SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2016, Vol. 59 102310:1–102310:16
doi: 10.1007/s11432-015-5481-9

Cross-layer transmission and energy scheduling under full-duplex energy harvesting wireless OFDM joint transmission

Hongjia LI[1*], Zejue WANG[1], Dan HU[2], Song CI[3] & Zhen XU[1]

[1]*State Key Lab of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China;*
[2]*Cisco System Incorporated, Beijing 100022, China;*
[3]*Department of Computer and Electronics Engineering, University of Nebraska-Lincoln, NE 68182, USA*

Received August 30, 2015; accepted October 3, 2015; published online April 26, 2016

Abstract This paper studies the design of the optimal and online cross-layer transmission and energy schedulings for a full-duplex energy harvesting wireless orthogonal frequency division multiplexing (OFDM) joint transmissions. Supported by today's power management integrated circuit, the full-duplex energy harvesting system becomes a reality, which can overcome the transmission time loss problem caused by the half-duplex constraint of the energy storage unit (ESU) in the serial Harvest-Store-Use system. However, its corresponding modeling is still unexplored. Therefore, the full-duplex energy harvesting system is first modeled and proved to be equivalent to a composition of energy behavior models of Harvest-Store-Use in fine-time granularity. Then, the convex optimization problem of cross-layer transmission and energy scheduling is formulated with the objective to maximize the sum of transmission throughput during successively multiple time units, which takes into account the temporal variance of energy harvesting rates and channel states, and the limited capacity of ESUs. The optimal power allocation with three dimensions of time, channel and antenna is solved by utilizing the dual decomposition method with the pre-known temporal variance, and the corresponding result of the system throughput provides the theoretical upper bound. Finally, to reduce the throughput degradation caused by channel state prediction errors, a non-convex online scheduling problem is formulated as the classical energy efficiency format. It is transformed into a convex optimization problem by exploiting the properties of fractional programming, and then, an efficiently iterative solution is designed. Numerical results show that the average throughput of the online algorithm is 24% greater than that of existing time-energy adaptive water-filling algorithm. The degradation of the average throughput is less than 19% with probability 90%, even as the channel prediction error reaches 20%. These results provide guidelines for the design and optimization for full-duplex energy harvesting joint transmission systems.

Keywords energy harvesting, wireless communication, power allocation, joint transmission, optimization, green communication

Citation Li H J, Wang Z J, Hu D, et al. Cross-layer transmission and energy scheduling under full-duplex energy harvesting wireless OFDM joint transmission. Sci China Inf Sci, 2016, 59(10): 102310, doi: 10.1007/s11432-015-5481-9

* Corresponding author (email: lihongjia@iie.ac.cn)

1 Introduction

Energy harvesting is gaining popularity due to its ability to capture and store energy from readily available ambient sources, including wind, solar, biomass, geotherm, tides, and even radio frequency signals, which are renewable and more environmentally friendly than that derived from fossil fuels [1], and is thus becoming a preferred choice supporting green wireless communications in the fields of wireless sensor networks (WSNs) [2, 3], WiFi networks [4], cellular networks [5, 6], etc.

Energy harvesting system often in the relevant literature contains two major architectures: energy generators directly connected to the load without energy storage units (ESU's) and energy generators bridged by ESU's to the load [7], the corresponding energy behavior models of which are Harvest-Use and Harvest-Store-Use, respectively. In Harvest-Use, the energy generator directly powers the load and as sufficient energy is not available the load is disabled. In Harvest-Store-Use, there is an ESU that stores the harvested energy and then powers the load. Because the energy storing capability of ESU enables temporal scheduling of harvested energy, Harvest-Store-Use is preferable in prior arts. However, the half-duplex constraint of ESU that charge and discharge cannot be executed simultaneously leads to transmission time loss problem, i.e., data transmissions have to be suspended during ESU is charged. Fortunately, combining these two architectures controlled by the power controller makes full-duplex energy harvesting enabled transmissions possible [8]. However, the fundamental problem of how to model the full-duplex energy harvesting system is still left open, which is named as motivation problem 1.

Due to the fact that the energy-arrival rate is determined by the changing surrounding environment, it is not trivial to design and optimize the energy harvesting enabled wireless communication system. Since the energy cannot be consumed before it is harvested, the opportunistic energy harvesting results in fluctuating power budget, namely, energy causality constraint. The energy causality constraint mandates that, at any time, the total consumed energy should be equal to or less than the sum of the total harvested energy and initial energy in ESU [2]. Furthermore, energy overflow constraint should also be considered to avoid the occurrence that the sum of unconsumed energy and newly arriving energy exceeds the finite capacity of ESU at any time [2].

In addition, to maximize the sum of transmission throughput and corresponding harvested energy utilization efficiency, it is essential to take the cross-layer radio resources scheduling according to profiles of the energy generation and channel states. Flashing back to the relevant literature, the first page appears energy harvesting enabled WSN, since sensor nodes are often deployed under complicated and adverse conditions where manual battery recharging is infeasible. In [2], an offline transmission policy is designed to improve the amount of data transmitted during a finite time period. As for the broadcasting scenario, a packet transmission scheduling scheme is studied under the additional white Gaussian noise (AWGN) broadcast channels in [3]. The second page appears energy harvesting enabled WiFi ad hoc communication system. For example, using historical solar isolation and traffic flows data, Ref. [4] achieves the max-min fair flow control subject to eliminating WiFi network outage. Recently, the harvesting technique is extensively studied in cellular networks. For example, an energy efficient resource allocation algorithm is proposed for an OFDMA downlink network in [5]. In [6], the throughput maximization problem is investigated for harvested energy powered relay nodes under Gaussian channel.

In order to improve throughput, and spectrum and energy efficiency, advanced transmission techniques are adopted for wireless transmissions and correspondingly standardized for extensive use. One of such advanced transmission techniques is the coordinated multipoint (CoMP) transmission, which is accepted in LTE-(A) specification and is becoming a promising technique for 5G networks [9,10]. There are various types of CoMP schemes as overviewed in [9], and one typical CoMP scheme, known as joint transmission, is to coordinate multiple transmission points (TP's) to jointly transmit to a receiver. To green the joint transmission system, Ref. [11] investigates the sum-rate maximization of the joint transmission system powered by hybrid energy without ESU, by jointly optimizing the transmit power allocations at cooperative TP's and their exchanged energy amounts. Ref. [12] presents a renewable energy aware cluster formation scheme to minimize the energy consumption in electric grid, with the support of hybrid energy supply in each TP. However, to our best knowledge, how to optimize the sum throughput of full-duplex energy harvesting joint transmissions in successively multiple time units (or transmission intervals (TI's) in this paper), jointly considering the temporal variance of energy harvesting rates and channel states, and the limited capacity of ESUs, is relatively unexplored, which is named as motivation problem 2.

What is more, conversion from optimal transmission and energy scheduling to online one is the basis of practical implementation. In the online scheduling, relative long-term prediction of channel state information (CSI) [13] is necessary, which inevitably draws into prediction inaccuracy and errors. Because the imperfect CSI will degrade the performance, especially for complex and CSI-dependent online operations, robust design for the online algorithm is of significant importance. Ref. [14] well supports this point and pioneeringly investigates the robust design of MIMO beamforming with partial CSI under the radio signal energy harvesting scenario. In this paper, the motivation problem 3 for the online operation is raised: How can the throughput degradation caused by channel prediction errors be mitigated in the online scheduling algorithm?

Motivated by above 3 motivation problems, the following contributions are made. We first model the energy behavior model of full-duplex energy harvesting system, and then prove the model can be equivalent to Harvest-Store-Use model as the time unit or TI is small enough. Then, the sum of transmission throughput optimization problem with multiple TI's is formulated, the optimization variables of which are power allocations with 3 dimensions of TI, antenna and channel. Proving the convexity of the formulated problem, we derive solution (named as optimal cross-layer transmission and energy scheduling) utilizing the dual decomposition method under the assumption of perfectly pre-known CSI, which provides the upper bound of the system performance. Finally, based on the intuition from the original optimization problem, the online algorithm problem is formulated, taking into account reducing the impact of CSI prediction errors, which is a non-convex optimization problem. The non-convex optimization problem is transformed into a convex optimization problem by exploiting the properties of fractional programming, on the basis of which an efficient iterative solution (named as online cross-layer transmission and energy scheduling) is proposed. Our solutions and results can be useful in the design and transmission optimization in areas of the energy harvesting enabled general wireless communication systems, and private wireless communication networks, such as power grid and oil pipeline monitoring communication systems, emergency wireless communication systems for disaster relief, satellites communication, etc.

The rest of this paper is organized as follows. Section 2 presents the system model. Section 3 illustrates the problem formulation and the solution methodology and algorithm of the optimal cross-layer transmission and energy scheduling. Section 4 provides the formulation and the solution methodology of the online cross-layer transmission and energy scheduling. Section 5 evaluates and analyzes the performance of our proposed algorithms. Finally, Section 6 concludes the paper.

2 System model

Consider the energy harvesting enabled wireless OFDM joint transmission system as shown in Figure 1. There are one central scheduling and signal processing unit (CSPU), $L \in \mathbb{N}$ remote radio units (RRU's) with single antenna connected to CSPU. The receiver-specific cluster formulation is considered, where the receivers are selected to be jointly served by joint transmission of L RRU's [15]. Each RRU is equipped with and powered by the energy harvester (e.g., the wind turbine, the electromagnetic radiation energy harvester, photovoltaic panels, etc.) and ESU (e.g., batteries or super capacitance) with finite capacity E_{C} Joule[1)]. The CSPU is powered by relatively stable power[2)], and gathers data to be transmitted from other network nodes. Time is slotted, and we divide the time into transmission frames (TF's) with T ms consisting of M time intervals (TI's). The size of TI with the integer index $i \in \mathbb{N}$ is τ ms. The cross-layer

1) The RRU's can also be powered by hybrid energy. Hereafter, because we focus on the harvested energy utilization, only the harvesting systems are associated with RRU's in Figure 1.

