$

Proof: Let $\sigma_k^2 = \frac{m}{k}\sigma_z^2$. From Lemma 1 for $q = m$, we have that $\mathbf{h} \mid \hat{\mathbf{h}}_k \sim \mathcal{CN}(\frac{1}{1+\sigma_k^2}\hat{\mathbf{h}}_k, \frac{\sigma_k^2}{1+\sigma_k^2}\mathbf{I}_m)$. From (16) and the independence between $\hat{\mathbf{h}}_k$ and the noise vector $\mathbf{z}_{k+1}$, we obtain the result after algebraic manipulations. ∎

We take the channel estimate power as a random variable V_k, i.e., $V_k \triangleq \|\hat{\mathbf{h}}_k\|^2$. From Lemma 3, conditioned on $V_k = v_k$, the random variable $\frac{2}{\bar{\sigma}_k^2}V_{k+1}$ follows the noncentral Chi-Square distribution with the degree $2m$ of freedom and the noncentrality parameter

$$\theta_k = \frac{2k^2(k+1+m\sigma_z^2)v_k}{m\sigma_z^2(k+m\sigma_z^2)}.$$

Moreover, the conditional probability density function (pdf) of V_{k+1} is thus given by [25]

$$f(v_{k+1} \mid v_k) = \frac{1}{\bar{\sigma}_k^2} \exp\left(-\frac{v_{k+1}}{\bar{\sigma}_k^2} - \frac{\theta_k}{2} \right)$$
$$\cdot \left(\frac{2v_{k+1}}{\theta_k \bar{\sigma}_k^2} \right)^{\frac{m-1}{2}} I_{m-1}\left(\frac{2\theta_k v_{k+1}}{\bar{\sigma}_k^2} \right), \quad (17)$$

where $I_{m-1}(\cdot)$ is the $(m-1)$-th order modified Bessel function of the first kind. The conditional mean is

$$\mathbb{E}_{V_{k+1} \mid V_k = v_k}[V_{k+1}] = \bar{\sigma}_k^2 \left(m + \frac{\theta_k}{2} \right). \quad (18)$$

Lemma 4: Given a sequence of LS channel estimates $\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_k$, the distribution of channel vector $\mathbf{h}$ conditioned on all channel estimates is simplified as

$$f(\mathbf{h} \mid \hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_k) = f(\mathbf{h} \mid \hat{\mathbf{h}}_k), \quad (19)$$

which is the Gaussian distribution $\mathcal{CN}(\frac{1}{1+\sigma_k^2}\hat{\mathbf{h}}_k, \frac{\sigma_k^2}{1+\sigma_k^2}\mathbf{I}_m)$, where $\sigma_k^2 = \frac{m}{k}\sigma_z^2$.

Proof: See proof in Appendix A. ∎

---

[5]With full feedback, the amount of feedback is still acceptable. This is because in the optimal DP scheme, the receiver needs to feed back only once in each frame, which will be shown in Theorem 1.

Lemma 4 suggests that the accuracy of channel estimation can not be increased by using all available channel estimates, compared to using only the most recent channel estimate. This observation will be used to show the structure of the optimal DP policy (see Theorem 1, later).

B. Problem Formulation

We formulate the optimization problem to maximize the total expected harvested energy, assuming that the transmitter uses constant transmit power for WPT. We first make the necessary definitions.

1) Decision (or Control) Variable: We denote the decision variable as $u_k \in \mathcal{C}$. The decision space $\mathcal{C}$ consists of only two elements s and c, that corresponds to stopping CE (i.e., requesting WP) or continuing CE, respectively. We initialize $u_{-1} = \textsf{c}$.

2) System State: We define the system state $\mathbf{x}_k$ as consisting of (i) δ_k which denotes the number of slots used so far for CE, and (ii) the most recently available channel estimate. Given u_k and current state $\mathbf{x}_k$, the next state is

$$\mathbf{x}_{k+1} = \begin{cases} \{\delta_k + 1, \hat{\mathbf{h}}_{k+1}\}, & \text{if } u_k = \textsf{c} \\ \mathbf{x}_k, & \text{if } u_k = \textsf{s}. \end{cases} \quad (20)$$

The initial state is $\mathbf{x}_0 = \{\delta_0, \hat{\mathbf{h}}_0\}$ with $\delta_0 = 0, \hat{\mathbf{h}}_0 = \emptyset$. We denote the space of all possible state as $\mathcal{S}$. From Lemma 4, this system state is sufficient to obtain the statistics of $\mathbf{h}$ even if all priori channel estimates were made available.

3) Policy: Define a policy π as a sequence of functions

$$\pi = \{\mu_k(\mathbf{x}_k), \quad \forall \mathbf{x}_k \in \mathcal{S}, k = 0, 1, \ldots, N-1\},$$

where $\mu_k : \mathcal{S} \to \mathcal{C}$ is a function that maps the state $\mathbf{x}_k$ into the decision variable in the next time slot, i.e., $u_k = \mu_k(\mathbf{x}_k)$. We denote the set of all possible policies as Π.

4) Reward: Given state $\mathbf{x}_k$ and decision $u_k = \mu_k(\mathbf{x}_k)$, we denote $g_k(\mathbf{x}_k, u_k)$ as the reward, given by the expected harvested energy in slot k. If $u_k = \textsf{s}$, we have from (15) with $\sigma_e^2 = \frac{m}{\delta_k}\sigma_z^2$ that

$$g_k(\mathbf{x}_k, u_k) = m \left(\frac{m\sigma_z^2}{\delta_k + m\sigma_z^2} + \frac{\delta_k^2 \|\hat{\mathbf{h}}_k\|^2}{(\delta_k + m\sigma_z^2)^2} \right), \quad (21)$$

and if $u_k = \textsf{c}$, then $g_k(\mathbf{x}_k, u_k) = 0$.

5) Dynamic Program and Optimal Policy: To maximize the total harvested energy, we thus have the optimization problem (P$_1$) given in (3). The optimal policy $\pi^\star$ is given by the functions $\{\mu_k(\cdot)\}$, i.e., the decision u_k given state $\mathbf{x}_k$, that satisfy the Bellman's equation [23]:

$$J_{N-1}(\mathbf{x}_{N-1}) = \max_{u_{N-1}} g_{N-1}(\mathbf{x}_{N-1}, u_{N-1}),$$
$$J_k(\mathbf{x}_k) = \max_{u_k} g_k(\mathbf{x}_k, u_k) + \mathbb{E}_{\hat{\mathbf{h}}_{k+1} | \hat{\mathbf{h}}_k}[J_{k+1}(\mathbf{x}_{k+1})],$$
$$(22)$$

for $k = N-2, \ldots, 0$. Here, $J_k(\mathbf{x}_k)$ is known as the value function which represents the harvested energy for the last $(N-k)$ time slots, conditioned on the current system state $\mathbf{x}_k$. Typically, the solution is obtained by backward recursion, by first solving for $\mu_{N-1}(\cdot)$ for slot $N-1$, then for $\mu_{N-2}(\cdot), \ldots, \mu_0(\cdot)$. The

maximum expected harvested energy with policy $\pi^\star$ is given by $J_1(\mathbf{x}_1)$.

C. Optimal Policy

Theorem 1 states that the Bellman's equation (22) can be reduced to an optimal stopping problem, for which a decision is changed at most from c to s once and fixed henceforth.

Theorem 1: Any decision sequence of the optimal policy $\pi^\star$ has the structure

$$(u_0^*, u_1^*, \ldots, u_{N-1}^*) = (\underbrace{\textsf{c}, \textsf{c}, \ldots, \textsf{c}}_{k=0,1,\ldots,k^*-1} , \underbrace{\textsf{s}, \textsf{s}, \ldots, \textsf{s}}_{k=k^*,\ldots,N-1}), \quad (23)$$

where $0 \leq k^* \leq N-1$. That is, the optimal policy initially performs only CE for the first k^* slots, then performs only WPT for the remaining slots.

