Received August 30, 2017, accepted October 1, 2017, date of publication October 6, 2017, date of current version November 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2760378

Delay and Delay-Constrained Throughput Performance of a Wireless-Powered Communication System

ZHIDU LI[1], (Student Member, IEEE), YUMING JIANG[2], (Senior Member, IEEE), YUEHONG GAO[1], PENGXIANG LI[1], LIN SANG[1], AND DACHENG YANG[1]

[1]Wireless Theories and Technologies Laboratory, Beijing University of Posts and Telecommunications, Beijing 100876, China

[2]Department of Information Security and Communication Technology, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Corresponding author: Zhidu Li (prclzd@126.com)

This work was supported in part by the EU FP7 Marie Curie Actions under Grant 607584, in part by the EU H2020 Marie Skłodowska-Curie Actions under Grant 699924, in part by the National Natural Science Foundation of China under Grant 61300185, in part by the BUPT-SICE Excellent Graduate Students Innovation Fund, and in part by the BUPT Ph.D. Short-term Learning and Exchanging Abroad Fund.

ABSTRACT In this paper, the delay and delay-constrained throughput performance of a point-to-point wireless-powered communication system is investigated. In this system, the wireless-powered node, e.g., a user equipment (UE), receives data at the same time when powered from the other node, e.g., an access point (AP), and uses the harvested wireless energy to send data to the other node. The investigation focuses on the delay performance of sending data in the downlink (DL) from the AP node to the UE node and that in the uplink (UL) from the UE node to the AP node, based on which the throughput performance on both directions when delay constraints are enforced is also studied. To this aim, the cumulative service capacity of the service process is first analyzed for both DL and UL, taking into consideration the delay caused by the nontransmission phase for the AP or UE in each charging cycle. Thereafter, a general upper bound on the delay distribution for stochastic traffic arrivals is obtained for both DL and UL, based on which the delay-constrained throughput performance is further studied. In addition, to ensure the delay performance, the required energy storage capacity and wireless charging rate are investigated. The obtained results are exemplified with two specific traffic types, and the accuracy of the analysis is validated by comparison with extensive simulation results. The analysis and results shed new light on the performance of wireless-powered communication systems.

INDEX TERMS Wireless powered communication system, performance analysis, delay performance, delay-constrained throughput.

I. INTRODUCTION

With the advance of wireless energy transfer technology, a newly emerging topic is wireless powered communication (WPC) [1], [2], which has recently attracted growing interest from both academia and industry. WPC has a great potential for use in a wide range of applications particularly in wireless sensor networks (WSNs) and Internet of Things/Everything (IoT/IoE) [2]. In a WPC system, a node, e.g. a user equipment (UE), harvests energy and may simultaneously also receive data from the ambient radio signal that may be purposely radiated by another node, e.g. an access point (AP), in its downlink (DL) to the UE node, and the UE node may use the harvested energy to transmit data in the uplink (UL) to the AP node, as shown in Fig. 1.

Typically, in a WPC system, the UL and the DL share the same frequency band, which implies that the system works in half-duplex mode. A fundamental issue of the WPC system is to decide how frequently the energy transfer should be conducted and how long each time the energy transfer should last. To provide answers to these questions, one has to investigate how much data needs to be sent by the AP node in the DL and how much by the UE node in the UL, or equivalently what data throughput or capacity the system is intended to achieve for the AP and UE respectively. In addition to throughput, delay guarantee is sometimes also required by delay-critical IoT devices [3]- [5]. For instance, in a traffic control system, the monitor sensors should upload the traffic information through high resolution images to the traffic control center within a small latency, especially when the traffic is heavy. As a consequence, if there is delay requirement on the data, the investigation of a WPC system should also take it into account. These constitute the objective of this paper.