2) This model considers the practical constraint of the high power consumption and stable power requirement of CSPU, including baseband processing equipments, cooling facilities, etc.

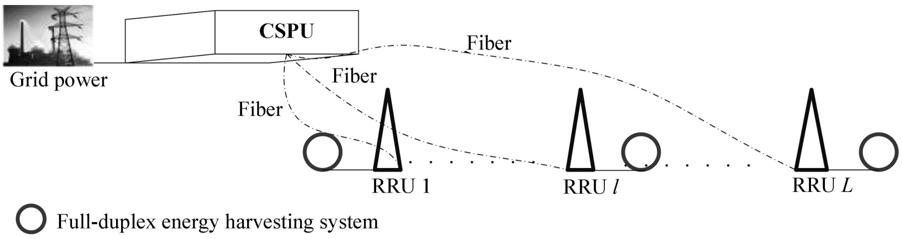

Figure 1 Full-duplex energy harvesting enabled joint transmission system.

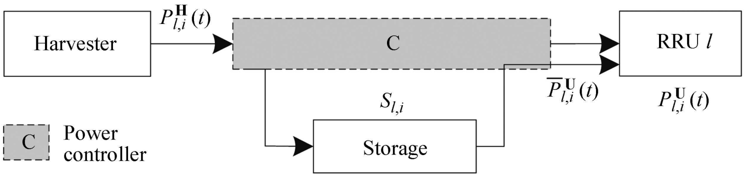

Figure 2 Full-duplex energy harvesting system.

transmission and energy scheduling is periodically made at the beginning of each TF. Therefore, for clear presentation, the following paper is discussed within one reference cycle, i.e., one TF.

2.1 Full-duplex energy harvesting model

The full-duplex energy harvesting enabled RRU is described in Figure 2, where harvesting, storage and usage are controlled by the power controller. $P^{\mathrm{H}}_{l,i}(t)$ is the output power of the harvester associated with RRU l, and $\bar{P}^{\mathrm{U}}_{l,i}(t)$ is the maximal output power of the ESU at time t in TI i, where $l \in \mathbb{N}$ is the integer index of RRU's. The power load of RRU l in TI i is $P^{\mathrm{U}}_{l,i}(t)$. The residual energy of ESU associated with RRU l at the beginning of TI i is $S_{l,i}$. Then, its basic energy behavior can be modeled, as the TI, i.e., τ, is small enough that the relationship among $P^{\mathrm{H}}_{l,i}(t)$, $P^{\mathrm{U}}_{l,i}(t)$ and $\bar{P}^{\mathrm{U}}_{l,i}(t)$ is fixed as the following 3 cases.

Case 1. If $P^{\mathrm{H}}_{l,i}(t) \geqslant P^{\mathrm{U}}_{l,i}(t)$ during τ, the output power of the energy harvester is enough to drive RRU l, and the surplus of harvested energy is stored into ESU with the capacity limitation E_{C}. Thus, the energy state of ESU evolves in accordance with the following recursion:

$$S_{l,i+1} = \min\left\{ \int_0^{\tau} P^{\mathrm{H}}_{l,i}(t) - P^{\mathrm{U}}_{l,i}(t)\mathrm{d}t + S_{l,i}, E_{\mathrm{C}} \right\}. \tag{1}$$

Case 2. If $P^{\mathrm{H}}_{l,i}(t) < P^{\mathrm{U}}_{l,i}(t)$ and $P^{\mathrm{H}}_{l,i}(t) + \bar{P}^{\mathrm{U}}_{l,i}(t) \geqslant P^{\mathrm{U}}_{l,i}(t)$ during τ, ESU and the energy harvester have to jointly power RRU l, and no harvested energy can be stored. Thus, the energy state of ESU evolves as the following recursion, where $[\cdot]^+ = \max(\cdot, 0)$:

$$S_{l,i+1} = \left[S_{l,i} - \int_0^{\tau} P^{\mathrm{U}}_{l,i}(t) - P^{\mathrm{H}}_{l,i}(t)\mathrm{d}t \right]^+. \tag{2}$$

Case 3. If $P^{\mathrm{H}}_{l,i}(t) + \bar{P}^{\mathrm{U}}_{l,i}(t) < P^{\mathrm{U}}_{l,i}(t)$ during τ, RRU l cannot be driven to work, and in turn all harvested energy is stored into ESU. Thus, the energy state of ESU evolves in accordance with the following recursion with the storage capacity limitation E_{C}:

$$S_{l,i+1} = \min\left\{ S_{l,i} + \int_0^{\tau} P^{\mathrm{H}}_{l,i}(t)\mathrm{d}t, E_{\mathrm{C}} \right\}. \tag{3}$$

Proposition 1. If TI is small enough that the relationship among $P^{\mathrm{H}}_{l,i}(t)$, $P^{\mathrm{U}}_{l,i}(t)$ and $\bar{P}^{\mathrm{U}}_{l,i}(t)$ is fixed, the full-duplex energy behavior model is in accordance with the format of the Harvest-Storage-Usage model [7], i.e., $S_{l,i+1} = \min\{(S_{l,i} + E_{l,i} - \mathcal{E}^{\mathrm{con}}_{l,i})^+, E_{\mathrm{C}}\}, \forall i \in \mathbb{N}$, where $E_{l,i}$ denotes the amount of harvested energy at RRU l during TI i; $\mathcal{E}^{\mathrm{con}}_{l,i}$ is the energy consumption of RRU l during TI i, which will be discussed in Subsection 2.2.

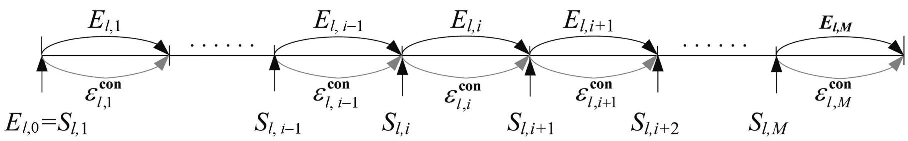

Figure 3 The sequences of energy behavior during each TF.

Proof. See Appendix A.

Following Proposition 1, the sequences of energy behavior during one TF are illustrated in Figure 3. $\{E_{l,1}, E_{l,2}, \ldots, E_{l,M}\}$ are assumed constant at each TI during one TF, which can be predicted at the beginning of each TF by using predictive analytic such as Bayesian forecasting [16]. Since prediction method is not the topic to be discussed, it is assumed that $E_{l,i}$ $(i \in \{1, 2, \ldots, M\})$ is perfectly pre-known at the beginning of each TF. According to Figure 3, states of ESU can be iteratively derived from the initial energy state (denoted as $E_{l,0} \triangleq S_{l,1}$) at the beginning of each TF as

$$S_{l,m} = \left[\sum_{i=0}^{m} \left(E_{l,i} - \mathcal{E}_{l,i}^{\mathrm{con}} \right) - E_{\mathrm{C}} \right]^{+}, \quad m \in \{1, 2, \ldots, M\}. \tag{4}$$

As the harvested energy is pre-known, the state of ESU evolves according to the energy consumption of RRU l, which is modelled in the following subsection.

2.2 Energy consumption model of joint transmission

Consider a block fading channel model where the channel state remains static within each TI, but becomes independent across different TI's. The system bandwidth is equally divided into N sub-channels with bandwidth B_{w}. Full user capacity is considered, which means N sub-channels are allocated to N receivers. L RRU's jointly transmit data on each sub-channel. With the perfect signal phase synchronization, the downlink signal received on sub-channel n $(\forall n \in \{1, 2, \ldots, N\})$ at TI i is

$$y_{n,i} = \sum_{l=1}^{L} \sqrt{G_{n,l,i}} \, |h_{n,l,i}| \, x_{n,l,i} + z_n. \tag{5}$$

In (5), $x_{n,l,i}$ is the symbol transmitted on sub-channel n from RRU l at TI i, the transmit power of which is $p_{n,l,i} = \mathrm{E}[|x_{n,l,i}|^2]$. $G_{n,l,i} \propto [d_{n,l,i}/d_0]^{-\alpha}$ denotes the large-scale path-loss from RRU l, where $d_{n,l,i}$ and d_0 are the distances between the receiver allocated to subchannel n and RRU l at TI i, and the reference distance, respectively. α is the pathloss exponent. $|h_{n,l,i}|$ stands for the gain of the Rayleigh fading channel at TI i, modeled by i.i.d. complex Gaussian with unit variance, i.e., $h_{n,l,i} \sim \mathcal{N}_{\mathcal{C}}(0, 1)$. z_n is the AWG noise, and the noise power for all sub-channels is denoted by σ^2.

Each receiver receives bits from multiple RRU's which cooperate with each other. Then, the achievable transmission bit rate on sub-channel $n \in \{1, 2, \ldots, N\}$ during TI i is given by

$$R_{n,i}(p_{n,l,i}) = B_{\mathrm{w}} \log_2 \left[1 + \left(\sum_{l=1}^{L} \gamma_{n,l,i} \sqrt{p_{n,l,i}} \right)^2 \right], \tag{6}$$

where $\gamma_{n,l,i}^2 = G_{n,l,i} |h_{n,l,i}|^2 / \sigma^2$ denotes carrier-to-noise ratio (CNR) of RRU l on sub-channel n at TI i. The total received data at all receivers during one TF can be calculated as

$$\mathcal{D}_{\mathrm{OP}}(p_{n,l,i}) = \sum_{n=1}^{N} \sum_{i=1}^{M} R_{n,i}(p_{n,l,i}) \cdot \tau. \tag{7}$$

The energy consumption of RRU l at TI i is modeled as

$$\mathcal{E}_{l,i}^{\mathrm{con}} = \left(\varpi \sum_{n=1}^{N} P_{l,i}^{\mathrm{U}} + P_{l}^{\mathrm{C}} \right) \cdot \tau = \left(\varpi \sum_{n=1}^{N} p_{n,l,i} + P_{l}^{\mathrm{C}} \right) \cdot \tau, \tag{8}$$

where ϖ is defined as $\varpi = \varepsilon/\eta$; ε denotes the peak-to-average power ratio (PAPR) and η denotes the power amplifier efficiency; P_l^{C} denotes the receiving and circuit power consumption of RRU l, which is relatively low compared to transmit power, and thus assumed constant for notational convenience. In practical applications, the upper bound of P_l^{C} can be statistically determined through measuring active RRU's.