Proof: See proof in Appendix B. ∎

Remark 1 (Optimality of an Implicit Feedback Protocol): Theorem 1 implies that the WP receiver only needs to feed back once, after it decides to request WP. Hence, an implicit feedback protocol in which the WP receiver feeds back only to stop preamble transmissions, is sufficient and can be used in practical implementation[6]. If there is no feedback received, the WP transmitter will assume that the decision at the WP receiver is to continue CE, and thus send preambles continuously.

Theorem 1 allows us to simplify the DP problem and obtain a solution that can be implemented with low complexity. Before we obtain the structure of the optimal policy in Theorem 2, we first state the expected harvested energy under different scenarios. Henceforth, we assume the optimal policy is employed.

Given state $\mathbf{x}_k$, if the receiver decides to request WP, i.e., $u_k = \textsf{s}$, then the expected harvested energy in the remaining slots is obtained from (9) and (15) as

$$\tilde{E}(\hat{\mathbf{h}}_k, k) = m(N-k)\mathbb{E}_{\mathbf{h}|\hat{\mathbf{h}}}\left[\mathbf{w}_{k,\text{opt}}^H \mathbf{h}\mathbf{h}^H \mathbf{w}_{k,\text{opt}}\right]$$
$$= A_k(B_k + C_k\|\hat{\mathbf{h}}_k\|^2), \quad (24)$$

where $A_k = m(N-k), B_k = \frac{m\sigma_z^2}{k+m\sigma_z^2}$, and $C_k = \frac{k^2}{(k+m\sigma_z^2)^2}$. If the decision is instead $u_k = \textsf{c}$, then the expected harvested energy in the last $(N-k)$ time slots, under all possible decisions made for subsequent slots, is

$$\bar{J}_{k+1}(\mathbf{x}_k) = \mathbb{E}_{\hat{\mathbf{h}}_{k+1}|\hat{\mathbf{h}}_k}[J_{k+1}(\mathbf{x}_{k+1})]. \quad (25)$$

For the special case in which the receiver decides to continue CE ($u_k = \textsf{c}$) at time instant k and stop CE ($u_{k+1} = \textsf{s}$) at time instant $k+1$, then from the conditional mean in (18), the expected harvested energy conditioned on $\hat{\mathbf{h}}_k$ is obtained after some algebraic manipulation as

$$\bar{E}(\hat{\mathbf{h}}_{k+1}, k+1) = \mathbb{E}_{\hat{\mathbf{h}}_{k+1}|\hat{\mathbf{h}}_k}[\tilde{E}(\hat{\mathbf{h}}_{k+1}, k+1)]$$
$$= D_{k+1}\frac{k^2(k+1+m\sigma_z^2)\|\hat{\mathbf{h}}_k\|^2 + F_{k+1}}{G_{k+1}},$$
$$(26)$$

---

[6]The WP transmitter will send slightly more preambles, due to the feedback delay induced by the channel and processing at the receiver and transmitter. However, the resulting performance degradation is small if the delay is small, compared to the frame length.

where $D_{k+1} = m(N - k - 1)$, $F_{k+1} = m\sigma_z^2(k + m\sigma_z^2)(k + m\sigma_z^2 + m)$, and $G_{k+1} = (k + 1 + m\sigma_z^2)(k + m\sigma_z^2)^2$.

Now, we state the optimal policy in Theorem 2.

Theorem 2: The optimal policy to Problem (P_1) depends only on the channel estimate power, i.e.,

$$u_k = \begin{cases} \mathsf{c}, & \text{if} \quad \|\hat{\mathbf{h}}_k\|^2 \in \mathcal{D}_{\mathsf{c},k} \\ \mathsf{s}, & \text{if} \quad \|\hat{\mathbf{h}}_k\|^2 \in \mathcal{D}_{\mathsf{s},k} \end{cases} \tag{27}$$

where the sets (intervals)

$$\mathcal{D}_{\mathsf{c},k} = [[0, \lambda_{k,1}) \cup [\lambda_{k,2}, \lambda_{k,3}) \cup \cdots \cup [\lambda_{k,M_k-1}, \lambda_{k,M_k})],$$
$$\mathcal{D}_{\mathsf{s},k} = [[\lambda_{k,1}, \lambda_{k,2}) \cup [\lambda_{k,3}, \lambda_{k,4}) \cup \cdots \cup [\lambda_{k,M_k}, +\infty)],$$

and $\lambda_{k,1} \leq \cdots \leq \lambda_{k,M_k}$ are the solutions to $\tilde{E}(\hat{\mathbf{h}}_k, k) = \bar{J}_{k+1}(\mathbf{x}_{k+1})$ with respect to the variable $\|\hat{\mathbf{h}}_k\|^2$.

Proof: From Theorem 1, to obtain the optimal policy for the original DP problem, if $u_{k-1} = \mathsf{s}$, then $u_k = \mathsf{s}$; if $u_{k-1} = \mathsf{c}$, it suffices to compare if $u_k = \mathsf{s}$ or c results in a larger output of the value function. Hence, from (21), the value function simplifies as

$$J_k(\mathbf{x}_k) = \max\{\tilde{E}(\hat{\mathbf{h}}_k, k), \bar{J}_{k+1}(\mathbf{x}_k)\} \tag{28}$$

for $k = N-2, N-1, \ldots, 0$, in a backward manner. Specifically, assuming $u_{k-1} = \mathsf{c}$, if $\tilde{E}(\hat{\mathbf{h}}_k, k) > \bar{J}_{k+1}(\mathbf{x}_k)$, then the receiver requests WP with $u_k = \mathsf{s}$, otherwise the receiver continues CE with $u_k = \mathsf{c}$. Hence, to obtain the optimal decision, we have to solve the equation

$$\tilde{E}(\hat{\mathbf{h}}_k, k) = \bar{J}_{k+1}(\mathbf{x}_k), \tag{29}$$

to compare the two terms in (28). From (24), the first term $\tilde{E}(\hat{\mathbf{h}}_k, k)$ is a monotonically increasing linear function of only the channel estimate power $\|\hat{\mathbf{h}}_k\|^2$. Moreover, the second term $\bar{J}_{k+1}(\mathbf{x}_k)$, given in (25), is also a function of only $\|\hat{\mathbf{h}}_k\|^2$. This claim is checked by induction based on backward recursion as follows. At time instant $k = N - 2$, we have $\bar{J}_{N-1}(\mathbf{x}_{N-2}) = \bar{E}(\hat{\mathbf{h}}_{N-1}, N - 1)$. From (26), the term $\bar{J}_{N-1}(\mathbf{x}_{N-2})$ is a linear function of $\|\hat{\mathbf{h}}_{N-2}\|^2$. It follows from (28) that $J_{N-2}(\mathbf{x}_{N-2})$ is a piecewise linear function of $\|\hat{\mathbf{h}}_{N-2}\|^2$. From (25) and the conditional distribution in Lemma 3, we thus obtain that $\bar{J}_{N-2}(\mathbf{x}_{N-3})$ is a function of $\|\hat{\mathbf{h}}_{N-3}\|^2$. By mathematical induction with decreasing slot index, we have that $\bar{J}_{k+1}(\mathbf{x}_k)$ is a function of $\|\hat{\mathbf{h}}_k\|^2$.

