Stochastic Optimal Control for Wireless Powered Communication Networks

Kae Won Choi, Senior Member, IEEE, and Dong In Kim, Senior Member, IEEE

Abstract—In this paper, we propose a stochastic optimal control algorithm for the wireless powered communication networks (WPCNs), in which the access point (AP) supplies energy to wireless nodes by means of the RF energy transfer technology. The energy beamforming is used to enhance the RF energy transfer efficiency by concentrating the radiated power on target nodes. Each wireless node is equipped with an energy queue and a data queue. We propose an algorithm that minimizes the expected energy transmission power from the AP while stabilizing the data queues of all nodes. The proposed algorithm is an online algorithm that adaptively decides the beamforming vector, the data scheduling, and the data transmission power, only based on the current state of the energy and the data queues. The proposed algorithm dynamically steers the energy beam to nodes that currently have low energy in the energy queue. We apply the Lyapunov optimization technique to design such an algorithm. We mathematically prove that the proposed algorithm achieves the optimal performance.

Index Terms—Wireless powered communication networks, Lyapunov optimization, RF energy transfer, energy beamforming.

I. INTRODUCTION

Recently, the radio frequency (RF) energy transfer has attracted great attention as an enabling technology for far-field wireless energy transfer [1]. The RF energy transfer technology is advantageous in many practical aspects compared to other wireless transfer technologies, e.g., magnetic resonant and inductive coupling [2]. Most of all, the charging range of the RF energy transfer is much wider than that of other technologies. In addition, the RF energy harvesting receiver has a small form factor and can be implemented at a low cost [3]. With these advantages, the RF energy transfer is considered as a promising technology for charging low-power Internet of Things (IoT) devices without cords [4]. Despite of its potential, the RF energy transfer technology suffers from low energy transfer efficiency mainly caused by a propagation loss.

Manuscript received April 12, 2015; revised September 4, 2015; accepted September 4, 2015. Date of publication September 9, 2015; date of current version January 7, 2016. This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIP) under Grant 2014R1A5A1011478. The associate editor coordinating the review of this paper and approving it for publication was Prof. Juan Manuel Romero-Jerez.

K. W. Choi is with the Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul 139-743, South Korea (e-mail: kaewon.choi@gmail.com).

D. I. Kim is with the School of Information and Communication Engineering, Sungkyunkwan University (SKKU), Suwon 440-746, South Korea (e-mail: dikim@skku.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2477351

Therefore, it is of great importance to enhance the energy transfer efficiency for fully realizing the potential of the RF energy transfer technology.

One of research directions for achieving high energy transfer efficiency is to concentrate the radiated power on the target wireless node by forming a microwave beam. This beamforming technique for power transfer has been extensively studied (e.g., [5]–[11]). These works consider a multiple-input-single-output (MISO) system with a simultaneous wireless information and power transfer (SWIPT) capability. The authors have formulated an optimization problem, the objective and the constraints of which are the transmission power, the signal-to-interference-noise ratio (SINR), and the harvested energy. The common approach to solve this optimization problem is the semidefinite relaxation (SDR) technique [5]–[9]. The suboptimal zero-forcing (ZF) beamforming is considered in [9], [10] and the dynamic programming (DP) is used when the channel state information (CSI) is imperfect [11].

The wireless powered communication network (WPCN) is another research direction of the RF energy transfer technology (e.g., [12]–[15]). In the WPCN, the access point (AP) transfers the RF energy to wireless nodes in the downlink, and the nodes sends data to the AP in the uplink by using the harvested energy. In [12], the authors have studied a throughput maximization scheme that optimizes the time allocation of the downlink RF energy transfer and the uplink data transmission. The authors of [13] have derived the limiting distribution of the stored energy in the energy buffer in the WPCN. There are a few works regarding the WPCN with downlink energy beamforming [14], [15]. These works solve the problem of jointly optimizing the time allocation and the downlink energy beamforming to maximize the throughput.

In this paper, we propose a stochastic optimal control algorithm for the WPCN with the energy beamforming, in the system model described in Fig. 1. The proposed algorithm minimizes the energy transmission power while satisfying the minimum data rates. The prominent feature of the proposed algorithm, compared to the other existing works (e.g., [14], [15]), is that it decides the beamforming vector, the data scheduling, and the power allocation, only based on the current states of the energy queue and the data queue of each node. The state of the energy queue means the remaining energy in a battery or a capacitor. In the proposed algorithm, the AP dynamically focuses a microwave beam on nodes that currently have low remaining energy in the energy queue. In addition, the proposed algorithm allocates more data transmission time to nodes with low remaining energy so that these nodes can use energy more efficiently. We will prove that this dynamic and
instantaneous operation of the proposed algorithm leads to the optimal performance.

In related work, stochastic control algorithms based on the state of the energy and the data queues have been proposed for various wireless networks with a renewable energy source (e.g., [16]–[20]). The WPCN with the energy and the data queues is modeled as a Markov decision process (MDP) in [16] and as a decentralized partially observable Markov decision process (Dec-POMDP) in [17]. The authors of [18] have formulated the cognitive radio system with the RF energy harvesting as an MDP. In [19] and [20], the optimal energy management scheme is studied for the energy harvesting sensor node with the energy and the data queues. However, none of these works considers the downlink energy beamforming.

In this paper, we use the Lyapunov optimization [21], [22] to design the stochastic optimal control algorithm. We define a quadratic Lyapunov function that grows larger as the data queue size increases or the energy queue size decreases. The queue sizes can be stabilized if the algorithm takes an action that minimizes the drift of the Lyapunov function (i.e., Lyapunov drift). The objective of the proposed algorithm is to minimize the energy transmission power from the AP. To achieve this objective, we define the drift–plus-penalty function that is the summation of the Lyapunov drift and the penalty on the energy transmission power. In each time frame, the proposed algorithm decides the beamforming vector, the data scheduling, and the data transmission power in such a way that the drift–plus-penalty function is minimized. The resulting algorithm will be presented in Algorithm 1 and 2 in Section IV-D. In Section IV-E, we will prove that the proposed Lyapunov optimization-based algorithm minimizes the expected energy transmission power while stabilizing the data queues.

The proposed Lyapunov optimization-based algorithm has the following two advantages over the existing optimal solution-based algorithms for the WPCN with the beamforming in [14] and [15]. First, the proposed algorithm does not need the prior knowledge of the channel state distribution and the data rate requirements of nodes. On the other hand, the existing algorithms should know this information to formulate and solve the optimization problem. Second, the proposed algorithm is an online algorithm that requires low-complexity computation at each time frame. However, the existing algorithms have a high computational overhead since they need to solve a semidefinite programming (SDP) problem. Moreover, the complexity of the existing algorithms can be prohibitively high when the channel state varies fast over time (i.e., coherence time is relatively short). Due to these advantages, the proposed algorithm is more appropriate for the practical application than the existing algorithms are.

The rest of the paper is organized as follows. We present the system model in Section II. In Section III, we introduce a simple energy transmission power minimization problem to provide a useful insight. The proposed stochastic optimal control algorithm is explained in Section IV. Section V presents numerical results, and the paper is concluded in Section VI.

**II. SYSTEM MODEL**

**A. Network Model**

We consider a wireless powered communication network (WPCN) with one access point (AP) and \( N \) nodes, as depicted in Fig. 1. The AP transmits RF energy beams to nodes in the downlink direction and receives data from nodes in the uplink direction. A node is a low power wireless node that is powered only by the RF energy from the AP. There is no power source other than the RF energy harvesting. A node sends data to the AP in a time-division multiple access (TDMA) manner by using the harvested RF energy.

The AP is equipped with \( K \) transmit antennas for RF energy transmission, while each node has only one antenna for energy harvesting.\(^1\) We do not consider the simultaneous wireless information and power transfer (SWIPT). Therefore, no information is conveyed by the energy signal from the AP. The energy transmission system in the AP is a phased antenna array that sends a tone signal. The energy signal is transmitted via a frequency channel with very small bandwidth since it is a tone signal.

Each node sends data to the AP in a TDMA manner. Time is divided into frames. The frame duration is \( T \) and the index of each frame is denoted by \( t \). Only one node can transmit data in each frame. For data transmission, nodes use a frequency channel of bandwidth \( W \). This frequency channel for data transmission is different from the one for RF energy transmission. The AP receives data from nodes by using a single receive antenna dedicated for data reception. Therefore, there are total \( K + 1 \) antennas at the AP: \( K \) transmit antennas for energy transmission and one receive antenna for data reception.