3 Optimal cross-layer transmission and energy scheduling

The goal of the following formulated problem is to design the optimal power allocation in 3 dimensions of sub-channel, RRU and TI under fluctuated energy supply such that the sum of transmission data across every TF (i.e., M TI's) is maximized:

$$\begin{aligned}
&\max_{p_{n,l,i}} \mathcal{D}_{\mathrm{OP}}(p_{n,l,i}) \\
&\text{s.t.} \ \sum_{i=1}^{m} \mathcal{E}_{l,i}^{\mathrm{con}} \geqslant \left(\sum_{i=0}^{m} E_{l,i} - E_{\mathrm{C}}\right)^{+}, m \in \{1,2,\ldots,M\}, \quad \text{(C1)} \\
&\qquad \sum_{i=1}^{m'} \mathcal{E}_{l,i}^{\mathrm{con}} \leqslant \sum_{i=0}^{m'-1} E_{l,i}, m' \in \{1,2,\ldots,M+1\}, \quad \text{(C2)} \\
&\qquad p_{n,l,i} \geqslant 0, \forall n, \forall l, \forall i, \quad \text{(C3)}
\end{aligned} \tag{9}$$

where $l \in \{1,2,\ldots,L\}$; constraints C1 and C2 are non-overflow constraint and causality constraint of RRU l, respectively. Specifically, seen from Figure 3, constraint C1 states that in order to prevent ESU of RRU l from energy overflowing, at least $(\sum_{i=0}^{m} E_{l,i} - E_{\mathrm{C}})^{+}$ amount of energy should be consumed at the end of TI m, where $m \in \{1,2,\ldots,M\}$; constraint C2 guarantees that the energy consumption of RRU l cannot exceed the amount of harvested and stored energy associated with RRU l during each TI. Constraint C3 represents the non-negative power allocation.

Then, the upper bound of the formulated problem is derived, assuming that CSI during each TF is perfectly pre-known.

Proposition 2. The objective function of (9) is concave.

Proof. See Appendix B.

Utilizing dual decomposition method, the Lagrangian associated with (9) is

$$\begin{aligned}
\mathcal{L}(\varsigma_{l,m}, \mu_{l,m'}, p_{n,l,i}) = &\sum_{n=1}^{N}\sum_{i=1}^{M+1} R_{n,i}(p_{n,l,i})\tau - \sum_{l=1}^{L}\sum_{m'=1}^{M+1} \mu_{l,m'} \sum_{i=1}^{m'} \mathcal{E}_{l,i}^{\mathrm{con}} + \sum_{l=1}^{L}\sum_{m=1}^{M} \varsigma_{l,m} \sum_{i=1}^{m} \mathcal{E}_{l,i}^{\mathrm{con}} \\
&+ \sum_{l=1}^{L}\sum_{m'=1}^{M+1} \mu_{l,m'} \sum_{i=0}^{m'-1} E_{l,i} - \sum_{l=1}^{L}\sum_{m=1}^{M} \varsigma_{l,m} \left(\sum_{i=0}^{m} E_{l,i} - E_{\mathrm{C}}\right)^{+},
\end{aligned} \tag{10}$$

where $\{\varsigma_{l,m}\}_{L\times M}$ and $\{\mu_{l,m'}\}_{L\times(M+1)}$ are Lagrange multiplier matrixes associated with C1 and C2, respectively. Then, the dual problem is given by

$$\min_{\varsigma_{l,m}, \mu_{l,m'} \geqslant 0} \ \max_{p_{n,l,i} \geqslant 0} \ \mathcal{L}(\varsigma_{l,m}, \mu_{l,m'}, p_{n,l,i}). \tag{11}$$

Making the derivative of $\mathcal{L}(\varsigma_{l,m}, \mu_{l,m'}, p_{n,l,i})$ in (10) respecting to $p_{n,l,i}$ equal to zero yields

$$\frac{\partial \mathcal{L}(\varsigma_{l,m}, \mu_{l,m'}, p_{n,l,i})}{\partial p_{n,l,i}} = \frac{\tau W_{n,i} \gamma_{n,l,i}}{\sqrt{p_{n,l,i}} \ln 2} - \mathcal{K}_{l,i} \xi \tau = 0, \tag{12}$$

where

$$W_{n,i} = \sum_{g=1}^{L} \gamma_{n,g,i} \sqrt{p_{n,g,i}} \Big/ \left[1 + \left(\sum_{g=1}^{L} \gamma_{n,g,i} \sqrt{p_{n,g,i}}\right)^{2}\right], \tag{13}$$

$$\mathcal{K}_{l,i}=\sum_{m'=i}^{M+1}\mu_{l,m'}-\sum_{m=i}^{M}\varsigma_{l,m}. \tag{14}$$

Define $\varphi_{n,i}=\{l|p_{n,l,i}>0,\ l=1,2,\ldots,L\}$, $\forall n\in\{1,2,\ldots,N\}$ as the set of RRU's which have the positive transmit power on sub-channel n at TI i. According to (12), there is

$$\frac{\gamma_{n,l_1,i}}{\mathcal{K}_{l_1,i}\sqrt{p_{n,l_1,i}}}=\cdots=\frac{\gamma_{n,l_\vartheta,i}}{\mathcal{K}_{l_\vartheta,i}\sqrt{p_{n,l_\vartheta,i}}}=\cdots=\frac{\gamma_{n,l_{L_n},i}}{\mathcal{K}_{l_{L_n},i}\sqrt{p_{n,l_{L_n},i}}}=\frac{\varpi\ln 2}{W_{n,i}}, \tag{15}$$

where $\vartheta\in\{1,2,\ldots,L_n\}$ and $l_1,\ldots,l_{L_n}\in\varphi_{n,i}$. $L_n=|\varphi_{n,i}|$ is the size of $\varphi_{n,i}$. Therefore, $p_{n,l_\vartheta,i}$ can be derived as

$$\sqrt{p_{n,l_\vartheta,i}}=\frac{\gamma_{n,l_\vartheta,i}}{\gamma_{n,l_1,i}}\cdot\frac{\mathcal{K}_{l_1,i}}{\mathcal{K}_{l_\vartheta,i}}\cdot\sqrt{p_{n,l_1,i}}, \tag{16}$$

where $\mathcal{K}_{l_1,i}\cdot\gamma_{n,l_\vartheta,i}$ can be viewed as weighted CNR of RRU l_ϑ. Then, as the transmit power of RRU 1 is determined, then that of RRU l_ϑ can be easily determined, which can effectively improve computation efficiency. Substituting (16) into (12) and (13) and making $p_{n,l,i}=0$ ($\forall l\notin\varphi_{n,i}$), the optimal solution of (9) can be expressed as

$$p^*_{n,l,i}=\begin{cases}\frac{\sum_{g\in\varphi^*_{n,i}}\frac{\gamma^2_{n,g,i}}{\mathcal{K}_{g,i}}}{\varpi H_{n,l,i}\mathcal{K}^2_{l,i}\ln 2}-\frac{1}{H_{n,l,i}\mathcal{K}^2_{l,i}}, & l\in\varphi^*_{n,i},\\ 0, & l\notin\varphi^*_{n,i},\end{cases} \tag{17}$$

where $\varphi^*_{n,i}=\{l|p^*_{n,l,i}>0,\ l=1,2,\ldots,L\}$ and

$$H_{n,l,i}=\frac{\left(\sum_{g\in\varphi^*_{n,i}}\frac{\gamma^2_{n,g,i}}{\mathcal{K}_{g,i}}\right)^2}{\gamma^2_{n,l,i}}.$$

Seen from (17), the optimal set $\varphi^*_{n,i}$ needs to be selected in determining $p^*_{n,l,i}$, which may need to search over 2^L cases for each TI. In the following, we will prove that there is only two possible cases for $\varphi^*_{n,i}$. Two Lemmas are first given as follows.

Lemma 1. The solution satisfying (17) with all $p^*_{n,l,i}>0$, i.e., $\varphi^*_{n,i}=\{1,2,\ldots,L\}$, is an extreme point solution. Moreover, it is also the optimal solution of (9), if exists, due to the concavity of the objective function in (9), c.f., Proposition 2.

Lemma 2. The solutions satisfying (17) with at least one $p^*_{n,l,i}=0$, i.e., $\varphi^*_{n,i}\subset\{1,2,\ldots,L\}$, are boundary point solutions. When the extreme point is not feasible, the optimal solution must be one of these solutions.

To solve the optimization problem (9), it is natural to first check whether the extreme point is feasible. If not, check each boundary point solution. So, based on Lemmas 1 and 2, the following proposition is provided.

Proposition 3. If the extreme point solution does not exist, the feasible solution of (9) must be the solution with all $p^*_{n,l,i}=0$, i.e., $\varphi^*_{n,i}=\emptyset$.