In the literature, a number of studies on the throughput performance of WPC systems can be found. Usually, these studies only focus on either UL or DL transmission. *On UL transmission*, Ju and Zhang [6] studied two types of throughput under a harvest-then-transmit protocol, which are the maximum system capacity and the throughput guaranteed for all devices at the same time. In [7], the focus was on spatial UL throughput maximization of a WPC network, by finding the optimal tradeoff between the energy transfer and information transfer. In [8], an optimization algorithm was proposed to maximize the system UL throughput in a multiuser multi-input-multi-output (MIMO) system through jointly optimizing the energy beamforming, receive beamforming and time slot allocation. In [9], under the requirement of a minimum UL throughput, the focus was on the energy efficiency through performing power control and time allocation jointly. *On DL transmission*, the studies usually take the technology of simultaneous wireless information and power transfer (SWIPT) into account. In [10], the relationship between the channel capacity and energy harvesting rate was studied for different power splitting schemes. It was found that the receiver architecture should be optimized for wireless power transfer since wireless information transmission operated with quite low power sensitivity. In [11], the rate-energy region was further studying for a MIMO broadcasting channel under the nonlinear radio frequency energy harvesting model. Yan and Liu [12] investigated opportunistic relaying and studied the maximum capacity of the secondary transmitter in a cooperative cognitive radio network. Differently, work [13] studied direct relaying and proposed a relay selection scheme to minimize the outage probability. In summary, all these existing works [6]- [13] focus mostly on throughput studies with fluid traffic, where, delay, though a key performance metric, receives little attention.

To the best of our knowledge, the state-of-the-art study of delay performance in WPC systems is rather limited. Weakly related, in [14], a method to control the power-delay performance on demand in a WPC system was proposed to minimize the time-averaged power consumption. In addition, work [15] proposed an adaptive harvest-then-cooperate protocol to minimize the average delay of UE by simulation method. However, the aims of [14] and [15] are both with little touch on maximizing the throughput or capacity performance as in [6]–[13]. Due to the limitation of hardware and energy transfer loss, the amount of harvested energy in a WPC system may be highly limited compared with the conventional systems powered through circuit [16]. Consequently, it is crucial to *study the delay and throughput performance together* for WPC systems. In our early work [17], we focused on the scenario with static channel, and the delay and delay-constrained throughput performance was analyzed for the UL transmission. In brief, a comprehensive, joint delay and throughput performance study for both DL and UL with the more general setting of stochastic channel is absent.

In this paper, we investigate the delay and delay-constrained throughput performance of a point-to-point WPC system. Specifically, our focus is on the delay performance of sending data in the DL from the AP to UE and that in the UL from the UE to the AP, and on their maximum throughput performance when delay constraints are enforced. In our investigation, finite energy storage capacity and general stochastic channel are taken into account. To this aim, the cumulative service capacity of the service process is first analyzed for both DL and UL, taking into consideration the delay caused by the non-transmission phase for the AP or UE in each charging cycle. Thereafter, two general upper bounds on the delay distribution for stochastic traffic arrivals are obtained for both DL and UL: While one is tight with high calculation complexity, the other may be generally not as tight but is much simple in expression and for small charging cycle cases is also accurate. Based on the delay analysis, the delay-constrained throughput performance is further studied. In addition, the minimum energy storage capacity and the minimum DL transmission power to ensure the UL delay performance are also studied. The obtained results are exemplified with two specific traffic types. Moreover, the accuracy of the analysis is validated by comparison with extensive simulation results.

The contributions of this paper are as follows. (1) A tractable framework is presented to study the delay and throughput performance together for both DL and UL transmissions in a WPC system, where finite energy storage capacity, stochastic channel and stochastic traffic arrivals are all taken into account. (2) More specifically, two useful upper bounds on delay distribution for both DL and UL transmissions are derived. The tight bound is verified to be accurate for all tested settings while the simplified bound is reasonably good and can hence be used to simplify the analysis when the charging cycle is small. It is worth highlighting that even the simplified bound for the UL transmission is more accurate than the bound obtained in [17] with identical conditions. (3) The delay-constrained throughput performance is studied, together with the minimum energy storage capacity and the minimum DL transmission power to ensure the UL delay performance. (4) The analysis and the simulation results explicitly reveal the impact of traffic and the impact of time allocation between UL and DL on the throughput performance when different delay requirements are enforced. They shed new light on the performance of WPC systems.

The remainder is organized as follows. In Section II, the system model is presented. In Section III, general analysis of the WPC system is conducted. In Section IV, the obtained results are exemplified with two specific traffic types, namely constant fluid traffic and stochastic traffic. In Section V, analytical and simulation results are presented, compared and discussed. Finally, we conclude the paper in Section VI.

II. SYSTEM MODEL

A. NOTATION

Throughout this paper, the following notation is adopted. The cumulative amount of arrival traffic and that of the data transmission capacity of the AP or the UE are both expressed in the form of $Y(t)$ during time $[0, t)$ and $Y(s, t)$ during time $[s, t)$ respectively, i.e., $Y(s, t) = Y(t) - Y(s)$. A variable with subscript 'D' means it is used in the DL while one with subscript 'U' means it is used in the UL. For a variable x, function $[x]^+$ means $\max\{x, 0\}$. Besides, $\lceil x \rceil$ and $\lfloor x \rfloor$ denote the ceiling function and the floor function of x respectively.