For the following reason, we assume the TDMA instead of the space-division multiple access (SDMA) for the uplink data transmission while the AP uses the transmit beamforming for the energy transmission. It is well known that the coverage areas of the energy transmission and the data communication are wildly different [23]. Therefore, it is justifiable to use the beamforming only for the energy transmission to extend the energy transmission coverage. Moreover, we can use a relatively simple analog circuit for the energy transmission since the energy

\(^1\)Although installing more transmit antennas at the AP for the energy transmission leads to an increased cost, more energy can be harvested by a node for the same energy transmission power as we increase the number of transmit antennas.
B. RF Energy Transmission Model

The AP is equipped with the phased antenna array system with $K$ transmit antennas for the energy transmission. We assume that the AP can change the beamforming vector on a frame-by-frame basis. We define the beamforming vector at frame $t$ as

$$\mathbf{w}_t = (w_{t,1}, \ldots, w_{t,K})^T \in \mathbb{C}^{K \times 1}. \quad (1)$$

Since the AP transmits a tone signal, the complex baseband signal from antenna $k$ is $w_{t,k}$, which is constant over time during a frame. The energy transmission power from the AP at frame $t$ is

$$P_t = \|\mathbf{w}_t\|^2_F = \mathbf{w}_t^H \mathbf{w}_t. \quad (2)$$

The energy channel vector from the AP to node $n$ is denoted by $\mathbf{g}_{n} \in \mathbb{C}^{1 \times K}$. We assume that the energy channel vector is independent and identically distributed (i.i.d.) over frames and nodes. At frame $t$, the baseband complex signal received by node $n$ is $\mathbf{g}_{n} \mathbf{w}_t$ and the received power is $\|\mathbf{g}_{n} \mathbf{w}_t\|^2$. If the energy harvesting efficiency is denoted by $\eta$, the harvested energy by node $n$ at frame $t$ is given by

$$R_{t,n} = \eta T |\mathbf{g}_{n} \mathbf{w}_t|^2. \quad (3)$$

We assume that the AP has an exact knowledge of the energy channel vector at the current frame. In practice, the energy channel vector can be estimated in various ways. For example, if the AP can send an orthogonal training symbol for each antenna, nodes can estimate the energy channel vector from the training symbols and send feedback to the AP. The AP can also estimate the energy channel vector via uplink training symbols in the same frequency by exploiting the channel reciprocity [24]. In [25], the one-bit feedback method is proposed for the case that nodes are limited in hardware implementation. The detailed implementation of the channel estimation method is out of the scope of this paper.

In our model, a node cannot harvest energy of the data transmission signal from other nodes for the following reason. Generally, an energy harvesting circuit is tuned to a specific frequency band by impedance matching. Then, most of the energy of the electromagnetic wave in the unintended frequency band is reflected from the energy harvesting antenna. Therefore, the energy harvesting efficiency of the data transmission signal from other nodes is very low since the data transmission signal uses a different frequency band.

C. Data Transmission Model

Each node transmits data to the AP in a TDMA manner. Let $H_{t,n}$ denote the scheduling indicator that is one if node $n$ is scheduled at time $t$; and is zero otherwise. Therefore, it should be satisfied that $H_{t,n} \in \{0, 1\}$ and $\sum_{n=1}^{N} H_{t,n} \leq 1$. When node $n$ is scheduled at frame $t$, it sends a data signal by using the data transmission energy $C_{t,n}$ during the frame duration of $T$. Note that the data transmission power is $C_{t,n}/T$. If $H_{t,n} = 0$, we have $C_{t,n} = 0$.

The data channel gain from node $n$ to the AP at frame $t$ is denoted by $h_{t,n}$. We assume that the data channel gain is i.i.d. over frames and nodes. The received power of the data signal from node $n$ is $|h_{t,n}|^2 C_{t,n}/T$. We assume that the data rate is decided according to the Shannon capacity formula. We define the data transmission rate, denoted by $D_{t,n}$, as the number of data bits successfully transmitted from node $n$ to the AP at frame $t$. We can calculate $D_{t,n}$ as

$$D_{t,n} = \theta(C_{t,n}; h_{t,n}), \quad (4)$$

where $\theta$ is a data rate function such that

$$\theta(x; h) = WT \log_2 \left(1 + \frac{|h|^2 x}{N_0 WT} \right). \quad (5)$$

In (5), $N_0$ is the noise spectral density.

D. Energy Queue and Data Queue Model

Each node has an energy queue for storing the harvested RF energy. The energy queue is a conceptual queue to represent an amount of energy stored in the node, and is physically a rechargeable battery or a supercapacitor. Let us define $B_{t,n}$ as the amount of the energy stored in the energy queue of node $n$ at the start of frame $t$.

When scheduled for data transmission at frame $t$, node $n$ uses energy stored in the energy queue for data transmission. The amount of the energy used for data transmission is $C_{t,n}$. Since a node cannot use more energy than the energy stored in the energy queue, it should be satisfied that

$$C_{t,n} \leq B_{t,n}. \quad (6)$$

Node $n$ harvests the RF energy transmitted from the AP and stores it in the energy queue. The amount of the energy harvested by node $n$ at frame $t$ is $R_{t,n}$ as given in (3). Therefore, the energy queue dynamics is

$$B_{t+1,n} = B_{t,n} - C_{t,n} + R_{t,n}. \quad (7)$$
Each node also has a data queue for keeping data until it is transmitted to the AP. Let $Q_{t,n}$ denote the data queue size (i.e., the number of bits in the data queue) of node $n$ at the start of frame $t$. At frame $t$, the number of bits transmitted from node $n$ to the AP is $D_{t,n}$ as given in (4). At each frame, new data is generated in each node and is stored in the data queue. Let $A_{t,n}$ denote the number of bits of new data generated in node $n$ at frame $t$. We assume that $A_{t,n}$ is i.i.d. over frames and nodes. We define a data generation rate as the expected number of bits of newly generated data during one frame. Let $a_n$ denote the data generation rate of node $n$ and let $a = (a_1, \ldots, a_N)^T$ denote the data generation rate vector. Then, we have

$$a_n = \mathbb{E}[A_{t,n}],$$

for all $t$. The data queue dynamics is

$$Q_{t+1,n} = [Q_{t,n} - D_{t,n}]^+ + A_{t,n},$$

where $[x]^+ = \max(0, x)$.

### III. Energy Transmission Power Minimization Problem

#### A. Optimization Problem Formulation

The final goal of this paper is to design a stochastic optimal control algorithm that minimizes the energy transmission power (i.e., $\limsup_{t \to \infty} \frac{1}{t} \sum_{i=1}^{t} P_i$), while stabilizing the data queue size (i.e., $Q_{t,n}$) for all nodes, under the assumption that the energy channel vectors (i.e., $g_{t,n}$) at the data channel gain (i.e., $h_{t,n}$) are random processes and the data generation rates (i.e., $a_n$) are unknown. This stochastic optimal control algorithm will be explained in Section IV.

Before designing the stochastic optimal control algorithm, we in this section introduce a simple deterministic optimization problem to provide a useful insight. In this optimization problem, we assume that $g_{t,n}$ and $h_{t,n}$ are deterministic and temporarily remove the frame index $t$ from all notations. The optimization problem is formulated as

$$\begin{align*}
\text{minimize} & \quad \mathbb{E}[P] \\
\text{subject to} & \quad \mathbb{E}[R_n] \geq \mathbb{E}[C_n], \quad \text{for all } n = 1, \ldots, N \quad (11) \\
& \quad \mathbb{E}[D_n] \geq a_n, \quad \text{for all } n = 1, \ldots, N. \quad (12)
\end{align*}$$

In (10), the optimization target is to minimize the energy transmission power at the AP. The constraint (11) enforces that the harvested energy should be no less than the data transmission energy. The data transmission rate is no less than the data generation rate due to the constraint (12).

In (10)–(12), the optimization variables are the beamforming vector (i.e., $w$), the scheduling indicator (i.e., $H_n$), and the data transmission energy (i.e., $C_n$). These optimization variables are all random variables and the expectations in (10)–(12) are taken over these optimization variables. We will derive each term in (10)–(12) by using the optimization variables in Sections III-B and III-C. Then, (10)–(12) will be reformulated in terms of the optimization variables in the optimization problem (30)–(34).