Hence, the decision policy depends only on $\|\hat{\mathbf{h}}_k\|^2$. Denote the solution(s) to (29) with respect to $\|\hat{\mathbf{h}}_k\|^2$ by $\lambda_{k,1}, \lambda_{k,2}, \ldots, \lambda_{k,M_k}$, assuming $\lambda_{k,1} \leq \cdots \leq \lambda_{k,M_k}$. The desired result is obtained. ∎

In general, the state value is of $(m + 1)$-dimension, and the complexity of obtaining the optimal policy and implementing it can be very high. Moreover, the memory required to store the policy too large. From Theorem 2, however, the optimal policy can be implemented for each slot by only comparing the channel estimate power to a scalar value, thus saving complexity in computation and storage of policy. The thresholds can be pre-computed and stored in a lookup table. During online implementation, the receiver refers to the table to make the decision.

D. Optimal Thresholds

In this section, we derive the optimal thresholds $\{\lambda_{k,j}\}$ in Theorem 2 in a backward manner, by solving (29) for $k = N - 1, N - 2, \ldots, 0$, assuming $u_{k-1} = \mathsf{c}$. At time instant $N - 1$, we have $\tilde{E}(\hat{\mathbf{h}}_{N-1}, N - 1) > 0$, $\forall \hat{\mathbf{h}}_{N-1} \in \mathbb{C}^m$. Thus, it is optimal to set $\lambda_{N-1} = 0$. This is because it is always optimal for the receiver to stop CE at time instant $N - 1$, since $\|\hat{\mathbf{h}}_{N-1}\|^2 > 0$ holds with probability one for fading channels.

For $k = N - 2$, we have $\bar{J}_{N-1}(\mathbf{x}_{N-2}) = \bar{E}(\hat{\mathbf{h}}_{N-1}, N - 1)$. The (29) for $k = N - 2$ thus reduces to $\tilde{E}(\hat{\mathbf{h}}_{N-2}, N - 2) = \bar{E}(\hat{\mathbf{h}}_{N-1}, N - 1)$. Observe that both the left-hand side (LHS) and right-hand side (RHS) are both monotonically increasing linear functions of $\|\hat{\mathbf{h}}_{N-2}\|^2$. Hence, the decision policy at this time instant depends on a single threshold. If $D_{N-1}(N - 2)^2 \neq A_{N-2}G_{N-1}C_{N-2}$, the threshold is

$$\lambda_{N-2} = \left[\frac{A_{N-2}B_{N-2}G_{N-1} - D_{N-1}F_{N-1}}{D_{N-1}(N - 2)^2 - A_{N-2}G_{N-1}C_{N-2}} \right]^+, \tag{30}$$

where the notation $[x]^+ = \max\{0, x\}$; and $\lambda_{N-2} = 0$, otherwise. For the subsequent slots $k = N - 3, N - 4, \ldots, 0$, we can obtain the thresholds by a numerical search as follows. By substituting (28) into (25), we note that the RHS of (29) is expressed in a recursive form

$$\bar{J}_{k+1}(\mathbf{x}_k)$$
$$= \mathbb{E}_{\hat{\mathbf{h}}_{k+1} \mid \hat{\mathbf{h}}_k}[\max\{\tilde{E}(\hat{\mathbf{h}}_{k+1}, k + 1), \bar{J}_{k+2}(\mathbf{x}_{k+1})\}]. \tag{31}$$

Hence, it is difficult to obtain a closed-form solution of the threshold $\|\hat{\mathbf{h}}_k\|^2$ that solves (29). Thus, we let the quantity $\|\hat{\mathbf{h}}_k\|^2$ take discrete values in the set $\mathcal{Q} \triangleq \{\Delta, 2\Delta, \ldots, M\Delta\}$, and search for the closest values(s) in the set $\mathcal{Q}$ that solves (29).

In general, there may be multiple solutions to (29), denoted as $\lambda_{k,j}$, $j = 1, \ldots, M_k$, since the LHS $\tilde{E}(\hat{\mathbf{h}}_k, k)$ is a monotonically increasing linear function of $\|\hat{\mathbf{h}}_k\|^2$, and the RHS is a function of $\|\hat{\mathbf{h}}_k\|^2$, which may not be a monotonic function. To get more insights, we give a numerical example on the thresholds.

Example 1: Let $N = 42, m = 3$ and the noise power $\sigma_z^2 = -63$ dBm, see the detailed parameter setting in Section VII. The thresholds are numerically computed and shown in Fig. 2. We numerically find that the threshold at each time index k is always unique, which can further simplify the decision process in practice. It is observed that the threshold monotonically decreases as the time index k increases. This observation is consistent with the intuition that if a channel estimate is good enough to be acceptable at time k for WP to be performed, it should also be acceptable at time $(k + 1)$ when there will be one more slot available for the channel estimate to be improved.

V. WPT WITH FIXED-LENGTH PREAMBLE

In this section, we consider the scenario wherein the preamble length is fixed in all frames. We first derive the optimal preamble length. This corresponds to the case of offline adaptation, in contrast to online adaptation in Section IV where the preamble length is varied over frames. Based on the optimized fixed-length preamble, we shall derive the optimal power allocation scheme in Section VI-B. We assume uncorrelated channel. We aslo assume that the receiver adopts an LS channel estimator
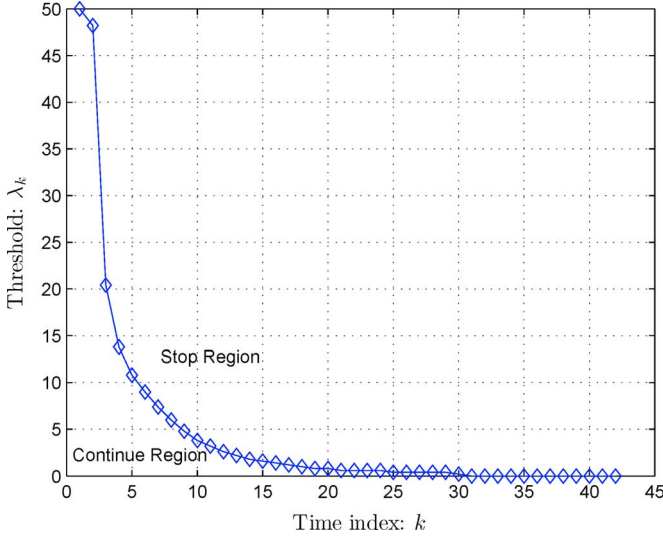
Fig. 2. Threshold over slot index k.

and performs q-dimensional feedback. See analogous results for correlated channel, and for an LMMSE estimator in [21].

Using the optimal beamformer in (14), from (9) and (15), the total harvested energy is rewritten as

$$E = (T - \tau)\left(\frac{\sigma_e^2}{1 + \sigma_e^2} + \frac{1}{(1 + \sigma_e^2)^2}\mathbb{E}_{\hat{\mathbf{h}}_q}[\|\hat{\mathbf{h}}_q\|^2]\right). \quad (32)$$

Before giving the result, we define a quantity that depends on only the number m of transmit antennas and the feedback dimension q as

$$G_{m,q} \triangleq \sum_{r=1}^{q} \frac{2m!}{(r-1)!}$$
$$\cdot \sum_{s=1}^{m-r+1} \frac{s(-1)^{s+1}}{(m-r+1-s)!s!(s+r-1)^2}. \quad (33)$$

It can be shown that $G_{m,q}$ increases as either m or q increases. In the case of full feedback, i.e., the receiver feeds back $\hat{\mathbf{h}}$ to the transmitter, we have $G_{m,m} = 2m$. In independent Rayleigh-fading MISO channels, by using the optimal beamformer in (14) at the transmitter, we have the following theorem.