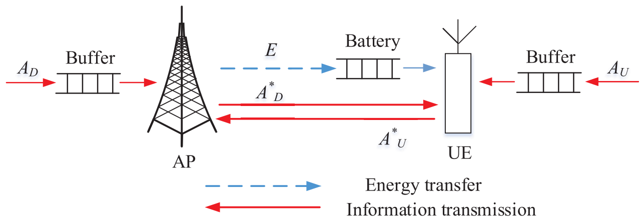


FIGURE 1. System model with wireless energy transfer in the downlink and wireless information transmission in both uplink and downlink.

B. SYSTEM MODEL

As shown in Fig. 1, we consider a point to point wireless powered communication system with an AP and a UE. It is assumed that the AP and UE are equipped with one single antenna each. In the DL, the AP transmits radio signal to the UE. The energy from the received signal is divided into two parts by the UE, where one part is used to recover the information and the other part is stored into a battery. In the UL, the stored energy is used to send data from the UE to the AP. The DL and UL are assumed to share the same frequency band, i.e., the system works in half-duplex mode.

We focus on the data transmission performance of the AP in the DL and that of the UE in the UL. The data traffic may arrive at the AP or the UE randomly and is FIFO-served. Both AP and UE are equipped with a buffer to store the data that cannot be served immediately. The buffer capacity of the AP and that of the UE are assumed to be sufficiently large such that no buffer overflow would happen.

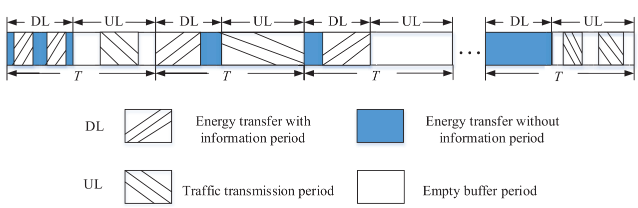


FIGURE 2. The time (allocation) model.

The time model consists of multiple consecutive time blocks (TBs) which are indexed by $1, 2, \cdots$. Each TB consists of a DL phase and a UL phase. The duration of each TB is fixed as T time units. The system adopts a DL-then-UL transmission protocol, as depicted in Fig. 2. The time allocation for the DL and UL in a TB is determined by a parameter τ $(0 \leq \tau \leq 1)$. Specifically, in each TB, the first τT time units are assigned to the AP to transfer wireless energy and possibly also data to the UE in the DL. A fixed amount of the harvested energy is used by the UE to recover the information while the remaining energy is stored into the battery to support data transmission in the UL. Thereafter, the remaining $(1 - \tau)T$ time units are assigned to the UE to send data to the AP as long as the UE has data to send. As used in the literature, we assume that the complex flat-fading channel gain denoted by $\tilde{h}$ is invariant within a TB while it is independent and identically distributed (i.i.d) over different TBs.

C. ENERGY HARVESTING AND DATA TRANSMISSION

The transmission power of the AP, denoted by p_A, holds as $p_A = \mathbb{E}[|x_A|^2]$, where $\mathbb{E}[\cdot]$ is the expectation function and x_A is the radio signal sent by the AP. During the DL phase of the ith TB, the transmission rate of the AP holds as

$$R_{D,i} = W \log_2(1 + \frac{p_A h_i}{N_0 W}), \tag{1}$$

where N_0 denotes the power spectral density of background noise and W denotes the bandwidth. Besides, h_i is the power gain in the ith TB and $h_i = |\tilde{h}_i|^2$ where $\tilde{h}_i$ is the corresponding channel gain. Here, the UE is assumed to be close to the AP such that the path loss is neglected.

It is assumed that p_A is constant and sufficiently large, such that the wireless energy harvested from the background noise and the energy loss due to DL information recovery can be negligible. Note that this assumption is reasonable since the power level of the background noise and that of information recovery are both lower than 10^{-2} milliwatt while p_A is usually higher than 10 milliwatt [10]. Thus, the amount of energy harvested by the UE in the ith TB can be expressed as

$$P_i = \zeta p_A h_i \tau T, \tag{2}$$

where $\zeta \in (0, 1]$ is the energy harvesting efficiency.