#### B. Energy Beamforming Analysis

In this subsection, we focus on the energy beamforming for the RF energy transmission. The AP controls the energy beam by adjusting the beamforming vector (i.e., $w$). For a given beamforming vector, we can calculate the energy transmission power (i.e., $P$) in (10) and the harvested energy (i.e., $R_n$) in (11) by using (2) and (3).

From (2) and (3), the expected energy transmission power and the expected harvested energy are respectively derived as

$$\begin{align*}
\mathbb{E}[P] &= \mathbb{E}[w^H w] = \text{tr}(\mathbb{E}[ww^H]) = \text{tr} (S),
\end{align*}$$

and

$$\begin{align*}
\mathbb{E}[R_n] &= \mathbb{E}[\eta T(g_n^w w_n^H g_n^H)] \\
&= \mathbb{E}[\eta T(g_n^H g_n^H)] = \eta T(g_n^H S g_n).
\end{align*}$$

where $\text{tr}$ is a trace operator, $S = \mathbb{E}[ww^H]$, and $g_n = g_n^H g_n$. Since $S$ is the expectation of arbitrary random rank-one positive semidefinite matrices, $S$ can be an arbitrary positive semidefinite matrix without any restriction in the rank (i.e., $S \succeq 0$).

We define the energy harvesting region as the set of the vectors of the expected harvested energy of each node, given that the expected energy transmission power is $P$. That is,

$$\Theta(P) = \{r(G) : \text{tr}(S) = P, \ S \succeq 0\},$$

where $G = (G_1, \ldots, G_N)$. In (15), $r(G)$ is the harvested energy function defined as

$$r(G) = (r_1(G), \ldots, r_N(G)),$$

where $r_n(G) = \eta T(r_n G_n)$. Since the linear combination of the positive semidefinite matrices is also a positive semidefinite matrix, it is easy to prove that the energy harvesting region, $\Theta(P)$, is a convex set. Moreover, if the energy transmission power is multiplied by a scalar $\gamma$, the energy harvesting region is expanded $\gamma$ times as well, that is, $\Theta(\gamma P) = \{\gamma x \in \Theta(P)\}$.

We can obtain the Pareto frontier of the energy harvesting region. A point $x$ in $\Theta$ is Pareto-optimal if there is no $y \in \Theta$ such that $y \geq x$ and $y \neq x$. Note that the inequality between two vectors (e.g., $y \geq x$) is a component-wise inequality in this paper. The Pareto frontier is the set of all Pareto-optimal points. Since $\Theta(P)$ is a convex set, we can obtain all Pareto-optimal points in $\Theta(P)$ by solving the following problem with all possible $v = (v_1, \ldots, v_N)^T \in \mathbb{R}^{N \times 1}$ such that $v_n \geq 0$ for all $n$.

$$\begin{align*}
\text{maximize} & \quad v^T r(G, S) = \eta T r(F(G, v) S) \\
\text{subject to} & \quad \text{tr}(S) = P \\
& \quad S \succeq 0, \quad (19)
\end{align*}$$

where $F(G, v)$ is defined as

$$F(G, v) = \sum_{n=1}^{N} v_n G_n.$$
semidefinite matrix, $F(\mathbf{G}, \mathbf{v})$ is also a positive semidefinite matrix. Therefore, the eigenvalue decomposition of $F(\mathbf{G}, \mathbf{v})$ is given as

$$F(\mathbf{G}, \mathbf{v}) = \mathbf{U}(\mathbf{G}, \mathbf{v})\mathbf{Z}(\mathbf{G}, \mathbf{v})\mathbf{U}(\mathbf{G}, \mathbf{v})^H,$$

(21)

where $\mathbf{U}(\mathbf{G}, \mathbf{v})$ is a unitary matrix and $\mathbf{Z}(\mathbf{G}, \mathbf{v})$ is a diagonal matrix. The unitary matrix $\mathbf{U}(\mathbf{G}, \mathbf{v})$ is

$$\mathbf{U}(\mathbf{G}, \mathbf{v}) = (\mathbf{u}_1(\mathbf{G}, \mathbf{v}), \ldots, \mathbf{u}_N(\mathbf{G}, \mathbf{v})),$$

(22)

which consists of the column vectors $\mathbf{u}_n(\mathbf{G}, \mathbf{v}) \in \mathbb{C}^{N \times 1}$. The diagonal matrix $\mathbf{Z}(\mathbf{G}, \mathbf{v})$ is

$$\mathbf{Z}(\mathbf{G}, \mathbf{v}) = \text{diag}(z_1(\mathbf{G}, \mathbf{v}), \ldots, z_N(\mathbf{G}, \mathbf{v})),$$

(23)

the main diagonal of which contains positive eigenvalues $z_n(\mathbf{G}, \mathbf{v}) \geq 0$. Without loss of generality, the main diagonal of $\mathbf{Z}(\mathbf{G}, \mathbf{v})$ is sorted in a decreasing order, that is, $z_j(\mathbf{G}, \mathbf{v}) \geq z_i(\mathbf{G}, \mathbf{v})$ for all $i \leq j$.

The solution to the optimization problem (17)–(19) is given in the following Lemma 1. The proof of Lemma 1 is similar to that of Proposition 2.1 in [26]. Let $\mathbf{S}^*(\mathbf{G}, \mathbf{v}, P)$ denote the solution to this optimization problem.

**Lemma 1:** The optimal solution to the optimization problem (17)–(19) is

$$\mathbf{S}^*(\mathbf{G}, \mathbf{v}, P) = P\mathbf{u}_1(\mathbf{G}, \mathbf{v})\mathbf{u}_1(\mathbf{G}, \mathbf{v})^H.$$  

(24)

**Proof:** See Appendix A.

The optimal solution in Lemma 1 is a rank-one positive semidefinite matrix. Therefore, the optimal solution can be achieved by a single beamforming vector such that $\mathbf{w} = \sqrt{P}\mathbf{u}_1(\mathbf{G}, \mathbf{v})$. Then, we do not need to take expectation over multiple beamforming vectors to achieve the optimal solution.

Let $\mathbf{r}^*((\mathbf{G}, \mathbf{v}, P)) = (\mathbf{r}_1^*((\mathbf{G}, \mathbf{v}, P)), \ldots, \mathbf{r}_N^*((\mathbf{G}, \mathbf{v}, P)))^T$ denote the Pareto-optimal point in $\Phi(P)$ obtained by $\mathbf{v}$. Then, we have

$$\mathbf{r}^*((\mathbf{G}, \mathbf{v}, P)) = \mathbf{r}(\mathbf{G}, \mathbf{S}^*((\mathbf{G}, \mathbf{v}, P))).$$

(25)

By varying $\mathbf{v}$, all Pareto-optimal points of $\Theta(P)$ can be obtained from $\mathbf{r}^*((\mathbf{G}, \mathbf{v}, P))$.

In Fig. 2, we show the Pareto frontier of an example energy harvesting region. For this figure, we assume a uniform linear array with four antenna elements and the antenna spacing is half the wavelength of the energy signal. The energy transmission power is $P = 2000$ mW. There are two nodes which are 3 m away from the AP. Node 1 is located perpendicular to the line of antenna elements and node 2 is located at 45 degrees clockwise rotation from node 1. All other simulation parameters are as given in Section V. The Pareto frontier in Fig. 2 is drawn by solving the optimization problem (17)–(19) with various $\mathbf{r}^*(\mathbf{G}, \mathbf{v}, P)$. In Fig. 2, we also show the harvested energy of nodes 1 and 2 when $\mathbf{v} = (2, 1)^T$, $\mathbf{v} = (1, 1)^T$, and $\mathbf{v} = (1, 2)^T$. In addition, the antenna radiation patterns for these $\mathbf{v}$’s are presented in Fig. 3. In these figures, we can see that the energy beam can be focused to node 1 when $\mathbf{v} = (2, 1)^T$, can be evenly distributed to two nodes when $\mathbf{v} = (1, 1)^T$, or can be focused to node 2 when $\mathbf{v} = (1, 2)^T$.