Theorem 3: Let T, m, σ_z^2 and $G_{m,q}$ be defined as before. When the channel is estimated by an LS estimator, the optimal length of preambles for channel estimation is unique and is given by

$$\tau^\star = \begin{cases} 0, & \text{if } \sigma_z^2 > \frac{T(G_{m,q}-2)}{2m^2} \\ \arg\max_{\tau \in \{\lfloor \tau_1 \rfloor, \lceil \tau_1 \rceil\}} E(\tau), & \text{otherwise} \end{cases} \quad (34)$$

where the notations $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor operation and the ceiling operation, respectively, the quantity

$$\tau_1 = -m^2\sigma_z^2 + m\sqrt{\frac{\sigma_z^2(m^2\sigma_z^2 + T)(G_{m,q} - 2)}{G_{m,q}}}$$

and the function

$$E(\tau) = (T - \tau)\frac{G_{m,q}\tau + 2m^2\sigma_z^2}{2(\tau + m^2\sigma_z^2)}, \quad (35)$$

Moreover, the corresponding maximal harvested energy $E_{\max} = E(\tau^\star)$.

Proof: See proof in Appendix C. ∎

Remark 2 (Effect of Feedback Dimension q): Theorem 3 implies that larger feedback dimension q leads to more harvested energy, which is as expected. This is because given m, the constant $G_{m,q}$ increases as q increases, which can be easily shown. Moreover, we observe that the optimal preamble length $\tau^\star$ also increases as q increases. That is, longer training time is required to obtain accurate q-dimensional CSI and thus high-efficiency WPT, when more feedback is allowed.

Remark 3 (Effect of Channel Variation on the Harvested Energy): So far, we assume the channel is constant in one frame. In practice, however, the channel may change, due to the feedback delay induced by the channel, and processing at the receiver and transmitter. Here, we investigate the effect of channel variation on the harvested energy.

For simplicity, we adopt m-dimensional feedback and an LS estimator. Let $\mathbf{h}_1 \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{I}_m)$ and $\mathbf{h}_2 \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{I}_m)$ be the channel in the CE phase and the WPT phase, respectively. We assume that the temporally correlated channels follow a first-order Gauss-Markov distribution according to

$$\mathbf{h}_2 = \sqrt{\alpha}\mathbf{h}_1 + (1 - \sqrt{\alpha})\mathbf{n}, \quad (36)$$

where the constant $\alpha \in [0, 1]$ is a temporal correlation coefficient, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_m, \mathbf{I}_m)$ is an innovation process. Then, we have the following lemma.

Lemma 5: Assuming the Gauss-Markov channel-varying model in (36), the optimal beamformer is $\mathbf{w}^\star(\hat{\mathbf{h}}_1) = \frac{\hat{\mathbf{h}}_1}{\|\hat{\mathbf{h}}_1\|}$. The harvested energy in the WPT phase is given by

$$E(\tau) = (T - \tau)\left[\frac{m\alpha(\tau + m\sigma_z^2)}{\tau + m^2\sigma_z^2} + (1 - \sqrt{\alpha})^2\right]. \quad (37)$$

Proof: Recall that $\hat{\mathbf{h}}_1 = \mathbf{h}_1 + \mathbf{e}$. From (36), we have

$$\mathbf{h}_2 = \sqrt{\alpha}\hat{\mathbf{h}}_1 - \sqrt{\alpha}\mathbf{e} + (1 - \sqrt{\alpha})\mathbf{n}. \quad (38)$$

We take the sum of the second and third term in (38) as noise. Following similar steps in Section III-B and the proof for Theorem 3, we obtain Lemma 5. ∎

Specially, when the channels are constant in one frame, i.e., $\alpha = 1$, the harvested energy reduces to the result in (35) for $G_{m,m} = 2m$. Moreover, we can obtain the optimal training length $\tau^\star$ for time-varying channels, by using the same proof scheme as Theorem 3.

VI. OPTIMAL POWER ALLOCATION

Based on the derived optimal preamble length, in this section, we derive the optimal power allocation schemes for the scenario of dynamic-length preamble and fixed-length preamble, respectively.

A. Dynamic-Length-Preamble Based Power Allocation

By using dynamic-length preamble, the preamble length is typically shorter if the channel condition in one frame is good, and longer if the channel condition is bad. Intuitively, we can maximize the harvested energy by adjusting the transmit power for WPT, according to different channel conditions. In this section, we derive the optimal power allocation scheme, assuming the use of the optimal policy $\pi^\star$ for adapting the preamble length, although our subsequent results does not depend on the actual policy π used.

As in Section IV-A, we take the channel estimate power in time slot k as a random variable denoted by V_k, i.e., $V_k = \|\hat{\mathbf{h}}_k\|^2 \in \mathbb{R}^+$. Under policy $\pi^\star$, the preamble length, denoted by κ time slots for convenience, is also random. When the receiver stops the channel estimation procedure at the end of the κ-th slot, we denote the corresponding channel estimate power by $\tilde{V}_\kappa$, i.e., $\tilde{V}_\kappa = \|\hat{\mathbf{h}}_\kappa\|^2 \in \mathcal{D}_{\mathsf{s},\kappa}$.

First, we derive the joint pdf of κ and $\tilde{V}_\kappa$, denoted by $f(\tilde{v}_\kappa, \kappa \,|\, \pi^\star)$, upon using the optimal policy $\pi^\star$. For convenience, we omit the notation $\pi^\star$ in the sequel. We denote the joint pdf of $V_1, V_2, \ldots, V_{\kappa-1}, \tilde{V}_\kappa$ and κ by $f(v_1, v_2, \ldots, v_{\kappa-1}, \tilde{v}_\kappa, \kappa)$. The joint pdf $f(\tilde{v}_\kappa, \kappa)$ is given by the following recursive relation

$$f(\tilde{v}_1, 1) = \frac{\tilde{v}_1^{m-1}}{2\Gamma(m)(1+m\sigma_z^2)^{m-1}} \exp\left(\frac{-\tilde{v}_1}{1+m\sigma_z^2}\right),$$

$$(39)$$

$$f(\tilde{v}_\kappa, \kappa) = \int_{v_1 \in \mathcal{D}_{\mathsf{c},1}} \cdots \int_{v_{\kappa-1} \in \mathcal{D}_{\mathsf{c},\kappa-1}} f(v_1, \ldots, v_{\kappa-1},$$
$$\tilde{v}_\kappa, \kappa) dv_1 \cdots dv_{\kappa-1}, \quad \text{for } \tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}$$
$$\stackrel{(a)}{=} \int_{v_1 \in \mathcal{D}_{\mathsf{c},1}} f(v_1) \left[\int_{v_2 \in \mathcal{D}_{\mathsf{c},2}} f(v_2 \,|\, v_1) \right.$$
$$\times \left(\cdots \int_{v_{\kappa-1} \in \mathcal{D}_{\mathsf{c},\kappa-1}} f(v_\kappa \,|\, v_{\kappa-1}) \right.$$
$$\left. \left. \times f(v_{\kappa-1} \,|\, v_{\kappa-2}) dv_{\kappa_1} \cdots \right) dv_2 \right] dv_1, \quad (40)$$

for $\kappa = 2, \ldots, N-1$, where $\Gamma(\cdot)$ is the Gamma function. Here, $f(v_1)$ is the same as (39) with the argument replaced by $v_1 \in \mathbb{R}^+$, and $f(v_{i+1} \,|\, v_i)$ is in (17), and (a) is from Lemma 4 and the fact $\tilde{v}_\kappa = v_\kappa$, for $v_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}$.