Further, since the traffic may arrive randomly over time, the energy may not be depleted at the end of some TBs. Let b denote the energy storage capacity of the battery and E_i denote the amount of energy left at the beginning of the ith TB. There holds

$$E_{i+1} = \min\{E_i + P_i, b\} - p_i u_i, \tag{3}$$

where p_i and u_i denote the transmission power of the UE and the total UL transmission time in the ith TB respectively. Here, we assume, at the UE, its harvested energy is mainly consumed by its data transmission, ignoring the other part of its functionalities. Additionally, the transmission power of the UE in the ith TB is determined as

$$p_i = \frac{\min\{E_i + P_i, b\}}{(1 - \tau)T}. \tag{4}$$

The intuition behind the power allocation scheme to the UE is to maximize the transmission power to use up the remaining energy and the harvested energy at the end of each TB, such that the data transmission capacity in each TB is maximized.

During the UL phase of the ith TB, the transmission rate of the UE holds as

$$R_{U,i} = W \log_2(1 + \frac{p_i h_i}{N_0 W}). \quad (5)$$

D. PERFORMANCE METRICS OF INTEREST

For ease of defining variables for the DL and UL transmissions simultaneously, we use subscript 'X' to denote the subscript of DL 'D' or that of UL 'U'. We assume the amount of traffic arrivals to the AP is i.i.d over time and so is to the UE. We use $A_X(s,t)$ to denote the cumulative amount of the traffic arrivals during time $[s,t)$ and use $A_X(t)$ to represent $A_X(0,t)$ for short as notated in Sec-II-A. Then, $A_X(t)$ has independent and stationary increments [18], i.e.,

- For $\forall 0 \le s \le t$, $\forall e \ge 0$, $A_X(s,t)$ and $A_X(s+e, t+e)$. have identical distribution.
- For $\forall 0 \le t_1 \le \cdots \le t_n$, $A_X(t_1,t_2), \cdots, A_X(t_{n-1},t_n)$ are independent of each other.

The corresponding departure process of $A_X(t)$ is denoted by $A_X^*(t)$. It is easily verified that, for a system with input $A_X(t)$ and output $A_X^*(t)$, there holds [19]

$$A_X^*(t) = \inf_{0 \le s \le t} \{A_X(s) + C_X(s,t)\}, \quad (6)$$

where $C_X(s,t)$ represents the cumulative service capacity within time $[s,t)$.

The objective of this paper is to study the delay and delay-constrained throughput performance for sending data from the AP to UE in the DL and that from the UE to AP in the UL. The delay $D_X(t)$ is defined as [19]

$$D_X(t) = \inf\{d_X : A_X(t) \le A_X^*(t + d_X)\}. \quad (7)$$

The *delay-constrained throughput* is defined as the maximum traffic rate that the system can sustain to meet the delay constraint [20]:

$$r_X^{\max} = \sup\{r_X : Pr\{D_X(t) > d_X\} \le \epsilon_X\}, \quad (8)$$

where r_X denotes the traffic arrival rate, and $Pr\{D_X(t) > d_X\} \le \epsilon_X$ denotes the delay constraint which means the probability for violating a given delay threshold d_X should be controlled within ϵ_X. Note that the delay constraints for the DL and UL may be different and may impact with each other through the time allocation parameter τ.

E. REMARKS

In the system model described above, several assumptions have been made in order to simplify the analysis and the representation of the analytical results. They include: the harvested energy is mainly consumed by data transmission, the Shannon capacity is used as the (maximum achievable) data transmission rate where there is no data loss, and the energy harvesting efficiency is constant. Note that many of these assumptions have been commonly adopted in the literature (see [6]– [9]). In addition, though these assumptions have led to simplified results, the fundamental essence of the analytical approach is not affected. In other words, the analysis can be extended when these assumptions are relaxed. Furthermore, the most appealing phenomenon, i.e. the delay-constrained throughput may significantly differ from the throughput without delay consideration, is sufficiently revealed by the obtained results, as later clearly seen from Fig. 4.

III. PERFORMANCE ANALYSIS

A. SYSTEM SERVICE CHARACTERIZATION

For the AP or UE, a TB always consists of a transmission phase and a non-transmission phase. The latter part will result in packet backlog and packet delay if there are packets requesting to be served. In general, the packet backlog is usually aggravated in the non-transmission phase while it is alleviated in the transmission phase for a long-term data transmission since the average service rate should be higher than the traffic arrival rate for a stable system. In what follows, we study the cumulative service capacity $C_X(t)$ with consideration of stochastic traffic arrivals.