---

**C. Data Transmission Analysis**

In this subsection, we analyze the data transmission from nodes to the AP. Let $q_n$ denote the scheduling probability of node $n$. The scheduling probability of node $n$ is defined as the probability that node $n$ scheduled, i.e., $q_n = \mathbb{E}[H_n]$. We also define $\mathbf{q} = (q_1, \ldots, q_N)^T$. It should be satisfied that $\mathbf{1}^T\mathbf{q} \leq 1$. A node should use the same data transmission energy when scheduled, in order to maximize the expected data transmission rate given the expected data transmission energy. We define $\rho_n$ as the data transmission energy of node $n$ when node $n$ is scheduled. Then, the expected data transmission energy is given as

$$\mathbb{E}[C_n] = q_n\rho_n,$$

(26)

and the expected data transmission rate is

$$\mathbb{E}[D_n] = q_n\theta(\rho_n; h_n).$$

(27)

We define the data transmission energy region as the set of the vectors of the expected data transmission energy required.
to achieve a given data generation rate. The data transmission energy region is

$$\Xi(a) = \{c(q, \rho)|q_n\theta(\rho_n; h_n) \geq a_n \text{ for all } n = 1, \ldots, N\}$$

where $\rho = (\rho_1, \ldots, \rho_N)^T$ and

$$c(q, \rho) = (q_1\rho_1, \ldots, q_N\rho_N)^T.$$  \hfill (28)

D. Solution to Energy Transmission Power Minimization Problem

From Sections III-B and III-C, the energy transmission power minimization problem in (10)–(12) is reformulated as

minimize $\text{tr}(S)$ \hfill (30)

subject to $\eta T \text{tr}(G_n S) \geq q_n\rho_n$, for all $n$ \hfill (31)

$q_n\theta(\rho_n; h_n) \geq a_n$, for all $n$ \hfill (32)

$S \succeq 0$ \hfill (33)

$I^T q \leq 1$. \hfill (34)

We can geometrically interpret this problem as in Fig. 4. For this figure, we use the same antenna configuration and node locations as those for Figs. 2 and 3. All other simulation parameters are as given in Section V. The optimization problem (30)–(34) is equivalent to finding the minimum $P$ subject to the condition that there exists $x \in \Theta(P)$ and $y \in \Xi(a)$ satisfying $x \succeq y$. In Fig. 4, we present the boundary of the data transmission energy region and the Pareto frontier of the energy harvesting regions when the energy transmission power is 250 mW, 344 mW, and 450 mW. We can see that the energy transmission power of 344 mW is the minimum power that can make the harvested energy no less than the data transmission energy for all nodes.

The optimization problem (30)–(34) can be converted to a convex optimization problem. By merging the constraints (31) and (32), we can make the following new constraint.

$$\eta T \text{tr}(G_n S) \geq \zeta_n(q_n).$$  \hfill (35)

where

$$\zeta_n(x) = x^{\beta - 1}(a_n/x; h_n) = \frac{N_0 WT x}{|h_n|^2}(2^{\frac{a_n}{T h_n}} - 1).$$  \hfill (36)

Since $\zeta"_n(x) \geq 0$ for $x \geq 0$, $\zeta_n(x)$ is a convex function. Therefore, the optimization problem (30), (35), (33), and (34) is a convex optimization problem. This nonlinear convex optimization problem can be solved by means of the interior point method. Let $S^*$ and $q^*$ denote the optimal solutions.

We explain how the AP decides the beamforming vector based on the optimal solution $S^*$. The eigenvalue decomposition of $S^*$ is $S^* = V^* X^* V^* H$, where $V^* = (v_1^*, \ldots, v_N^*)^T$ and $X^* = \text{diag}(\lambda_1^*, \ldots, \lambda_N^*)$. For each eigenvector $v_n^*$, we can decide the probability and the transmission power of $v_n^*$, which are denoted by $a_n^*$ and $\gamma_n^*$, respectively. We can arbitrarily decide $a_n^*$ and $\gamma_n^*$ as long as the following conditions are met: $x_n^* = a_n^* \gamma_n^*$ for all $n = 1, \ldots, N$ and $\sum_{n=1}^N a_n^* \leq 1$. The AP transmits the energy signal by using the beamforming vector $\sqrt{\gamma_n^*} v_n^*$ with the probability $a_n^*$.

IV. ENERGY AND DATA QUEUE SIZE-BASED STOCHASTIC OPTIMAL CONTROL ALGORITHM

A. Stochastic Optimal Control Problem Formulation

The optimal solution to the deterministic optimization problem (30)–(34) can be used for designing the joint energy beamforming and data scheduling algorithm. This optimal solution-based algorithm first calculates the optimal solution, i.e., the beamforming vectors, the scheduling probability, and the data transmission energy, based on the channel gains and the data generation rates. Then, the algorithm decides $w_t$, $H_{t,n}$, and $C_{t,n}$ at each frame $t$ from the optimal solution.

However, this optimal solution-based algorithm is not practical for the following reasons. First, the optimal solution-based algorithm requires the prior knowledge of the data generation rates of all nodes. Second, it is difficult to use the optimal solution-based algorithm in the case that the energy channel vector (i.e., $g_{t,n}$) and the data channel gain (i.e., $h_{t,n}$) are random processes. In this case, the optimization problem (30)–(34) can be reformulated as the optimization problem (40)–(46) in Section IV-B. This optimization problem is not convex and its complexity is too high to be solved online. Moreover, the complete knowledge of the distribution of the channel gains is required to formulate the optimization problem (40)–(46).

To overcome this difficulty, we propose a stochastic optimal control algorithm in this section. The stochastic optimal control algorithm decides $w_t$, $H_{t,n}$, and $C_{t,n}$ at each frame $t$ only based on the current state of the energy queue and the data queue (i.e., $Q_{t,n}$ and $Q_{t,n}$) and the current state of the energy channel vector and the data channel gain (i.e., $g_{t,n}$ and $h_{t,n}$). Let $H_t$, $C_t$, $B_t$, $Q_t$, $g_t$, and $h_t$ denote the vectors of $H_{t,n}$, $C_{t,n}$, $B_{t,n}$, $Q_{t,n}$, $g_{t,n}$, and $h_{t,n}$ for all $n = 1, \ldots, N$, respectively. We consider a stationary policy $\delta$ that maps the state of the energy queue, the data queue, the energy channel vector, and the data channel gain, i.e., $(B_t, Q_t, g_t, h_t)$, to the beamforming vector, the scheduling
indicator, and the data transmission energy, i.e., \((\mathbf{w}, \mathbf{H}, \mathbf{C})\). That is, at frame \(t\), the algorithm calculates

\[
(w_t, H_t, C_t) = \delta(B_t, Q_t, \mathbf{g}_t, \mathbf{h}_t).
\]

(37)

The optimal policy \(\delta^*\) is defined as the policy that minimizes the expected energy transmission power

\[
\limsup_{t \to \infty} \frac{1}{t} \sum_{t=1}^{r} E[P_t].
\]

(38)

while stabilizing the data queue as

\[
\sum_{n=1}^{N} \limsup_{t \to \infty} \frac{1}{t} \sum_{t=1}^{r} E[Q_{t,n}] < \infty.
\]

(39)

In the rest of this section, we will derive such an optimal policy and prove its optimality.

### B. Optimization Problem Formulation of Energy Transmission Power Minimization Problem in Random Channel Gain Case

In this subsection, we reformulate the optimization problem (30)–(34) for the case that the channel gains are random processes. We assume that there are \(M\) discrete channel states. Let \(p^{(m)}\) denote the probability of channel state \(m\). When the channel state is \(m\), the energy channel vector is \(\mathbf{g}_n^{(m)}\) and the data channel gain is \(h_n^{(m)}\). We adopt this discrete channel model rather than a continuous channel model for the ease of the mathematical derivation. The continuous channel state can still be approximated by sufficiently large \(M\). We denote \(h_n, G_n, S, q_n,\) and \(\rho_n\) in channel state \(m\) by \(h_n^{(m)}\), \(G_n^{(m)}\), \(S^{(m)}\), \(q_n^{(m)}\), and \(\rho_n^{(m)}\), respectively.