1) Optimal Length-and-Channel-Power Aware Power Allocation: In this section, we consider the scenario in which the power is allocated according to both the optimal preamble length κ and the channel estimate power $\tilde{v}_\kappa$. We refer to this scheme as length-and-channel-power aware power allocation (LCPA). With unit transmit power, the harvested energy is from (24)

$$\tilde{E}(\tilde{v}_\kappa, \kappa) = m(N-\kappa) \left(\frac{m\sigma_z^2}{\kappa + m\sigma_z^2} + \frac{\kappa^2 \tilde{v}_\kappa}{(\kappa + m\sigma_z^2)^2} \right). \quad (41)$$

We use $p(\tilde{v}_\kappa, \kappa)$ to denote the transmit power for WPT in the frame with optimal preamble length κ and channel estimate power $\tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}$. We assume that $p(\tilde{v}_\kappa, \kappa)$ can be dynamically allocated, subject to the per-frame transmission power constraint P_1 and the average transmission power constraint P_2

over frames. To maximize the total expected harvested energy, we have the following optimization problem

$$(\text{P}_3) \quad \max_{\{p(\tilde{v}_\kappa, \kappa)\}} \quad \mathbb{E}_{\tilde{v}_\kappa, \kappa}[\tilde{E}(\tilde{v}_\kappa, \kappa) p(\tilde{v}_\kappa, \kappa)] \quad (42a)$$

$$\text{s. t.} \quad \mathbb{E}_{\tilde{v}_\kappa, \kappa}[m(N-\kappa) p(\tilde{v}_\kappa, \kappa)] \le P_2, \quad (42b)$$

$$0 \le p(\tilde{v}_\kappa, \kappa) \le P_1, \quad (42c)$$

for $\tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}$, $\kappa = 0, \ldots, N-1$. In (42b), the transmit power is utilized for WPT only in the last $(N-\kappa)$ slots.

Define $\eta(\tilde{v}_\kappa, \kappa) = \frac{\tilde{E}(\tilde{v}_\kappa, \kappa)}{m(N-\kappa)}$. Here, $\eta(\tilde{v}_\kappa, \kappa)$ is the efficiency of power transfer in the frame with optimal length κ and channel estimate power $\tilde{v}_\kappa$, which will be used as a criterion for adjusting the transmit power for WPT among frames. The optimal solution can be obtained by a greedy procedure as stated in Theorem 4.

Theorem 4: The optimal power allocation for Problem (P_3) is to allocate as much energy (up to P_1) to the frame with highest $\eta(\tilde{v}_\kappa, \kappa)$ over all $\tilde{v}_\kappa$ and all κ, then to the frame with the second highest $\eta(\tilde{v}_\kappa, \kappa)$, and so on, until the average energy constraint P_2 is satisfied.

Proof: Define $a(\tilde{v}_\kappa, \kappa) = m(N-\kappa) f(\tilde{v}_\kappa, \kappa)$, and $x(\tilde{v}_\kappa, \kappa) = a(\tilde{v}_\kappa, \kappa) p(\tilde{v}_\kappa, \kappa)$. Problem (P_3) is rewritten as

$$\max_{\{x(\tilde{v}_\kappa, \kappa)\}} \quad \sum_{\kappa=0}^{N-1} \left[\int_{\tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}} \eta(\tilde{v}_\kappa, \kappa) x(\tilde{v}_\kappa, \kappa) d\tilde{v}_\kappa \right] \quad (43a)$$

$$\text{s. t.} \quad \sum_{\kappa=0}^{N-1} \left[\int_{\tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}} x(\tilde{v}_\kappa, \kappa) d\tilde{v}_\kappa \right] \le P_2, \quad (43b)$$

$$0 \le x(\tilde{v}_\kappa, \kappa) \le a(\tilde{v}_\kappa, \kappa) P_1, \quad (43c)$$

for $\tilde{v}_\kappa \in \mathcal{D}_{\mathsf{s},\kappa}$, $\kappa = 0, \ldots, N-1$.

The transformed problem is a linear programming (LP) problem. For convenience of exposure, the channel estimate power $\tilde{v}_\kappa$ is assumed to take discrete values. We use $\mathbf{d}_{\mathsf{s}}$ to denote the decreasing sorted vector of $\text{vec}(\{\eta(\tilde{v}_\kappa, \kappa)\})$, where $\text{vec}(\cdot)$ is the vectorization operator. Let $(\tilde{v}_{\kappa_i}, \kappa_i)$ be the pair of the channel estimate power and preamble length for the i-th element in $\mathbf{d}_{\mathsf{s}}$. Thus, the optimal solution to the transferred LP problem is obtained as follows: For the first consecutive $(L-1)$ elements in $\mathbf{d}_{\mathsf{s}}$, the power allocation is P_1; for the L-th element in $\mathbf{d}_{\mathsf{s}}$, the power allocation is such that the constraint (43b) is satisfied with equality; and for other remaining elements in $\mathbf{d}_{\mathsf{s}}$, no power is allocated. Here, L is chosen to be the maximally possible. This is because for $(\tilde{v}_{\kappa_1}, \kappa_1)$, the objective function is increased the most, by setting the transmit power corresponding to $(\tilde{v}_{\kappa_1}, \kappa_1)$ as the maximally possible, after which the transmit power for $(\tilde{v}_{\kappa_2}, \kappa_2)$ is set as the maximally possible, and so on, until the average power constraint P_2 is satisfied. ∎

Remark 4 (Complexity of Power Allocation Scheme): Following the greedy procedure in Theorem 4, the allocated power for each combination $(\kappa, \tilde{v}_\kappa)$ is pre-computed and stored in a two-dimensional look-up table. In implementation, the WP transmitter obtains the transmit power by referring to the lookup table. Hence, the complexity during implementation is relatively low.

2) Optimal Length-Aware Power Allocation: To further reduce implementation complexity, we consider a simplified power allocation scheme in which the power is allocated

according to only the optimal preamble length κ, referred as length-aware power allocation (LPA). Compared to the general Problem (P$_3$), we herein restrict $p(\tilde{v}_\kappa, \kappa) = p(\kappa)$, independent of the channel estimate power $\tilde{v}_\kappa$. As in the LCPA scheme, we employ the optimal policy $\pi^\star$. Then, the probability that the optimal preamble length is κ, is given by

$$q_\kappa = \int_{\tilde{v}_\kappa \in \mathcal{D}_{s,\kappa}} f(\tilde{v}_\kappa, \kappa) d\tilde{v}_\kappa. \tag{44}$$

With unit transmit power, the average harvested energy from frames of preamble length κ is given by

$$\bar{Q}_{\mathrm{harv},\kappa} = \mathbb{E}_{\tilde{v}_\kappa}[\tilde{E}(\tilde{v}_\kappa, \kappa)]. \tag{45}$$

Problem J$_1$ is thus simplified as

$$\begin{aligned} (\mathrm{P}_4) \quad & \max_{\{p(\kappa)\}} \quad \mathbb{E}_\kappa\left[p(\kappa)\bar{Q}_{\mathrm{harv},\kappa}\right] \\ & \text{s.t.} \quad \mathbb{E}_\kappa[m(N - \kappa)p(\kappa)] \leq P_2, \\ & \qquad 0 \leq p(\kappa) \leq P_1, \quad \forall \kappa. \end{aligned} \tag{46}$$

Let $\eta(\kappa) = \frac{\bar{Q}_{\mathrm{harv},\kappa}}{m(N-\kappa)}$. Similar to Theorem 4, the solution to Problem (P$_4$) is given without proof as in Proposition 1.

Proposition 1: The optimal power allocation for Problem (P$_4$) is to allocate as much energy (up to P_1) to the frame with highest $\eta(\kappa)$ over all κ, then to the frame with the second highest $\eta(\kappa)$, and so on, until the average energy constraint P_2 is satisfied.

B. Fixed-Length-Preamble Based Power Allocation

In the fixed-length preamble scenario considered here, the optimal preamble length (i.e., $\kappa = \frac{\tau^\star}{m}$ time slots) is obtained in Section V, and henceforth used for all frames. Here, we consider the power allocation according to only the channel estimate power $\tilde{v}_\kappa$, referred as channel-power-aware power allocation (CPA). For consistence, we use the same notations as the LCPA scheme in Section VI-A1, with the only difference here that the preamble length κ is fixed.