Without losing generality in a long-term transmission of DL, we focus on a TB which satisfies two points: 1) before the non-transmission phase of this TB, the backlogged packets have all been served; 2) the backlog occurs during the non-transmission phase of this TB. We let the backlog start time as time 0. Note that the backlog start time can be anytime during the non-transmission period since the traffic arrival process is stochastic. Let $(1-\frac{\sigma_D}{n}(1-\tau))T$ denotes the backlog start time from the beginning of this TB, where $\sigma_D \in \mathbf{Z}$, $n \in \mathbf{Z}^+$, $\mathbf{Z}$ and $\mathbf{Z}^+$ are the set of integers and that of positive integers respectively, with σ_D uniformly distributed within $[0,n]$. Based on the new start time, the cumulative data transmission capacity from time 0 to time t holds as

$$C_D(t) = \sum_{i=1}^{\lfloor \frac{t}{T} \rfloor} R_{D,i}\tau T + R_{D,\lceil \frac{t}{T} \rceil} \min\{(t - \lfloor \frac{t}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+, \tau T\}. \quad (9)$$

Similarly, in the UL, let $(1 - \frac{\sigma_U}{n})\tau T$ denotes the backlog start time from the beginning of this TB, where $\sigma_U \in \mathbf{Z}$, $n \in \mathbf{Z}^+$, and σ_U is uniformly distributed within $[0,n]$. Based on the new start time, the cumulative data transmission capacity from time 0 to time t holds as

$$C_U(t) = \sum_{i=1}^{\lfloor \frac{t}{T} \rfloor} R_{U,i}(1-\tau)T + R_{U,\lceil \frac{t}{T} \rceil} \min\{(t - \lfloor \frac{t}{T} \rfloor T - \frac{\sigma_U}{n}\tau T)^+, (1-\tau)T\}. \quad (10)$$

Note that the expression of $C_X(t)$ also applies to the case of deterministic traffic arrivals by choosing a suitable σ_X to characterize the maximum delay caused due to the existence of non-transmission phase for the AP or the UE.

B. DELAY PERFORMANCE

The following lemma provides a general expression for the delay violation probability to meet a given delay requirement.

Lemma 1: Consider a stable WPC system as depicted in Fig. 1, where the data transmission capacity of the node is characterized as $C_X(t)$ and the traffic arrival process is as $A_X(t)$. If $C_X(t)$ and $A_X(t)$ are independent and both have independent and stationary increments, then for a given delay requirement d_X, the corresponding delay violation probability is bounded by

$$Pr\{D_X(t) > d_X\} \leq \mathbb{E}[e^{-\theta_X C_X(d_X)}],$$

for some $\theta_X > 0$ which meets $\mathbb{E}[e^{\theta_X A_X(1)}]\mathbb{E}[e^{-\theta_X C_X(1)}] \leq 1$.

Proof: Please see Appendix A. ■

In Lemma 1, the condition $\mathbb{E}[e^{\theta_X A_X(1)}]\mathbb{E}[e^{-\theta_X C_X(1)}] \leq 1$ implies a sufficient stability condition for a system [19]

$$\frac{\log \mathbb{E}[e^{\theta_X A_X(1)}]}{\theta_X} \leq -\frac{\log \mathbb{E}[e^{-\theta_X C_X(1)}]}{\theta_X}$$
$$\Updownarrow$$
$$\lim_{t\to\infty}\frac{\log \mathbb{E}[e^{\theta_X A_X(t)}]}{\theta_X t} \leq -\lim_{t\to\infty}\frac{\log \mathbb{E}[e^{-\theta_X C_X(t)}]}{\theta_X t}, \quad (11)$$

where $\frac{\log \mathbb{E}[e^{\theta_X A_X(t)}]}{\theta_X}$ and $-\frac{\log \mathbb{E}[e^{-\theta_X C_X(t)}]}{\theta_X}$ are the statistical envelops of processes $A_X(t)$ and $C_X(t)$ respectively [19], [21]. And, $\alpha_{\theta_X} \triangleq \frac{\log \mathbb{E}[e^{\theta_X A_X(1)}]}{\theta_X}$ and $\beta_{\theta_X} \triangleq -\frac{\log \mathbb{E}[e^{-\theta_X C_X(1)}]}{\theta_X}$ denote the corresponding envelop rates respectively.