The reformulated optimization problem is

\[
\text{minimize} \quad \sum_{m=1}^{M} p^{(m)} \text{tr}(S^{(m)})
\]

subject to

\[
\sum_{m=1}^{M} p^{(m)} \eta \text{tr}(G_n^{(m)} S^{(m)}) \\
\geq \sum_{m=1}^{M} p^{(m)} q_n^{(m)} \rho_n^{(m)}, \quad \text{for all } n
\]

\[
\sum_{m=1}^{M} p^{(m)} q_n^{(m)} \theta(\rho_n^{(m)}, h_n^{(m)}) \geq a_n, \quad \text{for all } n
\]

\[
S^{(m)} \geq \mathbf{0}, \quad \text{for all } m
\]

\[
\mathbf{1}^T q^{(m)} \leq 1, \quad \text{for all } m, n
\]

\[
\rho_n^{(m)} \leq \rho_{\text{max}}, \quad \text{for all } m, n
\]

(40)–(46)

In this optimization problem (40)–(46), we limit the data transmission energy \(\rho_n^{(m)}\) to the maximum data transmission energy \(\rho_{\text{max}}\) as in (46). We define the capacity region as the set of all possible data generation rate vectors that can be supported. The capacity region \(\Lambda\) is

\[
\Lambda = \left\{ \mathbf{x} = (x_1, \ldots, x_N)^T | x_n \leq \sum_{m=1}^{M} p^{(m)} q_n^{(m)} \theta(\rho_n^{(m)}, h_n^{(m)}), \right. \\
1^T q^{(m)} \leq 1 \text{ for all } m, \rho_n^{(m)} \leq \rho_{\text{max}} \text{ for all } m, n \right\}
\]

(47)

A data generation rate vector \(\mathbf{a}\) can be supported if and only if the data generation rate vector is inside the capacity region (i.e., \(\mathbf{a} \in \Lambda\)).

The optimal solution to the optimization problem (40)–(46), given the data generation rate vector \(\mathbf{a}\), is denoted by \(S^{(\mathbf{a})}(\mathbf{a}), q_n^{(\mathbf{a})}(\mathbf{a})\), and \(\rho_n^{(\mathbf{a})}(\mathbf{a})\). The optimal value is denoted by \(P^*(\mathbf{a}) = \sum_{m=1}^{M} p^{(m)} \text{tr}(S^{(\mathbf{a})}(\mathbf{a}))\). We respectively define the harvested energy, the data transmission energy, and the data transmission rate of node \(n\) with the optimal solution as

\[
R_n^*(\mathbf{a}) = \sum_{m=1}^{M} p^{(m)} \eta \text{tr}(G_n^{(m)} S^{(\mathbf{a})}(\mathbf{a}))
\]

(48)

\[
C_n^*(\mathbf{a}) = \sum_{m=1}^{M} p^{(m)} q_n^{(\mathbf{a})}(\mathbf{a}) \rho_n^{(\mathbf{a})}(\mathbf{a})
\]

(49)

\[
D_n^*(\mathbf{a}) = \sum_{m=1}^{M} p^{(m)} q_n^{(\mathbf{a})}(\mathbf{a}) \theta(\rho_n^{(\mathbf{a})}(\mathbf{a}), h_n^{(\mathbf{a})}).
\]

(50)

### C. Lyapunov Optimization

We design the stochastic optimal control algorithm by means of the Lyapunov optimization method [21], [22]. Let us define the quadratic Lyapunov function as

\[
L_t = \frac{\mu}{2} \sum_{n=1}^{N} (\phi_n - B_{t,n})^2 + \frac{1}{2} \sum_{n=1}^{N} (Q_{t,n})^2,
\]

(51)

where \(\mu\) is a nonnegative constant and \(\phi_n\) is a perturbation value for the energy queue of node \(n\). We will explain how to set the perturbation value later in this section. We define the perturbation vector as \(\phi = (\phi_1, \ldots, \phi_N)^T\). We assume that the energy queue size does not exceed \(\phi_n\). Therefore, the energy queue dynamics is \(B_{t+1,n} = \min\{B_{t,n} - C_{t,n} + R_{t,n}, \phi_n\}\). One of the goal of the proposed stochastic optimal control algorithm is to stabilize the queues. The Lyapunov function (51) grows larger as the data queue size increases or the energy queue size decreases. Therefore, we can stabilize the queues by stabilizing the Lyapunov function.

The Lyapunov drift is the change of the Lyapunov function between two consecutive frames. The Lyapunov drift is given by

\[
\Delta_t = \mathbb{E}[L_{t+1} - L_t | Z_t],
\]

(52)

where \(Z_t = (B_t, Q_t)\) and the expectation is taken over the channel state (i.e., the energy channel vector and the data channel gain).

The objective of the proposed stochastic optimal control algorithm is to minimize the expected energy transmission power, while stabilizing the queues. To this end, we consider the
Lyapunov optimization that minimizes the drift-plus-penalty function. The drift-plus-penalty function is given by
\[ \Delta_t + \lambda E[P_t | Z_t], \] (53)
where \( \lambda \) is a non-negative constant that balances the trade-off between the expected queue size and the expected energy transmission power.

The following lemma holds for the drift-plus-penalty function.

**Lemma 2:** The drift-plus-penalty function satisfies
\[ \Delta_t + \lambda E[P_t | Z_t] \leq \Upsilon + \lambda E[P_t | Z_t] - \sum_{n=1}^{N} \mu(\phi_n - B_{t,n})E[R_{t,n} | Z_t] + \sum_{n=1}^{N} \mu(\phi_n - B_{t,n})E[C_{t,n} | Z_t] - \sum_{n=1}^{N} Q_{t,n}(E[D_{t,n} | Z_t] - a_n), \] (54)
where \( \Upsilon \) is a constant such that
\[ \Upsilon = (N/2)\left( \mu(R_{\max})^2 + \mu(C_{\max})^2 + (D_{\max})^2 + (A_{\max})^2 \right). \] (55)
In (55), \( R_{\max}, C_{\max}, D_{\max}, \) and \( A_{\max} \) are the maximum values of \( R_{t,n}, C_{t,n}, D_{t,n}, \) and \( A_{t,n} \), respectively.

**Proof:** See Appendix B.

**D. Energy Beamforming and Data Transmission Power Allocation**

In this subsection, we derive the optimal policy \( \delta^* \) that is used by the stochastic optimal control algorithm for deciding the energy beamforming vector, the scheduling indicator, and the data transmission energy at each frame. For minimizing the expected energy transmission power and stabilizing the queues, the stochastic optimal control algorithm minimizes the right side of (54) in Lemma 2. The problem of minimizing the right side of (54) is divided into the following two subproblems. The first subproblem is to decide the beamforming vector \( w_t \) that minimizes
\[ \lambda P_t - \sum_{n=1}^{N} \mu(\phi_n - B_{t,n})R_{t,n}. \] (56)
The second subproblem is to find the scheduling indicator \( H_{t,n} \) and the data transmission energy \( C_{t,n} \) that minimize
\[ \sum_{n=1}^{N} \mu(\phi_n - B_{t,n})C_{t,n} - \sum_{n=1}^{N} Q_{t,n}D_{t,n}. \] (57)
In the second subproblem, there is an additional constraint \( C_{t,n} \leq B_{t,n} \) as in (6). In Lemma 3 in Section IV-E, we will suggest the condition under which \( B_{t,n} \geq \rho_{\max} \) is satisfied for all \( t \). Under this condition, we always have \( C_{t,n} \leq B_{t,n} \). Therefore, we do not consider the constraint \( C_{t,n} \leq B_{t,n} \).

We solve the first subproblem of minimizing (56) based on Lemma 1. Let us define \( S_t = w_t w_t^H, G_t = \Phi_{t,n} \Phi_{t,n}^H, \) and \( \bar{G}_t = (\tilde{G}_{t,1}, \ldots, \tilde{G}_{t,N}) \). Then, we have \( R_{t,n} = r_n(\bar{G}_t, S_t) \) for all \( n \) and \( P_t = \text{tr}(S_t) \). We rewrite (56) as
\[ \lambda \text{tr}(S_t) - \mu(\phi - B_t)^T \textbf{r}(\bar{G}_t, S_t). \] (58)
We can obtain \( S_t \) minimizing (58) by first fixing \( S_t = P \) and then finding \( P \) which minimizes (58).