Let $p_\kappa(\tilde{v}_\kappa)$ denote the transmit power for WPT in the frame with channel estimate power $\tilde{v}_\kappa$. After obtaining $\tilde{v}_\kappa$ via feedback, the transmitter performs WPT with transmit power $p_\kappa(\tilde{v}_\kappa)$ in the current frame. With the same power constraint P_1 and P_2 as in Problem (P$_3$), we have the following problem formulation

$$\begin{aligned} (\mathrm{P}_5) \quad & \max_{\{p_\kappa(\tilde{v}_\kappa)\}} \quad \mathbb{E}_{\tilde{v}_\kappa}[\tilde{E}_\kappa(\tilde{v}_\kappa, \kappa)p_\kappa(\tilde{v}_\kappa)] \end{aligned} \tag{47a}$$
$$\text{s. t.} \quad m(N - \kappa)\mathbb{E}_{\tilde{v}_\kappa}[p_\kappa(\tilde{v}_\kappa)] \leq P_2, \tag{47b}$$
$$0 \leq p_\kappa(\tilde{v}_\kappa) \leq P_1, \quad \text{for } \tilde{v}_\kappa \in \mathbb{R}^+. \tag{47c}$$

We note that given κ, the harvested energy $\tilde{E}_\kappa(\tilde{v}_\kappa, \kappa)$ in (41) is a monotonically increasing function of the channel estimate power $\tilde{v}_\kappa$. Similar to Theorem 4, the solution to Problem (P$_5$) is given below without proof.

Theorem 5: The optimal power allocation for Problem (P$_5$) is to allocate as much energy (up to P_1) to the frame with highest channel estimate power $\tilde{v}_\kappa$ over all $\tilde{v}_\kappa$, then to the frame with the second highest $\tilde{v}_\kappa$, and so on, until the average energy constraint P_2 is satisfied.
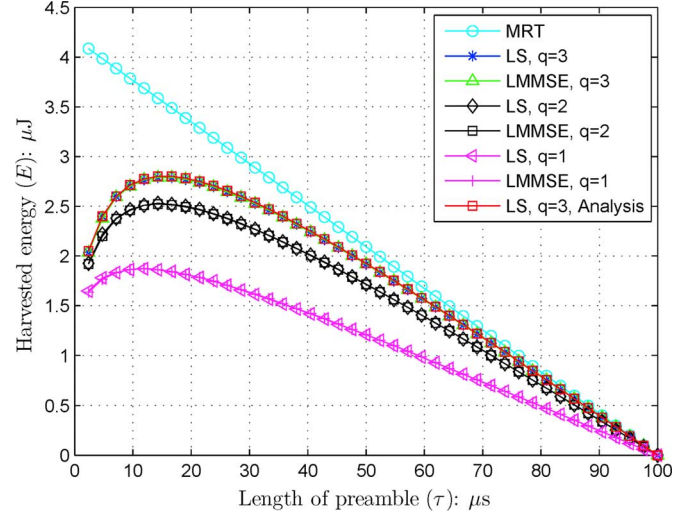

Fig. 3. Harvested energy in uncorrelated channel.

VII. NUMERICAL RESULTS

In this section, we present numerical results to validate our results. We set the number of transmit antennas $m = 3$. We assume the time duration for the CE and WPT phases in each frame is 100 μs, which consists of $T = 126$ symbol periods (equivalently, $N = 42$ time slots). The carrier frequency is 5 GHz, and the bandwidth is 100 KHz. We set the power spectrum density of noise as -113 dBm/Hz, which implies the noise power $\sigma_z^2 = -63$ dBm. We take the path loss model as $10^{-3}D^{-3}$, where the path loss exponent is 3, and $D = 2.1$ m is the distance between the WP transmitter and WP receiver. A 30 dB path loss is assumed at a reference distance of 1 m. We employ LS channel estimation given in this paper, as well as LMMSE channel estimation, see details in [21].

First, we simulate the harvested energy using the scheme based on the fixed-length preamble in Section V, but without the adaptive power allocation in Section VI-B. We fix the transmit power as $P_0 = 1$ Watt.

We start from an uncorrelated MISO channel. Fig. 3 plots the harvested energy for different dimension q of CSI feedback. With perfect CSI at the transmitter, the maximum ratio transmit (MRT) beamforming scheme harvests most energy, which provides an upper bound for all schemes that use fixed-length preamble. The $\square$-maker curve is plotted according to (35) in Theorem 3 for different preamble length τ. From (34) in Theorem 3, the optimal preamble length is $\tau^\star = 14$ μs, and the maximum harvested energy is 2.8 μJ. The simulation results ($*$-maker curve) coincide with the analytical results. Moreover, the harvested energy is reduced as the dimension q of CSI feedback decreases. Also, we observe that the LS based WPT achieves the same performance as the LMMSE-based WPT scheme as expected, since the channel is uncorrelated.

Next, we assume a correlated MISO channel, with channel correlation matrix that has the structure: $[\mathbf{R}]_{i,j} = \varrho^{|i-j|}$, $0 \leq \varrho < 1$, where i and j are the indices of the entries [20]. We set the correlation parameter $\varrho = 0.8$. The harvested energy is plotted in Fig. 4. We observe that the LMMSE-based scheme transfers more energy than the LS-based WPT in general, due
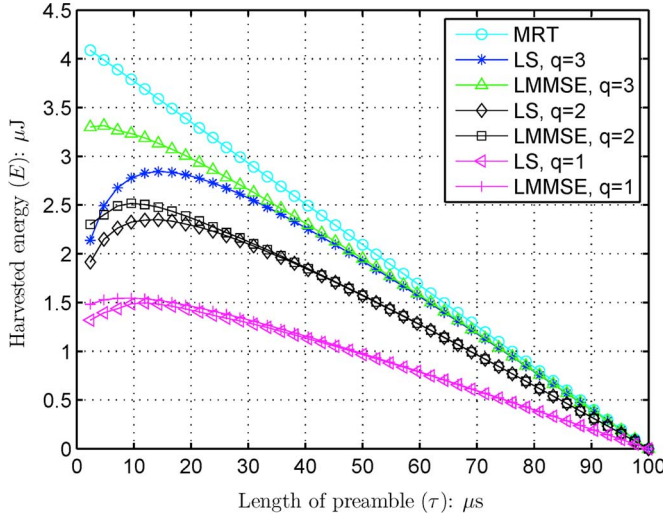
Fig. 4.　Harvested energy in correlated channel.

Fig. 5.　Comparison of FwoPA, LCPA, LPA and CPA.

to the fact that an LMMSE estimator achieves more accurate channel estimation than an LS estimator.

Second, as shown in Fig. 5, we compare the length-and-channel-power aware power allocation (LCPA) scheme in Section VI-A1, to the length-aware power allocation scheme (LPA) in Section VI-A2, as well as the scheme based on the optimized fixed-length preamble with channel-power-aware power allocation (CPA) in Section VI-B. We take the harvested energy by using the optimized fixed-length preamble without power allocation (FwoPA), as a benchmark. The schemes with power allocation are shown to achieve significant increase in harvested energy, compared to the FwoPA scheme. Moreover, the CPA scheme and the LCPA scheme harvest almost the same amount of energy. This is because in the CPA scheme, the optimal preamble length is obtained after averaging all possible channel realizations, and the dynamical nature of the channels is fully exploited by the CPA scheme. We also find it encouraging to observe that the energy harvested by using the low-complexity LPA scheme is close to that harvested by using the optimal LCPA scheme or the CPA scheme, although the dynamical nature of the channels is only partially exploited by the LPA scheme.