Proof. According to (17), if the extreme point solution does not exist, there is

$$p^*_{n,l,i}<0\Leftrightarrow\sum_{g\in\varphi^*_{n,i}}\frac{\gamma^2_{n,g,i}}{\mathcal{K}_{g,i}}<\xi\ln 2, \tag{18}$$

where $\varphi^*_{n,i}=\{1,2,\ldots,L\}$.

Suppose there exists a boundary point solution $\{p'_{n,l,i}\}$ satisfying (17) with $\varphi'_{n,i}=\{l|p'_{n,l,i}>0,\ l=1,2,\ldots,L\}\neq\emptyset$ and $\varphi'_{n,i}\subset\varphi^*_{n,i}$, then

$$\sum_{g\in\varphi'_{n,i}}\frac{\gamma^2_{n,g,i}}{\mathcal{K}_{g,i}}<\sum_{g\in\varphi^*_{n,i}}\frac{\gamma^2_{n,g,i}}{\mathcal{K}_{g,i}}<\xi\ln 2\Leftrightarrow p'_{n,l,i}<0,l\in\varphi'_{n,i}, \tag{19}$$

i.e., the solution $\{p'_{n,l,i}\}$ is not feasible. Hence, except the solution with $p^*_{n,l,i}=0$ and $\varphi^*_{n,i}=\emptyset$, none of the boundary point solutions is feasible.

Therefore, for the optimal solution of (9) in (17), $\varphi^*_{n,i}$ is equal to either $\{1,2,\ldots,L\}$ or $\emptyset$, and the optimal solution of (9) can be derived as

$$p^*_{n,l,i}=\left[\frac{1}{H_{n,l,i}\mathcal{K}^2_{l,i}}\left(\frac{\sum_{l=1}^{L}\frac{\gamma^2_{n,l,i}}{\mathcal{K}_{l,i}}}{\varpi\ln 2}-1\right)\right]^+,\forall n\in\{1,2,\ldots,N\},\forall l\in\{1,2,\ldots,L\},\forall i\in\{1,2,\ldots,M\}.\quad(20)$$

The gradient update equations for all entries in matrices $\{\varsigma_{l,m}\}_{L\times M}$ and $\{\mu_{l,m'}\}_{L\times(M+1)}$ are given:

$$\varsigma^{\upsilon+1}_{l,m}=\left[\varsigma^{\upsilon}_{l,m}-\kappa^{\upsilon}_1\left(\sum_{i=1}^{m}\mathcal{E}^{\text{con}}_{l,i}-\left(\sum_{i=0}^{m}E_{l,i}-E_{\text{C}}\right)^+\right)\right]^+,\quad(21)$$

$$\mu^{\upsilon+1}_{l,m'}=\left[\mu^{\upsilon}_{l,m'}-\bar{\kappa}^{\upsilon}_1\left(\sum_{i=0}^{m'-1}E_{l,i}-\sum_{i=1}^{m'}\mathcal{E}^{\text{con}}_{l,i}\right)\right]^+,\quad(22)$$

where $\forall l\in\{1,2,\ldots,L\}$; $\forall m\in\{1,2,\ldots,M\}$; $\forall m'\in\{1,2,\ldots,M+1\}$; $\upsilon\in\mathbb{N}$ is the iteration number; κ^{υ}_1 and $\bar{\kappa}^{\upsilon}_1$ are the sequence of scalar step sizes.

The optimal cross-layer transmission and energy scheduling algorithm is summarized in Algorithm 1, the convergence of which has been proved in [17].

Algorithm 1 Optimal cross-layer transmission and energy scheduling algorithm

Require: $\Delta\to 0^+$
  if υ=0 **then**
    $\{\varsigma^{\upsilon}_{l,m}\}_{L\times M}\Leftarrow\boldsymbol{I}$ and $\{\mu^{\upsilon}_{l,m'}\}_{L\times(M+1)}\Leftarrow\boldsymbol{I}$;
  end if
  flag $\Leftarrow 1$;
  while flag **do**
    $\{\varsigma^{\upsilon+1}_{l,m}\}_{L\times M}\Leftarrow\left[\varsigma^{\upsilon}_{l,m}-\kappa^{\upsilon}_1\left(\sum\limits_{i=1}^{m}\mathcal{E}^{\text{con}}_{l,i}-\left(\sum\limits_{i=0}^{m}E_{l,i}-E_C\right)^+\right)\right]^+_{L\times M}$;
    $\{\mu^{\upsilon+1}_{l,m'}\}_{L\times(M+1)}\Leftarrow\left[\mu^{\upsilon}_{l,m'}-\bar{\kappa}^{\upsilon}_1\left(\sum\limits_{i=0}^{m'-1}E_{l,i}-\sum\limits_{i=1}^{m'}\mathcal{E}^{\text{con}}_{l,i}\right)\right]^+_{L\times(M+1)}$;
    $\{p^{\upsilon+1}_{n,l,i}\}_{N\times L\times M}\Leftarrow\left[\frac{1}{H_{n,l,i}\mathcal{K}^2_{l,i}}\left(\frac{\sum_{l=1}^{L}\frac{\gamma^2_{n,l,i}}{\mathcal{K}_{l,i}}}{\varpi\ln 2}-1\right)\right]^+_{N\times L\times M}$;
    if $\mathcal{D}_{\text{OP}}(p^{\upsilon+1}_{n,l,i})-\mathcal{D}_{\text{OP}}(p^{\upsilon}_{n,l,i})<\Delta$ **then**
      $\{p^*_{n,l,i}\}_{N\times L\times M}\Leftarrow\{p^{\upsilon}_{n,l,i}\}_{N\times L\times M}$;
      flag $\Leftarrow 0$;
    end if
    $\upsilon\Leftarrow\upsilon+1$;
  end while

4 Online cross-layer transmission and energy scheduling

Although prediction techniques [13] have the potential to effectively make a good channel state prediction, inherent random factors in the wireless propagation inevitably lead to prediction errors. Moreover, the longer the prediction period, the greater the prediction errors [13]. To avoid the impact of greater channel state prediction errors on the objective function, we first design the online algorithm as that the optimal power allocation is executed at every TI with the help of predicted channel state at one TI ahead, as illustrated in Figure 4. Assuming the current TI index is k, the prediction errors at TI $k+1$ are defined as matrix $\mathfrak{E}_{k+1}\triangleq\{\mathfrak{e}_{n,l,k+1}\}_{N\times L}$. Then, the sum of transmission data, utilizing $\gamma_{n,l,k}$ and the predicted

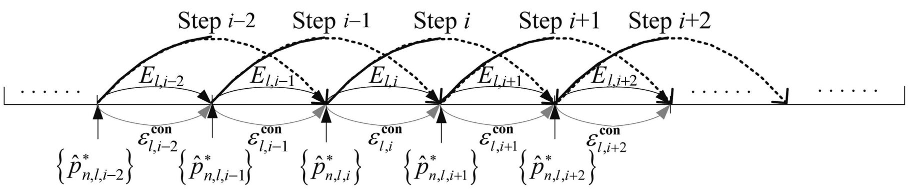

Figure 4 The diagram of the online cross-layer transmission and energy scheduling.

$\gamma_{n,l,k+1}$, is given by

$$\mathcal{D}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)=\sum_{n=1}^{N}\sum_{i=k}^{k+1}R_{n,i}(\hat{p}_{n,l,i})\tau=\tau\cdot\sum_{n=1}^{N}\sum_{i=k}^{k+1}B_{\mathrm{w}}\log_2\left[1+\left(\sum_{l=1}^{L}\gamma_{n,l,i}\sqrt{\hat{p}_{n,l,i}}\right)^2\right],\ \forall k, \tag{23}$$

where $\hat{p}_{n,l,i}$ $(i\in\{k,k+1\},\forall k\in\{1,2,\ldots,M\})$ are the allocated power on sub-channel n of RRU l at TI k and TI $k+1$. It is obvious that the sum of transmission data in (23) is a monotone increasing function of $\hat{p}_{n,l,i}$ $(i\in\{k,k+1\})$.

Recalling (8), the energy consumption of $\mathcal{D}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)$ in (23) is

$$\mathcal{Q}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)=\sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\mathrm{con}}_{l,i}(\hat{p}_{n,l,i})=\sum_{l=1}^{L}\sum_{i=k}^{k+1}\left(\varpi\sum_{n=1}^{N}\hat{p}_{n,l,i}+P_l^{\mathrm{C}}\right)\cdot\tau,\forall k\in\{1,2,\ldots,M\}, \tag{24}$$

which is also a monotone increasing function of $\hat{p}_{n,l,i}$ $(i\in\{k,k+1\})$.

Intuitively, utilizing characteristics that $\mathcal{D}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)$ and $\mathcal{Q}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)$ are monotonically increasing functions of the same variable $\hat{p}_{n,l,i}$, we reformulate the objective function as the ratio of $\mathcal{D}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)$ to $\mathcal{Q}_{\mathrm{ON}}(\hat{p}_{n,l,i})$ to mitigate the deviation caused by prediction errors $\mathfrak{e}_{n,l,k+1}{}'s$, c.f., (25):

$$\mathcal{F}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)=\frac{\bar{\mathcal{D}}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)}{\bar{\mathcal{Q}}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)}=\frac{\sum_{n=1}^{N}\sum_{i=0}^{k-1}B_{n,i}+\sum_{n=1}^{N}\sum_{i=k}^{k+1}R_{n,i}(\hat{p}_{n,l,i})\tau}{\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\mathrm{con}}_{l,i}+\sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\mathrm{con}}_{l,i}(\hat{p}_{n,l,i})},\forall k\in\{1,2,\ldots,M\}, \tag{25}$$

the physical meaning of which is the average transmission energy efficiency in unit [bit/Joule]. In (25), two constants are introduced to guarantee comparability in terms of time length of TF between the optimal and online algorithms, which are $\sum_{n=1}^{N}\sum_{i=0}^{k-1}B_{n,i}=\sum_{n=1}^{N}\sum_{i=0}^{k-1}\log_2[1+(\sum_{l=1}^{L}\gamma_{n,l,i}\sqrt{\hat{p}_{n,l,i}})^2]\tau$, denoting the total transmission data before TI $(k-1)$ in each TF, and $B_{n,0}=0$; $\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\mathrm{con}}_{l,i}=\sum_{l=1}^{L}\sum_{i=0}^{k-1}(\varpi\sum_{n=1}^{N}\hat{p}_{n,l,i}+P_{\mathrm{C}})\tau$, denoting the total energy consumed before TI k in each TF, and $\overline{\mathcal{E}}^{\mathrm{con}}_{l,0}=0$.