If we fix \( \text{tr}(S_t) = P \), then \( S_t \) minimizing (58) is the optimal solution of (17)–(19) for \( v = \phi - B_t \). From Lemma 1, \( S_t \) minimizing (58) subject to \( \text{tr}(S_t) = P \) is given by \( S_t = S^*(\bar{G}_t, \phi - B_t, P) \). From this solution, we can rewrite (58) as
\[ P \{ \lambda - \mu(\phi - B_t)^T \textbf{r}^*(\bar{G}_t, \phi - B_t, 1) \}, \] (59)
where \( \textbf{r}^*(\bar{G}_t, \phi - B_t, P) = P \cdot \textbf{r}^*(\bar{G}_t, \phi - B_t, 1) \).

Now, we derive \( P \) that minimizes (59). Let us define \( P_{\max} \) as the maximum energy transmission power. From (59), if \( \lambda - \mu(\phi - B_t)^T \textbf{r}^*(\bar{G}_t, \phi - B_t, 1) \leq 0 \), we can minimize (59) by letting \( P = P_{\max} \). In this case, we have \( S_t = S^*(\bar{G}_t, \phi - B_t, P_{\max}) \) and \( w_t = \sqrt{P_{\max}} u_1(\bar{G}_t, \phi - B_t) \). On the other hand, if \( \lambda - \mu(\phi - B_t)^T \textbf{r}^*(\bar{G}_t, \phi - B_t, 1) > 0 \), we can minimize (59) by letting \( P = 0 \). In this case, we have \( w_t = 0 \). In summary, the stochastic optimal control algorithm decides \( w_t \) at frame \( t \) according to Algorithm 1.

**Algorithm 1.** Calculating the beamforming vector at frame \( t \)
1. Calculate \( F \leftarrow \sum_{n=1}^{N} (\phi_n - B_{t,n})G_{t,n} \).
2. Derive the eigenvalue decomposition of \( F = \textbf{UZU}^H \), where \( \textbf{U} = (u_1, \ldots, u_N) \) and \( \textbf{Z} = \text{diag}(z_1, \ldots, z_N) \) satisfying \( z_i \geq z_j \) for all \( i \leq j \).
3. If \( \lambda - \sum_{n=1}^{N} \mu(\phi_n - B_{t,n}) \eta \text{tr}(G_{t,n} u_i u_i^H) \leq 0 \) then
   4. \( w_t \leftarrow \sqrt{P_{\max}} u_i \).
5. Else
   6. \( w_t \leftarrow 0 \).
7. End if
8. Return \( w_t \).

The second subproblem of minimizing (57) can be solved as follows. Since \( D_{t,n} = \theta(C_{t,n}; h_{t,n}) \), we can rewrite the second subproblem as
\[ \text{minimize} \sum_{n=1}^{N} f_n(C_{t,n}; h_{t,n}, B_{t,n}, Q_{t,n}) \] (60)
subject to \( 0 \leq C_{t,n} \leq \rho_{\max} H_{t,n} \) for all \( n \) (61)
\[ \sum_{n=1}^{N} H_{t,n} \leq 1, \] (62)
where
\[ f_n(x; h, B, Q) = \mu(\phi_n - B) x - Q(x; h) \]
\[ = \mu(\phi_n - B) x - Q W_T \log_2 \left( 1 + \frac{|h|^2 x}{N_0 W T} \right), \] (63)
from (5).

To solve the optimization problem (60)–(62), we first solve the following problem for each \( n \) to decide the data transmission energy when node \( n \) is scheduled.
\[ \text{minimize} \quad f_n(x; h_{t,n}, B_{t,n}, Q_{t,n}) \] (64)
subject to \( 0 \leq x \leq \rho_{\max} \). (65)
By differentiating the objective function (64), we can derive the optimal solution of the optimization problem (64), (65) as
\[ \lambda^*_n = \min \{ |y^*_n|^2, \rho_{\max} \}. \] (66)
where \( y_{i,n}^* \) is given by
\[
y_{i,n}^* = \frac{Q_{i,n} \mu}{\mu (\phi_n - B_{i,n}) \ln 2} - \frac{N_n \mu}{|h_{i,n}|^2}. \tag{67}
\]

Now, we select the node that minimizes the objective function (60) for scheduling at frame \( t \). Let us define \( n^* \) by
\[
n^* = \arg \min_{n=1,\ldots,N} f_n(x_{n}^*, h_{i,n}, B_{i,n}, Q_{i,n}). \tag{68}
\]

If \( f_n(x_{n}^*, h_{i,n}, B_{i,n}, Q_{i,n}) < 0 \), the stochastic optimal control algorithm decides \( n^* \) for frame \( t \). Therefore, \( H_{t,n}^* = 1 \) and \( C_{t,n}^* = x_{n}^* \) for the scheduled node \( n^* \), and \( H_{t,n} = 0 \) and \( C_{t,n} = 0 \) for all \( n \neq n^* \). On the other hand, if \( f_n(x_{n}^*, h_{i,n}, B_{i,n}, Q_{i,n}) = 0 \), no node is scheduled at frame \( t \), i.e., \( H_{t,n} = 0 \) and \( C_{t,n} = 0 \) for all \( n = 1, \ldots, N \). In summary, the stochastic optimal control algorithm decides \( H_t \) and \( C_t \) at frame \( t \) according to Algorithm 2.

**Algorithm 2.** Calculating the scheduling indicator and the data transmission energy at frame \( t \)

1: Calculate \( x_n = \min \left\{ \frac{Q_{i,n} \mu}{\mu (\phi_n - B_{i,n}) \ln 2} - \frac{N_n \mu}{|h_{i,n}|^2}, \rho_{\text{max}} \right\} \)
   for all \( n = 1, \ldots, N \).
2: Calculate \( f_n = \mu (\phi_n - B_{i,n}) x_n - Q_{i,n} \mu \ln 2 (1 + \frac{|h_{i,n}|^2 x_n}{N_n \mu}) \)
   for all \( n = 1, \ldots, N \).
3: Find \( n^* = \arg \min_{n=1,\ldots,N} f_n \).
4: if \( f_{n^*} < 0 \) then
5: \( H_{t,n^*} = 1 \) and \( H_{t,n} = 0 \) for all \( n \neq n^* \).
6: \( C_{t,n^*} = x_{n^*} \) and \( C_{t,n} = 0 \) for all \( n \neq n^* \).
7: else
8: \( H_{t,n} = 0 \) for all \( n = 1, \ldots, N \).
9: \( C_{t,n} = 0 \) for all \( n = 1, \ldots, N \).
10: end if
11: return \( H_t \) and \( C_t \).

**E. Performance Analysis**

In this section, we analyze the performance of the stochastic optimal control algorithm, i.e., Algorithms 1 and 2. This performance analysis is about the stability of the energy and data queues and the optimality of the energy transmission power.

We first present the following Lemma 3 about the energy queue size.

**Lemma 3:** If the beamforming vector is decided according to Algorithm 1, the energy queue size \( B_{t,n} \) satisfies that \( B_{t,n} \geq \rho_{\text{max}} \) for all \( t \) if the following conditions on the perturbation value \( \phi_n \) and the maximum energy transmission power \( P_{\text{max}} \) are met.

\[
\phi_n \geq \frac{\lambda}{\mu \min_m r_m^2 (\mathbf{G}^{(m)}, \psi_n, 1)} + 2 \rho_{\text{max}} \tag{69}
\]

\[
P_{\text{max}} \geq \frac{\rho_{\text{max}}}{\min_{m,n} \min_{\kappa_n} r_m^2 (\mathbf{G}^{(m)}, \kappa_n, 1)}. \tag{70}
\]

where \( m \) is the index of a channel state and \( \mathbf{G}^{(m)} = (\mathbf{G}_1^{(m)}, \ldots, \mathbf{G}_N^{(m)}) \). In (69), \( \psi_n = (\psi_1, \ldots, \psi_N)^T \), where \( \psi_n = 1 \) and \( \psi_i = 0 \) for all \( i \neq n \). In (70), the minimization over \( \kappa_n \) is done over \( \kappa_n = (\kappa_1, \ldots, \kappa_N)^T \) such that \( \kappa_n = \phi_n - 2 \rho_{\text{max}} \) and \( 0 \leq \kappa_i \leq \phi_n - \rho_{\text{max}} \) for \( i \neq n \).