VIII. Conclusion

The paper studies a MISO system where the transmitter delivers power to the receiver via energy beamforming, and the harvested energy is used by the receiver to do work. To maximize the harvested energy, we first derive the optimal energy beamformer. Then, we perform dynamic optimization for the preamble length, and also obtain the optimal offline (fixed) preamble length to reduce the complexity. Moreover, we derive the optimal power allocation schemes for wireless power transfer with dynamic-length preamble and fixed-length preamble, respectively. As future extension of this paper, we have considered the uplink data transmission powered by downlink WPT in a multiuser massive MIMO system that consists of a hybrid data-and-energy access point with a large number of antennas and multiple single-antenna users [26].
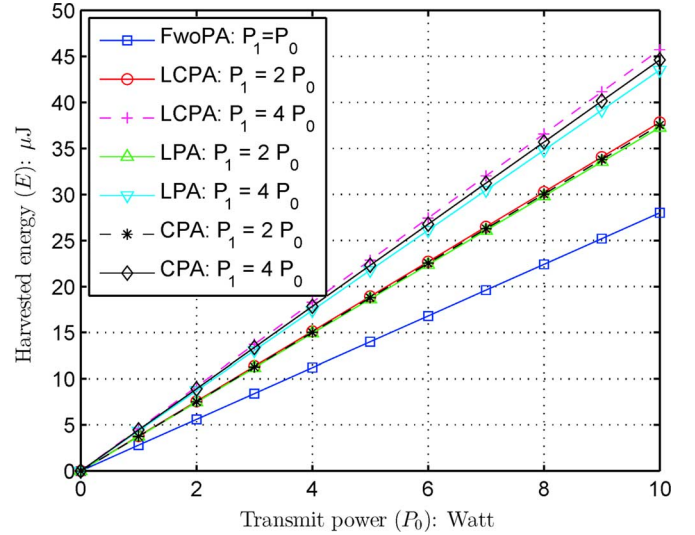
Appendix A
Proof for Lemma 4

Proof: For some antenna index, let $h, \hat{h}_r, \hat{h}_{r+1}$ be the channel coefficient, and the channel estimate in time slot r and $r + 1$, respectively. From Bayes' formula and Lemma 3, it is straightforward to show that

$$h \,|\, \hat{h}_r, \hat{h}_{r+1} \sim \mathcal{CN}\left(\frac{(r+1)\hat{h}_{r+1}}{r + 1 + m\sigma_z^2}, \frac{m\sigma_z^2}{r + 1 + m\sigma_z^2} \right). \quad (48)$$

From Lemma 1, we obtain that conditioned on $\hat{\mathbf{h}}_{r+1}$, the $\mathbf{h}$ has the same distribution as in (48). Thus, we have $f(h \,|\, \hat{h}_r, \hat{h}_{r+1}) = f(h \,|\, \hat{h}_{r+1})$. We further obtain by mathematical induction that $f(h \,|\, h_1, h_2, \ldots, h_k) = f(h \,|\, h_k)$, for $k = 1, \ldots, N - 1$. The independence between elements completes this proof. ∎

Appendix B
Proof for Theorem 1

Proof: We first consider the Policies 1 and 2, as follows. Policy 1 has a decision sub-sequence $(\mathsf{c}, \mathsf{s}, \mathsf{c})$ over slots $r - 1, r, r + 1$. The corresponding states are $\mathbf{x}_r, \mathbf{x}_{r+1}$ and $\mathbf{x}_{r+2}$, where $\mathbf{x}_{r+1} = \mathbf{x}_r$ because from (20) the state value remains the same when $u_r = \mathsf{s}$. Policy 2 is exactly the same policy as Policy 1, except that given state $\mathbf{x}_r$ in slot r, Policy 2 performs CE followed by WP regardless of the state in slot $k + 1$. Thus, the decision subsequence becomes $(\mathsf{c}, \mathsf{c}, \mathsf{s})$. We aim to show that Policy 2 has strictly higher expected harvested energy than Policy 1. Both policies are statistically equivalent in slot $r + 2$ and onwards because both have used the same number of slots for CE; hence the expected harvested energy in slot $r + 2$ onwards are the same. It thus suffices to compare the expected harvested energy of Policy 1 in slot r, denoted by $E_1^r(\mathbf{x}_r)$, and that of Policy 2 in slot $r + 1$, denoted by $E_2^{r+1}(\mathbf{x}_r)$. For Policy 1, from (21), the expected harvested energy is

$$E_1^r(\hat{\mathbf{h}}_r) = m\left(\frac{m\sigma_z^2}{r + m\sigma_z^2} + \frac{r^2 \left\| \hat{\mathbf{h}}_r \right\|^2}{(r + m\sigma_z^2)^2} \right). \quad (49)$$

For Policy 2, the channel estimate $\hat{\mathbf{h}}_{r+1}$ in the next slot $k+1$ is introduced. Hence, the expectation for the harvested energy is taken over the conditional distribution $p(\mathbf{h}, \hat{\mathbf{h}}_{r+1} \,|\, \hat{\mathbf{h}}_r)$ as follows

$$
\begin{aligned}
E_2^{r+1}&(\mathbf{x}_r) \\
&\triangleq m\mathbf{E}_{\hat{\mathbf{h}}_{r+1}\,|\,\hat{\mathbf{h}}_r}\left[\mathbf{E}_{\mathbf{h}\,|\,\hat{\mathbf{h}}_{r+1},\hat{\mathbf{h}}_r}[(\mathbf{w}_{r+1}^\star)^H\mathbf{h}\mathbf{h}^H\mathbf{w}_{r+1}^\star]\right] \\
&\stackrel{(a)}{=} m\mathbf{E}_{\hat{\mathbf{h}}_{r+1}\,|\,\hat{\mathbf{h}}_r}\left[\mathbf{E}_{\mathbf{h}\,|\,\hat{\mathbf{h}}_{r+1}}[(\mathbf{w}_{r+1}^\star)^H\mathbf{h}\mathbf{h}^H\mathbf{w}_{r+1}^\star]\right] \\
&\stackrel{(b)}{=} m\left(\frac{m\sigma_z^2}{r+1+m\sigma_z^2} + \frac{(r+1)^2\mathbf{E}_{\hat{\mathbf{h}}_{r+1}\,|\,\hat{\mathbf{h}}_r}[\|\hat{\mathbf{h}}_{r+1}\|^2]}{(r+1+m\sigma_z^2)^2}\right) \\
&\stackrel{(c)}{=} m\left(\frac{m\sigma_z^2\left(r+m+m\sigma_z^2\right)}{(r+m\sigma_z^2)(r+1+m\sigma_z^2)} + \frac{r^2\|\hat{\mathbf{h}}_r\|^2}{(r+m\sigma_z^2)^2}\right), \quad (50)
\end{aligned}
$$

where (a) comes from Lemma 4, (b) follows (15), and (c) is from the conditional mean in (18). We conclude that Policy 1 is strictly worse than Policy 2, since $E_2^{r+1}(\mathbf{x}_r) - E_1^r(\mathbf{x}_r) = \frac{m^2(m-1)\sigma_z^2}{(r+m\sigma_z^2)(r+1+m\sigma_z^2)} > 0$.