According to (9) and (10), for a time varying channel, $C_X(t)$ is i.i.d over different TBs but it may not be identically distributed over time. However, the length of a TB (i.e., T) should be small for a practical system since otherwise the packet delay which is caused due to the existence of non-transmission phase would be high. Thus, the cumulative transmission capacity $C_X(t)$ can be approximately considered to be i.i.d [18]. And we let $\beta_{\theta_X} \approx \lim_{t\to\infty} -\frac{\log \mathbb{E}[e^{-\theta_X C_X(t)}]}{\theta_X t}$ in this paper, i.e., there holds

$$\beta_{\theta_D} \approx -\frac{\log \mathbb{E}[e^{-\theta_D R_D \tau}]}{\theta_D}$$
$$\beta_{\theta_U} \approx -\frac{\log \mathbb{E}[e^{-\theta_U R_U (1-\tau)}]}{\theta_U}. \quad (12)$$

We highlight that later in Section V, this approximation is validated to be reasonable by comparing with the simulation results even when T is set to be as large as 10 time units.

Additionally, according to Lemma 1, the delay violation probability bound decreases as θ_X increases, which means a tighter bound would be achieved with a larger θ_X. However, θ_X is constrained by (11), i.e.,

$$\alpha_{\theta_X} \leq \beta_{\theta_X} \quad (13)$$

for both DL and UL data transmissions. Hence, the optimal θ_X can be found out according to the following expression,

$$\theta_X^{opt} = \max\{\theta_X : \alpha_{\theta_X} \leq \beta_{\theta_X}\}. \quad (14)$$

With Lemma 1, the delay performance of the DL and that of the UL can be analyzed specifically. The following theorems summarize the obtained bounds.

Theorem 1: For the DL transmission of the AP, the delay violation probability is upper bounded by

$$Pr\{D_D(t) > d_D\} \leq \lim_{n\to\infty}\frac{\sum_{\sigma_D=0}^{n}\mathbb{E}[e^{-\theta_D R_D \tau (d_D - \frac{\sigma_D}{n}(1-\tau)T)^+}]}{n+1}$$

and a simplified bound is as:

$$Pr\{D_D(t) > d_D\} \leq \mathbb{E}[e^{-\theta_D R_D \tau (d_D-(1-\tau)T)^+}].$$

Proof: Please see Appendix B. ■

Theorem 2: For the UL transmission of the UE, the delay violation probability is upper bounded by

$$Pr\{D_U(t) > d_U\} \leq \lim_{n\to\infty}\frac{\sum_{\sigma_U=0}^{n}\mathbb{E}[e^{-\theta_U R_U (1-\tau)(d_U - \frac{\sigma_U}{n}\tau T)^+}]}{n+1}$$

and a simplified bound is as:

$$Pr\{D_U(t) > d_U\} \leq \mathbb{E}[e^{-\theta_U R_U (1-\tau)(d_U-\tau T)^+}]$$

Proof: Please see Appendix C. ■

C. DELAY-CONSTRAINED THROUGHPUT

With the information of the delay constraint including the delay requirement d_X and the delay violation probability ϵ_X, the maximum sustained throughput of a traffic type can be derived. According to Theorems 1 and 2, the delay violation probability is actually a function of θ_X, which can be denoted as $f_X(\theta_X) = \epsilon_X$. If the delay constraint is given, the corresponding parameter θ_X holds as

$$\theta_X = f_X^{-1}(\epsilon_X), \quad (15)$$

where $f_X^{-1}(\cdot)$ is the inverse function of $f_X(\cdot)$.

Replacing the obtained θ_X into (13), the maximum traffic envelop rate denoted by $\alpha_{\theta_X}^{\max}$ is obtained. Furthermore, α_{θ_X} is a function in traffic arrival rate r_X, which can be denoted as $g(r_X) = \alpha_{\theta_X}$ [19]. Therefore, the maximum sustained throughput to meet the given delay constraint can be represented as:

$$r_X^{\max} = g^{-1}(\alpha_{\theta_X}^{\max}), \quad (16)$$

where $g^{-1}(\cdot)$ is the inverse function of $g(\cdot)$.

D. ENERGY STORAGE CAPACITY AND DL TRANSMISSION POWER

While small energy storage capacity restricts the UL performance, large capacity may lead to high economic cost. It is consequently worth finding out an energy storage capacity as small as possible to ensure the UL performance requirement.

Suppose the traffic arrival process is characterized by envelop rate α_{θ_U} and the delay constraint is given as (d_U, ϵ_U). According to (12) and Theorem 2, the envelop rate of C_U and the delay violation probability are both functions related to the UL transmission rate R_U and the parameter θ_U, which can be denoted by $\beta_{\theta_U}(R_U, \theta_U)$ and $f(R_U, \theta_U)$ respectively.