**Proof:** See Appendix C.

We choose \( \phi_n \) and \( P_{\text{max}} \) that are large enough to satisfy the conditions (69) and (70). Then, from Lemma 3, it is satisfied that \( B_{t,n} \geq \rho_{\text{max}} \) for all \( t \) and \( n \). Since \( B_{t,n} \geq \rho_{\text{max}} \) and \( C_{t,n} \leq \rho_{\text{max}} \), the constraint \( C_{t,n} \leq B_{t,n} \) in (6) is always satisfied.

**Lemma 4:** If Algorithms 1 and 2 are used and the data generation rate vector \( a \) is in the strict interior of the capacity region \( \Lambda \), the following inequality holds for some \( \epsilon > 0 \).

\[
\mathbb{E} \left[ \Delta_1 \right] + \lambda \mathbb{E} [P_t] \leq \Upsilon + \lambda P^* (a + \epsilon) - \epsilon \sum_{n=1}^N \mathbb{E} [Q_{t,n}], \tag{71}
\]

where \( \epsilon \) is the column vector the components of which are all \( \epsilon \).

**Proof:** See Appendix D.

Finally, we present the following theorem that shows the optimality of Algorithms 1 and 2.

**Theorem 1:** If Algorithms 1 and 2 are used and \( a \) is in the strict interior of \( \Lambda \), the expected data queue size and the expected energy transmission power satisfies

\[
\sum_{n=1}^N \limsup_{r \to \infty} \frac{1}{r} \sum_{t=1}^r \mathbb{E} [Q_{t,n}] \leq \frac{\mathbb{E} [\Delta_1] + \lambda \mathbb{E} [P_t] - \mathbb{E} [Q_{t,n}]}{\epsilon} \tag{72}
\]

\[
\limsup_{r \to \infty} \frac{1}{r} \sum_{t=1}^r \mathbb{E} [P_t] \leq \frac{\Upsilon + \lambda P^* (a + \epsilon)}{\lambda}. \tag{73}
\]

**Proof:** See Appendix E.

The inequality (72) means that the data queue size is stable for all nodes. In (73), we can see that the expected energy transmission power can be arbitrarily close to the optimal energy transmission power \( P^* (a) \) as we increase \( \lambda \) and choose a small \( \epsilon \).

**V. Numerical Result**

In this section, we present numerical results to show the performance of the proposed algorithm. The simulation parameters are as follows. The frame duration is \( T = 1 \) ms. We consider a polar coordinate system in which the AP is located at the center. The location of node \( n \) is given by the distance \( d_n \) and the angle \( \chi_n \). We assume a simple pathloss model of \( 1/jd_n - \alpha \), where \( d \) is a distance (in meter) between a transmitter and a receiver, the reference distance is \( d_0 = 1 \) m, the transmit and receive antenna gains and the carrier frequency are considered in the constant \( \Psi = 10^{-3} \), and a pathloss exponent is \( \alpha = 3 \). The number of antennas for the energy transmission is four (i.e., \( K = 4 \)). The maximum energy transmission power is \( P_{\text{max}} = 2000 \) mW. The energy harvesting efficiency is \( \eta = 0.5 \). The bandwidth for the data transmission is \( W = 1 \) MHz. The noise spectral density is \( N_0 = -124 \) dBm/Hz. The maximum data transmission energy is \( \rho_{\text{max}} = 0.2 \) mJ. The perturbation value for the energy queue is \( \phi_n = 1 \) mJ for all \( n \). Unless noted otherwise, we use the constants \( \lambda = 10^7 \) and \( \mu = 10^{15} \) when \( P_t \) is in mW, \( R_{t,n} \), \( C_{t,n} \), and \( B_{t,n} \) are in mJ, and \( D_{t,n} \) and \( Q_{t,n} \) are in bit.
We consider two channel models: the deterministic and the random channel models. For the deterministic channel model, we consider a uniform linear array and the antenna spacing is half the wavelength of the energy signal. We also assume that there is no multipath fading in both the energy and the data channels. On the other hand, in the random channel model, each component of the energy channel vector and the data channel gain of each node are subject to the independent Rayleigh fading. We assume that the number of channel states is $M = 100$ and the probabilities of channel states are all equal, i.e., $p^{(m)} = 0.01$ for all $m$. The energy channel vector and the data channel gain for each channel state are randomly generated according to the Rayleigh fading.

In Fig. 5, we present the time evolution of the energy transmission power, the data transmission and harvested energy, the energy queue size, the data transmission and data generation rate, and the data queue size.

![Fig. 5](image)

**Fig. 5.** The time evolution of the energy transmission power, the data transmission and harvested energy, the energy queue size, the data transmission and data generation rate, and the data queue size.

In Figs. 6 and 7, we present the energy transmission power and the energy and the data queue sizes of the proposed algorithm in the deterministic channel model. There are three nodes in this simulation. These figures are plotted according to the data generation rates which are the same for all nodes. We consider the following three scenarios. In scenario 1, we set $d_1 = 3m$, $d_2 = 3m$, $d_3 = 3m$, $\chi_1 = 0^\circ$, $\chi_2 = 45^\circ$, and $\chi_3 = 90^\circ$. In scenario 2, we set $d_1 = 2m$, $d_2 = 3m$, $d_3 = 4m$, $\chi_1 = 0^\circ$, $\chi_2 = 45^\circ$, and $\chi_3 = 90^\circ$. In scenario 3, we set $d_1 = 2m$, $d_2 = 3m$, $d_3 = 4m$, $\chi_1 = 0^\circ$, $\chi_2 = 11.25^\circ$, and $\chi_3 = 22.5^\circ$. Fig. 6 compares the energy transmission power of the proposed algorithm with the optimal solution of (30)–(34). We can see that the proposed algorithm achieves the optimality in terms of the energy transmission power.

![Fig. 6](image)

**Fig. 6.** The comparison between the optimal energy transmission power and the energy transmission power of the proposed algorithm in the deterministic channel model.

In Fig. 7, we see that the energy queue size and the data queue size are stabilized.

![Fig. 7](image)

**Fig. 7.** The energy queue size and the data queue size of the proposed algorithm in the deterministic channel model.
the energy and the data queue sizes of the proposed algorithm are stable.

In Fig. 8, we show the energy transmission power and the data queue size of the proposed algorithm according to $\lambda$. In this figure, we assume the random channel model. There are ten nodes whose distances from the AP are the same. The data generation rate of node $n$ is set to $100 \times n$ bit/frame. The optimal energy transmission power in this figure is the dual optimal value of the optimization problem (40)–(46). Note that the dual optimal value is no higher than the optimal value of the primal problem. Although it is difficult to solve (40)–(46), we can obtain the dual optimal value by using the subgradient method. We can see that the energy transmission power of the proposed algorithm converges to the dual optimal value as $\lambda$ increases. This means that the duality gap of (40)–(46) is zero, and the proposed algorithm achieves the optimality if $\lambda$ is sufficiently high, as claimed in Theorem 1.

In Fig. 8, we can see that, as $\lambda$ increases, the data queue size increases while the energy transmission power converges to the optimal value, as stated in Theorem 1. Since an arbitrarily large data queue size is permitted as long as the data queue size is stable, we can increase $\lambda$ until the energy transmission power becomes close enough to the optimal value. However, in real communication systems, the maximum data queue size is limited because of a limited memory space, and we have to set $\lambda$ to a moderate value which can best balance the tradeoff between the energy transmission power and the data queue size.

Figs. 9 and 10 present the energy transmission power and the energy and data queue sizes of the proposed algorithm in the random channel model. There are ten nodes with equal distances from the AP. We consider the following three scenarios according to the data generation rates. The data generation rate of node $n$ is set to $50 \times n$ bit/frame in scenario 1, is set to $75 \times n$ bit/frame in scenario 2, and is set to $100 \times n$ bit/frame in scenario 3. In Fig. 9, we can see that the energy transmission power of the proposed algorithm is equal to the dual optimal value of (40)–(46). Therefore, in this figure, the optimality of the proposed algorithm is shown in the random channel model. We have proved that the proposed algorithm minimizes the energy transmission power while stabilizing the data queue size. By simulation, we have shown that the proposed algorithm achieves the optimality in both the deterministic and the random channel models.

In future work, robust beamforming techniques [27], [28] can be applied to the proposed algorithm for the case of imperfect channel estimation. For applying the robust beamforming, we can incorporate an error vector in the energy channel vector, reformulate the optimization problem (17)–(19), and modify Algorithm 1 accordingly.