The same argument extends to the case in which there are more than one slot with decision s in between the two slots with decision c. Theorem 1 must then hold; otherwise, there exists a decision subsequence with a structure that was shown to be suboptimal. ∎

APPENDIX C
PROOF OF THEOREM 3

Proof: From (9) and (15), using the optimal beamformer in (7), the total harvested energy is

$$
\hat{E}(\beta) = (T - \beta m^2)\left(\frac{\sigma_z^2}{\beta + \sigma_z^2} + \frac{\beta^2\sum_{r=1}^{q}\mathbb{E}_{\hat{h}_{i_r}}\left[\left|\hat{h}_{i_r}\right|^2\right]}{(\beta + \sigma_z^2)^2}\right).
$$

Denote $\mathbf{u} = \frac{2}{1+\sigma_z^2/\beta}[|\hat{h}_1|^2 \;\ldots\; |\hat{h}_m|^2]^T$. Elements of $\mathbf{u}$ are independent Chi-Square random variables, since $\hat{h}_i$'s are independent zero-mean complex Gaussian random variables with variance $\left(1+\frac{\sigma_z^2}{\beta}\right)$. Let u_{i_r} denote the random variable corresponding to the r-th largest observation of the m original random variables. Denote $C_{m,r} = \frac{m!}{(m-r+1)!(r-1)!}$. From order statistics, we have

$$
\mathbb{E}[u_{i_r}] = \frac{2m!}{(r-1)!}\sum_{s=1}^{m-r+1}\frac{s(-1)^{s+1}}{(m-r+1-s)!s!(r+s-1)^2}.
$$

Denote $G_{m,q} \triangleq \sum_{r=1}^{q}\mathbb{E}[u_{i_r}]$. We have that $g_{m,1}$ is no less than 2. Moreover, $G_{m,m} = 2m$, since $\sum_{r=1}^{m}\mathbb{E}[u_{i_r}]$ is the variance of a (m degrees of freedom) Chi-Square random variable. Then we have

$$
\hat{E}(\beta) = (T - \beta m^2)\frac{G_{m,q}\beta + 2\sigma_z^2}{2(\beta + \sigma_z^2)}. \quad (51)
$$

Moreover, it is standard to show (51) is quasi-concave function of β. Setting the first-order derivative of $\hat{E}(\beta)$ to be zero,

the $\hat{E}(\beta)$ is maximized at the unique positive solution which is given as follows

$$
\beta_1 = -\sigma_z^2 + \sqrt{\frac{\sigma_z^2}{m^2 G_{m,q}}\left(m^2\sigma_z^2 + T\right)\left(G_{m,q} - 2\right)} \quad (52)
$$

Define $E(\tau) \triangleq \hat{E}'(\tau/m^2)$. Let $\tau_1 = m^2\beta_1$.

Since the preamble length should be multiples of the number of transmit antennas m, we obtain the optimal preamble length as $\tau^\star = \arg\max_{\tau\in\{\lfloor\tau_1\rfloor,\lceil\tau_1\rceil\}} E(\tau)$, if $\sigma_z^2 \leq \frac{T(G_{m,q}-2)}{2m^2}$; and $\tau^\star = 0$ otherwise. Thus, the maximum harvested energy $E_{\max} = E(\tau^\star)$. ∎

REFERENCES

[1] N. Shinohara, "Power without wires," *IEEE Microw. Mag.*, vol. 12, pp. S64–S73, Dec. 2011.

[2] N. Shinohara, *Wireless Power Transfer via Radiowaves*. New York, NY, USA: Wiley, 2014.

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

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

[5] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Commun.*, vol. 9, pp. 1326–1336, Apr. 2010.

[6] G. Yang, V. Y. F. Tan, C. K. Ho, S. H. Ting, and Y. L. Guan, "Wireless compressive sensing for energy harvesting sensor nodes," *IEEE Trans. Signal Process.*, vol. 61, pp. 4491–4505, Sep. 2013.

[7] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 1989–2001, May 2013.

[8] L. Liu, R. Zhang, and K. C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 288–300, Jan. 2013.

[9] K. Huang and E. G. Larsson, "Simultaneous information and power transfer for broadband wireless systems," *IEEE Trans. Signal Process.*, vol. 61, pp. 5972–5986, Dec. 2013.

[10] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 418–428, Jan. 2014.

[11] X. Kang, C. K. Ho, and S. Sun, "Full-duplex wireless-powered communication network with energy causality," [Online]. Available: arxiv.org/abs/1404.0471, submitted for publication

[12] Z. Xiang and M. Tao, "Robust beamforming for wireless information and power transmission," *IEEE Wireless Commun. Lett.*, vol. 1, pp. 372–375, Aug. 2012.

[13] D. Li, C. Shen, and Z. Qiu, "Sum rate maximization and energy harvesting for two-way AF relay systems with imperfect CSI," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Vancouver, Canada, May 2013, pp. 4958–4962.

[14] D. W. K. Ng, E. S. Lo, and R. Scholar, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, pp. 1–18, Apr. 2014, vol. PP, preprint.

[15] S. X. Wu, Q. Li, W.-K. Ma, and A. M. C. So, "Robust transmit designs for an energy harvesting multicast system," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)* [Online]. Available: www1.se.cuhk.edu.hk/~manchoso/papers/robust_SBF_EH-icassp14.pdf, to be published

[16] X. Chen, C. Yuen, and Z. Zhang, "Wireless energy and information transfer tradeoff for limited feedback multi-antenna systems with energy beamforming," *IEEE Trans. Veh. Technol.*, vol. 63, pp. 407–412, Jan. 2014.

[17] J. Xu and R. Zhang, "Energy beamforming with one-bit feedback," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)* [Online]. Available: arxiv.org/abs/1312.1444, to be published

[18] G. Yang, C. K. Ho, and Y. L. Guan, "Utility maximization for wirelessly powered multiple-antenna systems," in *Proc. Workshop on Energy Harvesting Commun. 2014 IEEE Int. Conf. Sensing, Commun., Netw. (SECON)*, to be published.

[19] A. P.-Iserte, D. P. Palomar, A. I. P.-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, pp. 346–360, Jan. 2006.

[20] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, pp. 884–893, Mar. 2006.

[21] G. Yang, C. K. Ho, and Y. L. Guan, "Dynamic resource allocation for multiple-antenna wireless power transfer," [Online]. Available: arxiv.org/abs/1311.4111, full version.

[22] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[23] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA, USA: Athena Scientific, 2005.

[24] C. K. Ho, J. Oostveen, and J.-P. M. G. Linnartz, "Rate adaptation using acknowledgement feedback in finite-state Markov channels with collisions," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 3226–3239, June 2009.

[25] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2007.

[26] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan, "Throughput optimization for massive MIMO systems powered by wireless energy transfer," [Online]. Available: arxiv.org/abs/1403.3991, submitted for publication

harvesting constraints, wireless information and power transfer, compressive sensing and statistical signal processing. He has served as a reviewer for journals such as IEEE JSAC, IEEE JSTSP, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, etc.

Chin Keong Ho (S'05–M'07) received the B.Eng. (First-Class Hons., Minor in Business Admin.), and M.Eng. degrees from the Department of Electrical Engineering, National University of Singapore in 1999 and 2001, respectively. He obtained his Ph.D. degree at the Eindhoven University of Technology, The Netherlands, where he concurrently conducted research work in Philips Research. Since August 2000, he has been with Institute for Infocomm Research (I^{2}R), A*STAR, Singapore. He is Lab Head of Energy-Aware Communications Lab, Department of Advanced Communication Technology, in I^{2}R. His research interest includes green wireless communications with focus on energy-efficient solutions and with energy harvesting constraints; cooperative and adaptive wireless communications; and implementation aspects of multi-carrier and multi-antenna communications.

Gang Yang (S'13) received the B.Eng. and M.Eng. degrees (First-Class Hons.) in communication engineering, communication and information systems in 2008 and 2011, respectively, from the University of Electronic Science and Technology of China (UESTC), Chengdu, China. Since 2011, he has been pursuing a Ph.D. degree in Infocomm Centre of Excellence of the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include green wireless communications with energy

Yong Liang Guan (M'99) received his Ph.D. degree from the Imperial College of Science, Technology and Medicine, University of London, in 1997, and B.Eng. degree with first class honors from the National University of Singapore in 1991. He is now an associate professor at the School of Electrical and Electronic Engineering, Nanyang Technological University. His research interests include modulation, coding and signal processing for communication, information security and storage systems. The author's homepage is available online at http://www3.ntu.edu.sg/home/eylguan.