Therefore, the online optimization problem for determining optimal power allocation at TI k is formulated as

$$\begin{aligned}\max_{p_{n,l,i}}\ & \mathcal{F}_{\mathrm{ON}}(\hat{p}_{n,l,i},k)\\ \text{s.t.}\ & \sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\mathrm{con}}_{l,i}+\sum_{i=k}^{k+1}\mathcal{E}^{\mathrm{con}}_{l,i}(\hat{p}_{n,l,i})\geqslant\left(\sum_{i=0}^{k}E_{l,i}-E_{\mathrm{C}}\right)^{+}, && \text{(C4)}\\ & \sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\mathrm{con}}_{l,i}+\sum_{i=k}^{u}\mathcal{E}^{\mathrm{con}}_{l,i}(\hat{p}_{n,l,i})\leqslant\sum_{i=0}^{u-1}E_{l,i},u\in\{k,k+1\}, && \text{(C5)}\\ & \hat{p}_{n,l,i}\geqslant 0,\forall n,\forall l,\forall i, && \text{(C6)}\end{aligned} \tag{26}$$

where $k\in\{1,2,\ldots,M\}$. Because the objective function of (26) is a nonlinear fractional programming problem [18], it can be associated with a parametric programming (PP) problem as follows:

$$\max_{p_{n,l,i}}\left[\left(\sum_{n=1}^{N}\sum_{i=0}^{k-1}B_{n,i}+\sum_{n=1}^{N}\sum_{i=k}^{k+1}R_{n,i}(\hat{p}_{n,l,i})\tau\right)-q\left(\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\mathrm{con}}_{l,i}+\sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\mathrm{con}}_{l,i}\right)\right], \tag{27}$$

where $q\in\mathbb{R}$ is referred as a parameter which determines the relative weight of the total energy consumption of system and can be interpreted as the overhead caused by energy consumption.

Proposition 4. Let $\hat{p}^*_{n,l,i}$ be the optimal variable in (26) with a fixed q. $\hat{p}^*_{n,l,i}$ is also optimal for (27), if and only if

$$\sum_{n=1}^{N}\sum_{i=0}^{k-1} B_{n,i} + \sum_{n=1}^{N}\sum_{i=k}^{k+1} R_{n,i}(\hat{p}^*_{n,l,i})\tau - q\left(\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i}\right) = 0. \tag{28}$$

Proof. Because $\{B_{n,i}\}$ and $\{\overline{\mathcal{E}}^{\rm con}_{l,i}\}$ are constant sequences and $\sum_{l=1}^{L}\sum_{i=k}^{k+1} E^{\rm con}_{l,i}$ is the linear sum of $\hat{p}^*_{n,l,i}$, which satisfies the convexity, the denominator of (25) is obviously convex. Recalling the proof of Proposition 2, the numerator of (25), i.e., $(\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i})$, can be similarly proved to be concave.

Therefore, according to Dinkelbach's theory, Proposition 4 can be proved.

Then, following Proposition 4, the primal problem in (26) is equivalent with

$$\max_{\hat{p}_{n,l,i}}\left(\sum_{n=1}^{N}\sum_{i=1}^{k-1} B_{n,i} + \sum_{n=1}^{N}\sum_{i=k}^{k+1} R_{n,i}(\hat{p}_{n,l,i})\tau - q\left(\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i}\right)\right) \quad \text{s.t.} \quad \text{C4, C5, C6}. \tag{29}$$

The Lagrangian of (29) is given by

$$\begin{aligned}\mathcal{L}(\rho_l,\psi_{l,u},p_{n,l,i}) = {}& \sum_{l=1}^{L}\rho_l\left(\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i} - \left(\sum_{i=0}^{k+1}E_{l,i} - E_C\right)^+\right)\\ &+ \sum_{l=1}^{L}\sum_{u=k}^{k+1}\psi_{l,u}\left(\sum_{i=0}^{u-1}E_{l,i} - \left(\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{i=k}^{u}\mathcal{E}^{\rm con}_{l,i}\right)\right)\\ &+ \sum_{n=1}^{N}\sum_{i=0}^{k-1}B_{n,i} + \sum_{n=1}^{N}\sum_{i=k}^{k+1}R_{n,i}(\hat{p}_{n,l,i})\tau - q\left(\sum_{l=1}^{L}\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{l=1}^{L}\sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i}\right),\end{aligned} \tag{30}$$

where $\{\rho_l\}_{L\times 1}$ and $\{\psi_{l,u}\}_{L\times 2}$ denote the Lagrange multiplier vector and matrix associated with constraints C4 and C5, respectively. Thus, the dual optimization problem is given by

$$\min_{\rho_l,\psi_{l,u}\geqslant 0}\ \max_{\hat{p}_{n,l,i}}\ \mathcal{L}(\rho_l,\psi_{l,u},\hat{p}_{n,l,i}). \tag{31}$$

By using dual decomposition approach as (12)–(20), the optimal solution of (29) with a fixed q can be described as

$$\hat{p}^*_{n,l,i} = \left[\frac{1}{\widetilde{H}_{n,l,i}\widetilde{\mathcal{K}}^2_{l,i}}\left(\sum_{l=1}^{L}\frac{\gamma^2_{n,l,i}}{\widetilde{\mathcal{K}}_{l,i}}\Big/(\xi\ln 2) - 1\right)\right]^+, \tag{32}$$

where $\forall n\in\{1,2,\ldots,N\}$, $\forall l\in\{1,2,\ldots,L\}$ and $\forall i\in\{k,k+1\}$; $\widetilde{\mathcal{K}}_{l,i}$ is given by

$$\widetilde{\mathcal{K}}_{l,i} = \begin{cases}\sum\limits_{u=k}^{k+1}\psi_{l,u} - \rho_l + q, & i=k,\\ \psi_{l,k+1} + q, & i=k+1,\end{cases} \tag{33}$$

and $\widetilde{H}_{n,l,i}$ is described as

$$\widetilde{H}_{n,l,i} = \frac{\left(\sum_{g\in\varphi^*_{n,i}}\frac{\gamma^2_{n,g,i}}{\widetilde{\mathcal{K}}_{g,i}}\right)^2}{\gamma^2_{n,l,i}}. \tag{34}$$

The gradient update equations for $\{\rho_l\}_{L\times 1}$ and $\{\psi_{l,u}\}_{L\times 2}$ are given as

$$\rho_l^{v+1} = \left[\rho_l^v - \kappa_2^v\left(\sum_{i=0}^{k-1}\overline{\mathcal{E}}^{\rm con}_{l,i} + \sum_{i=k}^{k+1}\mathcal{E}^{\rm con}_{l,i} - \left(\sum_{i=0}^{k+1}E_{l,i} - E_{\rm C}\right)^+\right)\right]^+, \tag{35}$$

Table 1 Simulation parameters

Simulation parameter	Value	Simulation parameter	Value
Number of RRU's (L)	3	Static circuit power consumption ($P_{\rm C}$ dBm)	40
Number of sub-channels (N)	12	Power amplifier efficiency (η)	30%
Length of each time interval (τ ms)	300	Peak-to-average power ratio (ε dB)	12
Path-loss exponent (α)	3	Noise power (σ^2 dBm)	-128
Reference distance (d_0 m)	10		

$$\psi_{l,u}^{\upsilon+1} = \left[\psi_{l,u}^{\upsilon} - \bar{\kappa}_2^{\upsilon}\left(\sum_{i=0}^{u-1} E_{l,i} - \sum_{i=0}^{k-1} \overline{\mathcal{E}}_{l,i}^{\rm con} - \sum_{i=k}^{u} \mathcal{E}_{l,i}^{\rm con}\right)\right]^+, \tag{36}$$

where $l \in \{1,2,\ldots,L\}$; $u \in \{k,k+1\}$; υ is the integer iteration number; κ_2^{υ} and $\bar{\kappa}_2^{\upsilon}$ are the sequences of scalar step size.

The online cross-layer transmission and energy scheduling algorithm is presented in Algorithm 2. The convergence of the algorithm converges with a superlinear convergence rate, which is proved in [18].