VI. CONCLUSION

In this paper, we have proposed the stochastic optimal control algorithm for the WPCN with the energy beamforming. The proposed algorithm is designed based on the Lyapunov optimization technique. The proposed algorithm is an online algorithm that adaptively steers the energy beam to nodes with lower energy in the energy queue and schedules nodes with larger data in the data queue for data transmission. The proposed algorithm opportunistically exploits the time-varying channel state so that the optimality is achieved in the random channel model. We have proved that the proposed algorithm minimizes the energy transmission power while stabilizing the data queue size. By simulation, we have shown that the proposed algorithm achieves the optimality in both the deterministic and the random channel models.

In future work, robust beamforming techniques [27], [28] can be applied to the proposed algorithm for the case of imperfect channel estimation. For applying the robust beamforming, we can incorporate an error vector in the energy channel vector, reformulate the optimization problem (17)–(19), and modify Algorithm 1 accordingly.
**APPENDIX A
PROOF OF LEMMA 1**

In this proof, for brevity, we omit \( \overline{G} \) and \( v \) from the notations \( F(G, v), U(G, v), U_t(G, v), Z(G, v), \) and \( z_n(G, v) \). Since \( S \) is a positive semidefinite matrix, the eigenvalue decomposition of \( S \) is

\[
S = V \Sigma V^H,
\]

(74)

where \( V \) is a unitary matrix and \( \Sigma \) is a diagonal matrix such that \( \Sigma = \text{diag}(\lambda_1, \ldots, \lambda_N) \). From (21) and (74), we have

\[
\text{tr}(FS) = \text{tr}(UZU^HVXX^H) = \text{tr}(V^HUX).
\]

(75)

where \( Y = (y_{ij}^*)_{i,j} = V^H U \).

We define a doubly stochastic matrix \( \Phi = (\phi_{ij})_{i,j} \), each element of which is \( \phi_{ij} = |y_{ij}|^2 \). Then, from (75), we can rewrite \( \text{tr}(FS) \) as

\[
\text{tr}(FS) = x^T \Phi z,
\]

(76)

where \( x = (x_1, \ldots, x_N)^T \) and \( z = (z_1, \ldots, z_N)^T \) are the vectors of main diagonal elements of \( X \) and \( Z \), respectively. In addition, we have

\[
\text{tr}(S) = 1^T x.
\]

(77)

Now, the optimization problem (17)–(19) becomes to maximize \( x^T \Phi z \) subject to \( 1^T x = P \) if we omit \( \eta T \) from the objective function. The objective function is maximized when \( V = U, x_1 = P, \) and \( x_n = 0 \) for \( n \neq 1 \). In this case, the value of the objective function is \( P z_1 \). We can prove that \( x^T \Phi z \leq P z_1 \) for \( x \) such that \( 1^T x = P \) as follows. All elements in \( \Phi \) are not greater than \( z_1 \) since \( z_1 \geq z_n \) for all \( n \) and \( \Phi \) is a doubly stochastic matrix. Since \( 1^T x = P \), the maximum value of \( x^T \Phi z \) is \( P z_1 \).

If \( V = U, x_1 = P, \) and \( x_n = 0 \) for \( n \neq 1 \), we have \( S = P u_1 u_1^H \), which proves Lemma 1.

**APPENDIX B
PROOF OF LEMMA 2**

From (7), (9), (51), and (52), we have

\[
\Delta_t = (1/2)E[\mu \sum_{n=1}^N (\phi_n - B_t - C_t, n + R_t, n, \phi_n)]^2
- (\phi_n - B_t, n)^2] + \sum_{n=1}^N ((Q_t, n - D_t, n)^T + A_t, n)^2

- (Q_t, n)^2 \geq 0.
\]

(78)

Adding \( \lambda E[P_t|Z_t] \) to the both side of (78), we obtain the inequality (54).

**APPENDIX C
PROOF OF LEMMA 3**

We prove Lemma 3 by proving that \( B_t + i \geq \rho_{\max} \) for all \( i \) if \( B_t + i \geq \rho_{\max} \) for all \( i \), under the conditions (69) and (70). Suppose that \( B_t + i \geq \rho_{\max} \) for all \( i \) and let us focus on a reference node \( n \). If \( B_{t,n} \geq \rho_{\max} \), we have \( B_{t+1,n} \geq \rho_{\max} \) since \( C_{t,n} \leq \rho_{\max} \).

We consider the case that \( \rho_{\max} \leq B_{t,n} \leq 2 \rho_{\max} \). In this case, we have \( \phi_n - B_{t,n} \geq \lambda/(\mu \min_n r_n^*(G_{t,n}^m, \psi_n, 1)) \) from (69). Therefore, we have

\[
\lambda - \mu(\phi - B_t)^T r_n^*(G_{t,n}, \phi - B_t, 1)
\]

\[
\leq \lambda - \mu(\phi_n - B_{t,n}) r_n^*(G_{t,n}, \psi_n, 1)
\]

\[
\leq \lambda (1 - r_n^*(G_{t,n}, \psi_n, 1)) \min_n r_n^*(G_{t,n}^m, \psi_n, 1) \leq 0.
\]

(79)

Then, Algorithm 1 sets \( P_t = \rho_{\max} \). From (70), the harvested energy by node \( n \) satisfies

\[
R_{t,n} = \rho_{\max} r_n^*(G_{t,n}, \phi - B_t, 1)
\]

\[
\geq \rho_{\max} \min_n r_n^*(G_{t,n}^m, \kappa_n, 1) \geq \rho_{\max}.
\]

(80)

Finally, we can conclude that

\[
B_{t+1,n} = B_{t,n} - C_{t,n} + R_{t,n} \geq B_{t,n} \geq \rho_{\max}.
\]

(81)

**APPENDIX D
PROOF OF LEMMA 4**

If \( a \) is in the strict interior of \( \Lambda \), there exists \( \epsilon > 0 \) such that \( a + \epsilon \in \Lambda \). When the data generation rate vector is \( a + \epsilon \), the optimal value of the optimization problem (40)–(46) is denoted by \( P^+(a + \epsilon) \). When this optimal solution is applied, the expected harvested energy, the expected data transmission energy, and the expected data transmission rate of node \( n \) are respectively \( R_n^*(a + \epsilon) \), \( C_n^*(a + \epsilon) \), and \( D_n^*(a + \epsilon) \) from (48)–(50). From the constraints (42) and (43), we have \( R_n^*(a + \epsilon) \geq C_n^*(a + \epsilon) \) and \( D_n^*(a + \epsilon) \geq a_n + \epsilon \) for all \( n \). Since Algorithms 1 and 2 minimize the right side of (54), we have

\[
\Delta_t = \lambda E[P_t|Z_t]
\]

\[
\leq \min \left\{ \gamma + \lambda E[P_t|Z_t]
- \sum_{n=1}^N \mu(\phi_n - B_{t,n})(E[R_{t,n}|Z_t] - E[C_{t,n}|Z_t])
- \sum_{n=1}^N Q_{t,n}(E[D_{t,n}|Z_t] - a_n)\right\}
\]

\[
\leq \gamma + \lambda P^+(a + \epsilon)
- \sum_{n=1}^N \mu(\phi_n - B_{t,n})(R_{n}^*(a + \epsilon) - C_{n}^*(a + \epsilon))
- \sum_{n=1}^N Q_{t,n}(D_{n}^*(a + \epsilon) - a_n)
\]

\[
\leq \gamma + \lambda P^+(a + \epsilon) - \sum_{n=1}^N Q_{t,n} \epsilon.
\]

(82)

We derive (71) after taking expectation over \( Z_t \) in (82).
APPENDIX E

Proof of Theorem 1

By taking \( \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} \) on the both side of (71), we have

\[
\limsup_{\tau \to \infty} \frac{1}{\tau} \left( E[L_{t+1}] - E[L_1] \right) + \lambda \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} E[P_t] 
\leq \Upsilon + \lambda P^*(a + \epsilon) - \epsilon \sum_{n=1}^{N} \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} E[Q_{n,t}] \tag{83}
\]

Then, we can obtain

\[
\lambda \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} E[P_t] + \epsilon \sum_{n=1}^{N} \limsup_{\tau \to \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} E[Q_{n,t}] 
\leq \Upsilon + \lambda P^*(a + \epsilon). \tag{84}
\]

From (84), we can derive both (72) and (73).

REFERENCES