Besides, the amount of the energy stored before the UL transmission is upper bounded by the energy storage capacity b according to (3). Thus, further according to expression (5) and the stability condition (13), there holds

$$\begin{aligned} \beta_{\theta_U}(R_U, \theta_U) &\geq \alpha_{\theta_U} \\ f(R_U, \theta_U) &= \epsilon_U \\ W \log_2(1 + \frac{bh}{N_0 W}) &\geq R_U, \end{aligned} \tag{17}$$

Solving equation set (17) with expression (12) and Theorem 2, the minimum required energy storage capacity $b^{\min}$ and the corresponding θ_U are obtained.

When the energy storage capacity is sufficiently large, the energy stored before the UL transmission in a TB, which is $E_i + P_i$, according to (3) can be simplified as

$$E_i + P_i = E_i + \zeta p_A h_i \tau T.$$

According to (2), the amount of harvested energy depends on the DL transmission power p_A and the time allocation parameter τ. For guaranteeing the UL performance, it is better to adjust p_A rather than τ since τ also has effect on the DL performance. Therefore, we can find the required minimum DL transmission power $p_A^{\min}$ to ensure the UL performance in average, i.e.,

$$\begin{aligned} \mathbb{E}[E_i + \zeta p_A h_i \tau T] &\geq b^{\min} \\ &\Updownarrow \\ p_A &\geq \frac{b^{\min} - \mathbb{E}[E]}{\zeta \tau T \mathbb{E}[h]} \\ &\Downarrow \\ p_A^{\min} &= \frac{b^{\min}}{\zeta \tau T \mathbb{E}[h]}. \end{aligned} \tag{18}$$

Here, we have removed E from the expression, i.e., let $E = 0$. The reason is that the probability distribution of E is generally difficult to obtain from (2), (3) and (5). In addition, lessening E to 0 gives a conservative estimation on both $p_A^{\min}$ and the UL performance. Moreover, the analytical results are still accurately enough while comparing with the simulation results to be shown in Section V.

IV. EXAMPLES

As indicated in Theorems 1 and 2, the delay performance depends on both the cumulative capacity of the DL and UL and the traffic characteristics. In order to further understand the analysis and the results, we exemplify with two specific traffic types. One is constant fluid traffic and the other is a stochastic traffic type, as defined in the following.

Traffic 1 (Fluid Traffic): The packet size is infinitesimal and the arrival rate is fixed as $r_{X,1} = \lambda_X L_X$. In this case, the statistical envelop rate of Traffic 1 holds as

$$\alpha_{\theta_X,1} = r_{X,1}. \tag{19}$$

Traffic 2 (Stochastic Traffic): The arrival interval between two packets are exponentially distributed with parameter $1/\lambda_X$. The packet size L_X is assumed to be constant. The statistical envelop rate, as introduced in (11), of Traffic 2 can be calculate as

$$\alpha_{\theta_X,2} = \frac{\lambda_X}{\theta_X}(e^{\theta_X L_X} - 1). \tag{20}$$

In addition, for ease of expression, we present below the simplified delay bounds obtained in Theorem 1 and Theorem 2 and assume constant power gain that implies constant data transmission rate R_X. The analysis with time-varying channel and the tight bound can also be performed through a similar way but without closed form expression of the results.

A. ANALYSIS FOR FLUID TRAFFIC

For Traffic 1, the arrival process and service process are both deterministic, based on which, the delay is always deterministically bounded as

$$\begin{aligned} D_{D,1}(t) &\leq (1-\tau)T \\ D_{U,1}(t) &\leq \tau T, \end{aligned}$$

on the condition that $r_{D,1} \leq R_D \tau$ and $r_{U,1} \leq R_U(1-\tau)$ for system stability.

According to (13), the delay-constrained throughput of Traffic 1 holds as

$$\begin{aligned} r_{D,1}^{\max} &= R_D \tau \\ r_{U,1}^{\max} &= R_U(1-\tau), \end{aligned} \tag{21}$$

as long as the delay requirement $d_D \geq (1-\tau)T$ and $d_U \geq \tau T$. Interestingly, if the requirement is $d_D < (1-\tau)T$ and $d_U < \tau T$, then the delay-constrained throughput is zero, i.e., the requirement cannot be met.