Algorithm 2 Online cross-layer transmission and energy scheduling algorithm

Require: $\Delta \to 0^+$ and $q \to 0^+$;
 if υ=0 **then**
  $\{\rho_l^{\upsilon}\}_{L\times 1} \Leftarrow \boldsymbol{I}$ and $\left\{\phi_{l,u}^{\upsilon}\right\}_{L\times 2} \Leftarrow \boldsymbol{I}$;
 end if
 $k \Leftarrow 1$;
 while $k \leqslant 12$ **do**
  flag $\Leftarrow 1$;
  while flag **do**
   $\{\rho_l^{\upsilon+1}\}_{L\times 1} \Leftarrow \left[\rho_l^{\upsilon} - \kappa_2^{\upsilon}\left(\sum\limits_{i=0}^{k-1}\overline{\mathcal{E}}_{l,i}^{\rm con} + \sum\limits_{i=k}^{k+1}\mathcal{E}_{l,i}^{\rm con} - \left(\sum\limits_{i=0}^{k+1}E_{l,i} - E_{\rm C}\right)^+\right)\right]_{L\times 1}^+$;
   $\{\psi_{l,u}^{\upsilon+1}\}_{L\times 2} \Leftarrow \left[\psi_{l,u}^{\upsilon} - \bar{\kappa}_2^{\upsilon}\left(\sum\limits_{i=0}^{u-1}E_{l,i} - \sum\limits_{i=0}^{k-1}\overline{\mathcal{E}}_{l,i}^{\rm con} - \sum\limits_{i=k}^{u}\mathcal{E}_{l,i}^{\rm con}\right)\right]_{L\times 2}^+$;
   $\{\widetilde{\mathcal{K}}_{l,k}\}_{L\times 1} \Leftarrow \left\{\sum\limits_{u=k}^{k+1}\psi_{l,u} - \rho_l + q\right\}_{L\times 1}$ and $\{\widetilde{\mathcal{K}}_{l,k+1}\}_{L\times 1} \Leftarrow \{\psi_{l,k+1}+q\}_{L\times 1}$;
   $\{\widetilde{H}_{n,l,i}\}_{N\times L\times 1} \Leftarrow \left\{\left(\sum\limits_{g\in\varphi_{n,i}^*}\frac{\gamma_{n,g,i}^2}{\widetilde{\mathcal{K}}_{g,i}}\right)^2 \Big/ \gamma_{n,l,i}^2\right\}_{N\times L\times 1}$;
   $\{\hat{p}_{n,l,i}^{\upsilon+1}\}_{N\times L\times 1} \Leftarrow \left[\frac{1}{\widetilde{H}_{n,l,i}\widetilde{\mathcal{K}}_{l,i}^2}\left(\sum\limits_{l=1}^{L}\frac{\gamma_{n,l,i}^2}{\widetilde{\mathcal{K}}_{l,i}}/(\xi\ln 2) - 1\right)\right]_{N\times L\times 1}^+$;
   $\bar{\mathcal{D}}_{\rm ON} \Leftarrow \sum\limits_{n=1}^{N}\sum\limits_{i=0}^{k-1}B_{n,i} + \sum\limits_{n=1}^{N}\sum\limits_{i=k}^{k+1}R_{n,i}(\hat{p}_{n,l,i})\tau$ and $\bar{\mathcal{Q}}_{\rm ON}(\hat{p}_{n,l,i},k) \Leftarrow \sum\limits_{l=1}^{L}\sum\limits_{i=0}^{k-1}\overline{\mathcal{E}}_{l,i}^{\rm con} + \sum\limits_{l=1}^{L}\sum\limits_{i=k}^{k+1}\mathcal{E}_{l,i}^{\rm con}(\hat{p}_{n,l,i})$;
   $\mathcal{F}_{\rm ON}(\hat{p}_{n,l,i},k) \Leftarrow \frac{\bar{\mathcal{D}}_{\rm ON}(\hat{p}_{n,l,i},k)}{\bar{\mathcal{Q}}_{\rm ON}(\hat{p}_{n,l,i},k)}$;
   if $\left|q - \mathcal{F}_{\rm ON}(\hat{p}_{n,l,i}^{\upsilon+1},k)\right| < \Delta$ **then**
    $\{\hat{p}_{n,l,k}^*\}_{L\times 1} \Leftarrow \{\hat{p}_{n,l,k}^{\upsilon+1}\}_{L\times 1}$;
    $flag \Leftarrow 0$;
   end if
   $\upsilon \Leftarrow \upsilon+1$ and $k \Leftarrow k+1$;
  end while
 end while

5 Numerical analysis

Numerical results of the optimal and the online cross-layer transmission and energy scheduling solution[3)] are presented under given profile of energy arrival rates and Rayleigh channels. Without loss of generality, simulation parameters obtained from [3] and [5] are listed in Table 1.

3) They are named as the optimal algorithm and the online algorithm for brevity in the following, respectively.

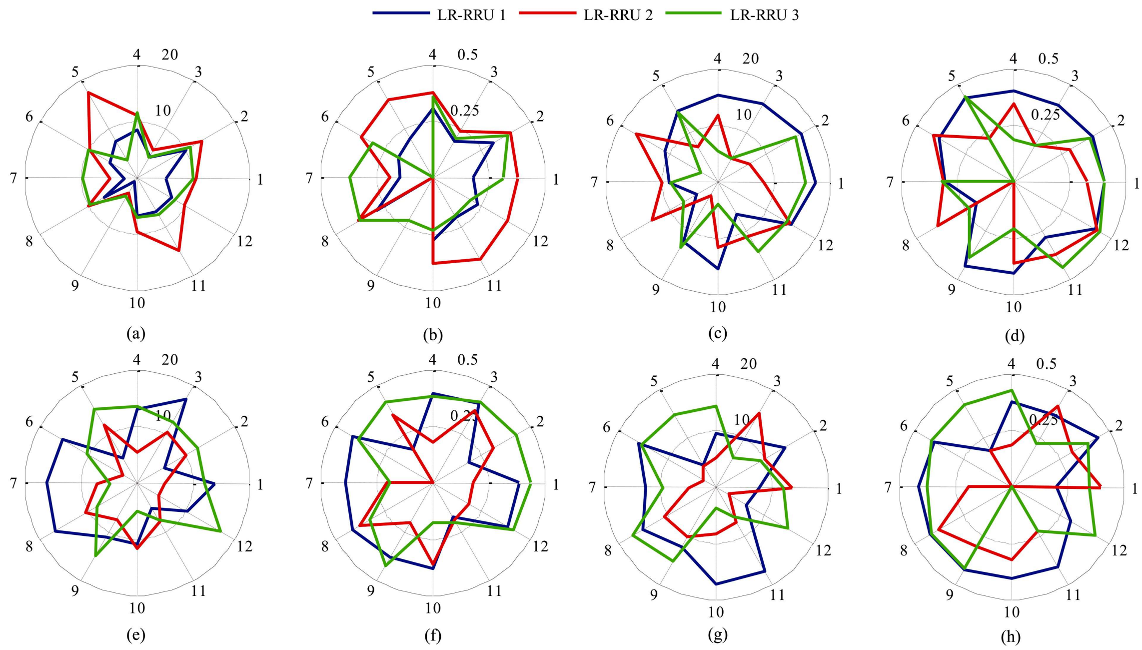

Figure 5 Weighted CNR's and allocated power during 10 TI's during one randomly selected TF for 3 RRU's. (a) Weighted CNRs during the 1st TI; (b) allocated power during the 1st TI; (c) weighted CNRs during the 3rd TI; (d) allocated power during the 3rd TI; (e) weighted CNRs during the 5th TI; (f) allocated power during the 5th TI; (g) weighted CNRs during the 7th TI; (h) allocated power during the 7th TI.

5.1 Snapshots of cross-layer transmission and energy scheduling

Setting $M = 10$, 5 snapshots of one randomly selected TF show characteristics of power adaptation with the weighted CNR's through the optimal algorithm, c.f., Figure 5. Radar plots of Figure 5(a), (c), (e) and (g) track weighted CNR's of 3 RRU's, and radar plots of Figure 5(b), (d), (f) and (h) track the allocated power, where the radii denote the values of weighted CNR's or normalized allocated power, and the dial at every other $\pi/6$ rotated from the x-axis reflects sub-channel index of 3 RRU's distinguished by different colors. Comparing radar plots of weighted CNR's and allocated power at the same TI, e.g., Figure 5(a) and (b), it can be observed that changing trend of tracked picture contours is the same. Specifically, for each sub-channel, the larger the weighted CNR of any RRU, the more the allocated transmit power; the greater the sum of weighted CNR's of any sub-channel at 3 RRU's, the greater the sum of allocated transmit power on the sub-channel; as the weighted CNR of any sub-channel are below certain values, e.g., 3.4 for RRU 2 in Figure 5(a), the allocated power of corresponding RRU is zero, which validates the result of (15). Correspondingly, the conclusion can be drawn that if the power allocation of any RRU l on a certain sub-channel at any TI is obtained, the power allocation of other RRU's can be directly obtained with their corresponding weighted CNR's as given in (15), which brings along significant operation efficiency by reducing $L - 1$ times the amount of optimal power allocation calculation. The results of the online algorithm also comply with the results, which are not provided due to the space limitation.

5.2 Statistical results of the optimal and online algorithms

To numerically illustrate the relationship between the average throughput and channel state prediction errors, and show the performance of proposed algorithms, the following metrics derived from the above results are first presented. In the following, $p^*_{n,l,i}$ and $\hat{p}^*_{n,l,i}$ are the optimal power allocations obtained through the optimal and online algorithm, respectively.

Definition 1. Average throughput of the optimal algorithm in M TI's is defined as

$$\mathcal{D}^*_{\mathrm{OP}} = \mathcal{D}_{\mathrm{OP}}(p^*_{n,l,i})/(M \cdot \tau). \tag{37}$$

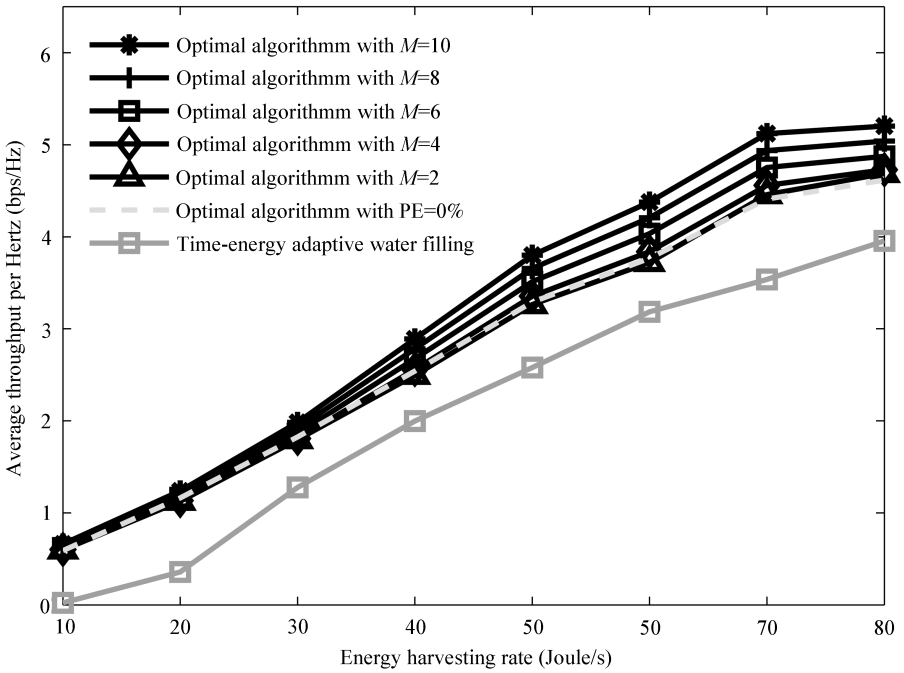

Figure 6 Average throughput vs. energy harvesting rate.

Definition 2. Average throughput of the online algorithm in M TI's is defined as

$$\mathcal{D}_{\rm ON}^{*}=\sum_{n=1}^{N}\sum_{i=1}^{M}\log_2\left[1+\left(\sum_{l=1}^{L}\gamma_{n,l,i}\sqrt{\hat{p}_{n,l,i}^{*}}\right)^2\right]\Big/M. \tag{38}$$

Definition 3. Average energy efficiency of the optimal algorithm is defined as

$$\mathcal{F}_{\rm OP}^{*}=\mathcal{D}_{\rm OP}(p_{n,l,i}^{*})\Big/\sum_{l=1}^{L}\sum_{i=1}^{N}\mathcal{E}_{l,i}^{\rm con}(p_{n,l,i}^{*}). \tag{39}$$

Definition 4. Average energy efficiency of the online algorithm in M TI's is defined as

$$\mathcal{F}_{\rm ON}^{*}=\mathcal{F}_{\rm ON}(\hat{p}_{n,l,i}^{*},M-1). \tag{40}$$

Definition 5. The normalized deviation of the online average throughput is defined as

$$R_1\big|_{({\rm PE}=x)}=(\mathcal{D}_{\rm OP}^{*}-\mathcal{D}_{\rm ON}^{*}\big|_{({\rm PE}=x)})/\mathcal{D}_{\rm OP}^{*}, \tag{41}$$

where $\mathcal{D}_{\rm ON}^{*}\big|_{({\rm PE}=x)}$ denotes the sum of transmission data of the online algorithm operated with prediction errors randomly distributed in range $[0\ x]$; x is valued in set $\{0\%,1\%,5\%,10\%,15\%,20\%\}$.

Definition 6. The normalized deviation of the online average energy efficiency is defined as

$$R_2\big|_{({\rm PE}=x)}=(\mathcal{F}_{\rm OP}^{*}-\mathcal{F}_{\rm ON}^{*}\big|_{({\rm PE}=x)})/\mathcal{F}_{\rm OP}^{*}, \tag{42}$$

where $\mathcal{F}_{\rm ON}^{*}\big|_{({\rm PE}=x)}$ denotes the online algorithm with all prediction errors randomly distributed in range $[0\ x]$; x is valued in set $\{0\%,1\%,5\%,10\%,15\%,20\%\}$.

Figure 6 provides $\mathcal{D}_{\rm OP}^{*}$ and $\mathcal{D}_{\rm ON}^{*}$ with varied energy harvesting rates and M's, which numerically exhibits the relationship among the energy harvesting rate, the TF size, and the average throughput. As for the optimal algorithm, it can be first observed that $\mathcal{D}_{\rm OP}^{*}$ increases with the TF size increasing. This is because more global optimization gains on $\mathcal{D}_{\rm OP}^{*}$ can be obtained as TF size increases. Interestingly enough, as the energy harvesting rate is lower than 20 Joule/TI, $\mathcal{D}_{\rm ON}^{*}$ approximates to $\mathcal{D}_{\rm OP}^{*}$. As the

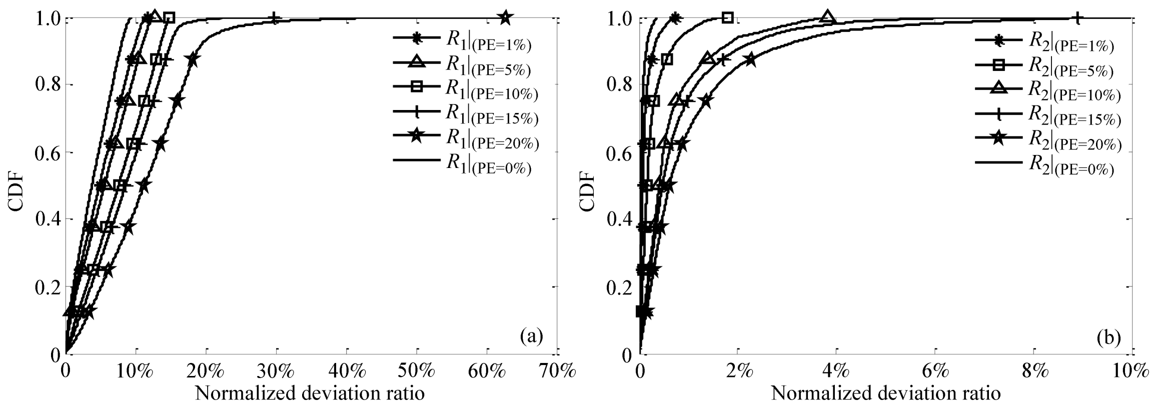

Figure 7 CDF of normalized deviation. (a) CDF of normalized deviation of the online average throughput; (b) CDF of normalized deviation of the online average energy efficiency.

energy harvesting rate further increases, the maximal difference between $\mathcal{D}^*_{\mathrm{ON}}$ and $\mathcal{D}^*_{\mathrm{OP}}$ is only 8%. Compared with time-energy adaptive water filling algorithm in [19], $\mathcal{D}^*_{\mathrm{ON}}$ is greater by at least 24%, due to transmission efficiency improvement brought by joint transmissions and full-duplex energy harvesting. The above results correspond to our motivation problems 2 and 3, which not only provide the optimal design limitation of cross-layer transmission and energy scheduling, but also show the better performance in terms of the throughput in comparison with the existing algorithm.

Seen from CDF curves in Figure 7(a), R_1's with PE $= 0\%, 1\%, 5\%, 10\%, 15\%$ can be guaranteed below 8.35%, 11.9%, 12.95%, 14.5% and 21% with probability 100%. Even as PE $=$ 20%, R_1 can be guaranteed below 19% with probability 90%. It can be seen from CDF curves in Figure 7(b) that R_2's with PE $= 1\%, 5\%, 10\%$ can be guaranteed below 1.5%, 3.6%, and 7.4% with probability 100%. Even as PE $= 15\%, 20\%$, R_2's can be assured below 8% with probability 93% and 97.5%, respectively. Seen from Figure 7(a) and (b), the impact of x on $\mathcal{F}^*_{\mathrm{ON}}$ is well mitigated by the ratio format of (25), compared with its impact on $\mathcal{D}^*_{\mathrm{ON}}$. Corresponding to the motivation problem 3 of mitigating the performance degrading effects brought by CSI prediction errors, above numerical results illustrate that the proposed online algorithm achieves effectively robust to the CSI prediction error, even as it reaches 20%.

6 Conclusion

In this paper, we have studied the optimal and online cross-layer transmission and energy scheduling under full-duplex energy harvesting wireless OFDM joint transmissions. Corresponding to three motivation problems raised in Section 1, we can draw the following conclusions. (1) To aid the analysis and design of energy harvesting communication system, this paper has significantly advanced the state of the art by modeling the energy behavior of full-duplex energy harvesting system, and proving that the model is equivalent to Harvest-Store-Use as TI is small enough. (2) The optimal cross-layer transmission and energy scheduling algorithm is proposed to provide a theoretical design limitation. In addition, we find that if the power allocation of any RRU on a certain sub-channel at any TI is obtained, the power allocation of other RRU's can be directly obtained with their corresponding weighted CNR's as given in (15), which brings along significant operation efficiency by reducing $L-1$ times the amount of optimal power allocation calculation. (3) The numerical analysis illustrates the degradation effect to the throughput under different channel prediction errors, and shows that the novel design of the online algorithm has the performance advantage on reducing the throughput degradation caused by channel prediction errors. Interesting topics for future work include studying the cross-layer transmission and energy scheduling under full-duplex energy harvesting wireless OFDM joint transmission for differential QoS requirements.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61302108) and Strategic Pilot Project of Chinese Academy of Sciences (Grant No. XDA06010300).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Han T, Ansari N. On greening cellular networks via multicell cooperation. IEEE Wirel Commun, 2013, 20: 82–89
2 Tutuncuoglu K, Yener A. Optimum transmission policies for battery limited energy harvesting nodes. IEEE Trans Wirel Commun, 2012, 11: 1180–1189
3 Ozel O, Yang J, Ulukus S. Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery. IEEE Trans Wirel Commun, 2012, 11: 2193–2203
4 Badawy G H, Sayegh A A, Todd T D. Fair flow control in solar powered WLAN mesh networks. In: Proceedings of IEEE Wireless Communications and Networking Conference, Budapest, 2009. 1–6
5 Ng K W D, Lo E S, Schober R. Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station. IEEE Trans Wirel Commun, 2013, 12: 3412–3427
6 Huang C, Zhang R, Cui S. Throughput maximization for the gaussian relay channel with energy harvesting constraints. IEEE J Sel Area Commun, 2013, 31: 1469–1479
7 Jiang X, Polastre J, Culler D. Perpetual environmentally powered sensor networks. In: Proceedings of the 4th International Symposium on Information Processing in Sensor Networks, Los Angeles, 2005. 463–468
8 Kuo Y C, Tung W H, Liu L J. Smart integrated circuit and system design for renewable energy harvesters. IEEE J Photovoltaics, 2013, 3: 401–406
9 Lee D, Seo H, Clerckx B, et al. Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges. IEEE Commun Mag, 2012, 50: 148–155
10 Ma Z, Zhang Z Q, Ding Z G, et al. Key techniques for 5G wireless communications: network architecture, physical layer, and MAC layer perspectives. Sci China Inf Sci, 2015, 58: 041301
11 Xu J, Zhang R. CoMP meets smart grid: a new communication and energy cooperation paradigm. IEEE Tran Veh Tech, 2015, 64: 2476–2488
12 Chiang Y H, Liao W. Renewable energy aware cluster formation for CoMP transmission in green cellular networks. In: Proceedings of IEEE Global Communications Conference, Austin, 2014. 4611–4616
13 Lyman J R. Optimal mean-square prediction of the mobile-radio fading envelope. IEEE Tran Signal Process, 2003, 51: 819–824
14 Xing C, Wang N, Ni J, et al. MIMO beamforming designs with partial CSI under energy harvesting constraints. IEEE Signal Process Lett, 2013, 20: 363–366
15 Marsch P, Fettweis G P. Coordinated Multi-Point in Mobile Communications: From Theory to Practice. London: Cambridge University Press, 2011. 5–6
16 West M, Harrison J. Bayesian Forecasting and Dynamic Models. Berlin: Springer, 1997. 270–304
17 Bertsekas D. Convex Optimization Theory. Belmont: Athena Scientific, 2009. 347–364
18 Dinkelbach W. On nonlinear fractional programming. Manage Sci, 1967, 13: 492–498
19 Ozel O, Tutuncuoglu K, Yang J, et al. Transmission with energy harvesting nodes in fading wireless channels: optimal policies. IEEE J Sel Area Commun, 2011, 29: 1732–1743

Appendix A Proof of Proposition 1

When τ is small enough, the allocated power to RRU l and the amount of harvested power are constant. Then, Eqs. (1)–(3) can be revised as

$$\text{Case1}: S_{l,i+1} = \min\left\{\int_0^{\tau} P_l^{\rm H}(t) - P_l^{\rm U}(t){\rm d}t + S_{l,i}, E_{\rm C}\right\} = \min\left\{E_{l,i} - \mathcal{E}_{l,i}^{\rm con} + S_{l,i}, E_{\rm C}\right\}, \tag{A1}$$

$$\text{Case2}: S_{i+1} = \left[S_{l,i} - \int_0^{\tau} P_l^{\rm U}(t) - P_l^{\rm H}(t){\rm d}t\right]^+ = \left[S_{l,i} - \mathcal{E}_{l,i}^{\rm con} + E_{l,i}\right]^+, \tag{A2}$$

$$\text{Case3}: S_{i+1} = \min\left\{S_{l,i} + \int_0^{\tau} P_l^{\rm H}(t){\rm d}t, E_{\rm C}\right\} = \min\left\{S_{l,i} + E_{l,i}, E_{\rm C}\right\}. \tag{A3}$$

As shown in Figure A1, the format of Harvest-Storage-Usage model [2, 7] is given as follows, which only considers two half-duplex operation processes: storing the harvesting energy and powering the RRU or the load by the storage.

$$S_{l,i+1} = \min\left\{\left(S_{l,i} + E_{l,i} - \mathcal{E}_{l,i}^{\rm con}\right)^+, E_{\rm C}\right\}. \tag{A4}$$

The format of Harvest-Storage-Usage model can be extensively applied in the full-duplex energy harvesting system. Under the condition of Case 1 that $P_{l,i}^{\rm H}(t) \geqslant P_{l,i}^{\rm U}(t)$ during τ, $(S_{l,i} + E_{l,i} - \mathcal{E}_{l,i}^{\rm con})$ must be greater than 0, and thus Eq. (A4) can be revised to (A1). Under the condition of Case 2 that $P_{l,i}^{\rm H}(t) < P_{l,i}^{\rm U}(t)$ and $P_{l,i}^{\rm H}(t) + \bar{P}_{l,i}^{\rm U}(t) \geqslant P_{l,i}^{\rm U}(t)$ during τ, the energy harvester and ESU jointly power RRU l, and no energy is stored, i.e., $S_{l,i+1}$ must be less than $E_{\rm C}$. Thus, Eq. (A4) can be reformulated as (A2). Under the condition of Case 3 that $P_{l,i}^{\rm H}(t) + \bar{P}_{l,i}^{\rm U}(t) < P_{l,i}^{\rm U}(t)$ during τ, RRU l cannot be driven to work and all harvested energy is stored into ESU, i.e., $\mathcal{E}_{l,i}^{\rm con} = 0$. Thus, Eq. (A4) can be revised to (A3).

Therefore, Eq. (A4) is the combination of ESU behaviors of (A1)–(A3). Proposition 1 is proved.

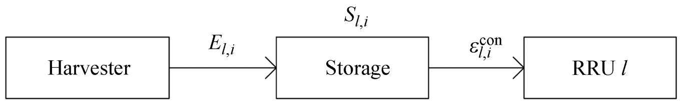

Figure A1 The Harvest-Storage-Usage process of energy harvesting enabled RRU l.

Appendix B Proof of Proposition 2

Make $D\left(\{p_{n,l,i}\}\right) = \sum_{n=1}^{N}\sum_{i=1}^{M+1} R_{n,i}(p_{n,l,i})\tau = \sum_{n=1}^{N}\sum_{i=1}^{M+1} D_{n,i}$, where

$$D_{n,i} = \log_2\left(1 + \left(\sum_{l=1}^{L}\gamma_{n,l,i}\sqrt{p_{n,l,i}}\right)^2\right)\tau = \log_2\left(1 + \sum_{l=1}^{L}\gamma_{n,l,i}^2 p_{n,l,i} + \sum_{l=1}^{L}\sum_{k\neq l}\gamma_{n,l,i}\gamma_{n,k,i}\sqrt{p_{n,l,i}p_{n,k,i}}\right)\tau. \tag{B1}$$

Since $D\left(\{p_{n,l,i}\}\right)$ is a linear combination of $D_{n,i}$, $D\left(\{p_{n,l,i}\}\right)$ is concave if $D_{n,i}$ is concave.

Make $\mathrm{G} = \log_2(\cdot)$, $f_{n,i} = \sum_{l=1}^{L}\gamma_{n,l,i}^2 p_{n,l,i}$ and $g_{n,i} = \sum_{l=1}^{L}\sum_{k\neq l}\gamma_{n,l,i}\gamma_{n,k,i}\sqrt{p_{n,l,i}p_{n,k,i}}$, and then $D_{n,i}$ can be written as $D_{n,i} = G(1 + f_{n,i} + g_{n,i})$, where $n \in \{1, 2, \ldots, N\}$ and $i \in \{1, 2, \ldots, M+1\}$, respectively. Since $\log_2(\cdot)$ is concave and non-decreasing, and $f_{n,i}$ is linear, $D_{n,i}$ is concave if $g_{n,i}$ is concave according to the composition rules of convexity [17]. The concavity of $g_{n,i}$ is verified with the first order condition [17].

Making $\boldsymbol{x}_1 = \left[p_{k,1,i}^{x_1}, p_{k,2,i}^{x_1}, \ldots, p_{k,L,i}^{x_1}\right]^{\mathrm{T}}$ and $\boldsymbol{x}_2 = \left[p_{k,1,i}^{x_2}, p_{k,2,i}^{x_2}, \ldots, p_{k,L,i}^{x_2}\right]^{\mathrm{T}}$, we have

$$\begin{aligned}
&g_{n,i}(\boldsymbol{x}_1) - g_{n,i}(\boldsymbol{x}_2) - \nabla g_{n,i}(\boldsymbol{x}_2)^{\mathrm{T}}(\boldsymbol{x}_1 - \boldsymbol{x}_2) \\
&= \sum_{k\neq l}\left[\sqrt{p_{n,l,i}^{x_1}p_{n,k,i}^{x_1}} - p_{n,l,i}^{x_1}\sqrt{\frac{p_{n,k,i}^{x_2}}{p_{n,l,i}^{x_2}}} - p_{n,l,i}^{x_1}\sqrt{\frac{p_{n,l,i}^{x_2}}{p_{n,k,i}^{x_2}}}\right]\gamma_{n,l,i}\gamma_{n,k,i} \\
&= -\sum_{k\neq l}\left[\sqrt{p_{n,l,i}^{x_1}}\left(\frac{p_{n,k,i}^{x_2}}{p_{n,l,i}^{x_2}}\right)^{\frac{1}{4}} - \sqrt{p_{n,l,i}^{x_1}}\left(\frac{p_{n,l,i}^{x_2}}{p_{n,k,i}^{x_2}}\right)^{\frac{1}{4}}\right]^2\gamma_{n,l,i}\gamma_{n,k,i} \leqslant 0.
\end{aligned} \tag{B2}$$

From (B2), we can see that $g_{n,i}(\boldsymbol{x}_1) - g_{n,i}(\boldsymbol{x}_2) \leqslant \nabla g_{n,i}(\boldsymbol{x}_2)^{\mathrm{T}}(\boldsymbol{x}_1 - \boldsymbol{x}_2)$.

Hence, $g_{n,i}$ is concave, and consequently $D\left(\{p_{n,l,i}\}\right)$ is proved to be concave.