In addition, according to (12) and (19), the equation set (17) holds as

$$R_U(1-\tau) \geq \lambda_U L_U,$$

i.e., we have $R_{U,1}^{\min} = \frac{\lambda_U L_U}{1-\tau}$. Thus, the minimum energy storage capacity needed is:

$$b_1^{\min} = \frac{(2^{\frac{R_{U,1}^{\min}}{W}} - 1) N_0 W (1-\tau) T}{h}. \tag{22}$$

Furthermore, the required minimum DL transmission power is obtained according to (18), i.e.,

$$p_{A,1}^{\min} = \frac{b_1^{\min}}{\zeta \tau T h}. \tag{23}$$

B. ANALYSIS FOR STOCHASTIC TRAFFIC

Applying Theorem 1 and Theorem 2, the delay violation probability for Traffic 2 holds as

$$\begin{aligned} Pr\{D_{D,2}(t) > d_D\} &\leq e^{-\theta_D^{opt} R_D \tau (d_D - (1-\tau)T)} \\ Pr\{D_{U,2}(t) > d_U\} &\leq e^{-\theta_U^{opt} R_U (1-\tau)(d_U - \tau T)}, \end{aligned} \tag{24}$$

where $\theta_D^{opt} = \max\{\theta_D : \frac{\lambda_D}{\theta_D}(e^{\theta_D L_D} - 1) \leq R_D \tau\}$ and $\theta_U^{opt} = \max\{\theta_U : \frac{\lambda_U}{\theta_U}(e^{\theta_U L_U} - 1) \leq R_U(1-\tau)\}$ according to (14).

Conversely, for a given delay violation probability ϵ_X, the corresponding delay is bounded as

$$D_{D,2}(t) \leq -\frac{\ln \epsilon_D}{\theta_D^{opt} R_D \tau} + (1-\tau)T$$
$$D_{U,2}(t) \leq -\frac{\ln \epsilon_U}{\theta_U^{opt} R_U (1-\tau)} + \tau T. \tag{25}$$

In what follows, we derive the delay-constrained throughput under a given delay constraint which is represented by two parameters: the delay requirement d_X, and the violation probability ϵ_X.

Firstly, we ascertain parameter θ_X according to (15) and (24):

$$\theta_D = -\frac{\ln \epsilon_D}{R_D \tau (d_D - (1-\tau)T)}$$
$$\theta_U = -\frac{\ln \epsilon_U}{R_U (1-\tau)(d_D - \tau T)}.$$

Then, combining (13), (16) and (20), the delay-constrained throughputs become:

$$r_{D,2}^{\max} = \lambda_D^{\max} L_D = \frac{R_D \tau \theta_D L_D}{e^{\theta_D L_D} - 1}$$
$$r_{U,2}^{\max} = \lambda_U^{\max} L_U = \frac{R_U (1-\tau) \theta_U L_U}{e^{\theta_U L_U} - 1}. \tag{26}$$

Next, when the traffic throughput and the delay constraint in the UL are enforced to the system, the required minimum energy storage capacity and minimum DL transmission power are derived in the following.

According to (12), (20) and (24), the equation set (17) now holds as

$$R_U(1-\tau) \geq \frac{\lambda_U}{\theta_U}(e^{\theta_U L_U} - 1)$$
$$e^{-\theta_U R_U (1-\tau)(d_U - \tau T)} = \epsilon_U$$
$$W \log_2(1 + \frac{bh}{N_0 W}) \geq R_U$$

with which, we have

$$R_{U,2}^{\min} = \frac{\lambda_U}{\theta_U (1-\tau)}(e^{\theta_U L_U} - 1),$$

where $\theta_U = \ln(1 - \frac{\ln \epsilon_U}{\lambda_U (d_U - \tau T)})/L_U$. Then, the energy storage capacity to sustain $R_{U,2}^{\min}$ can be calculated:

$$b \geq \frac{(2^{\frac{R_{U,2}^{\min}}{W}} - 1) N_0 W (1-\tau) T}{h}.$$

Thus, the minimum energy storage capacity needed becomes

$$b_2^{\min} = \frac{(2^{\frac{R_{U,2}^{\min}}{W}} - 1) N_0 W (1-\tau) T}{h}. \tag{27}$$

Furthermore, the required minimum DL transmission power is obtained according to (18), i.e.,

$$p_{A,2}^{\min} = \frac{b_2^{\min}}{\zeta \tau T h}. \tag{28}$$

C. DISCUSSION

It is worth highlighting that the performance for stochastic traffic is different from that for fluid traffic according to their analytical expressions. However, it is easily verified that the performance of the stochastic traffic converges to that of the fluid traffic by loosening the delay constraint, i.e.,

$$\lim_{\epsilon_X \to 1 \text{ and } d_X \to \infty} r_{X,2}^{\max} = r_{X,1}^{\max}$$
$$\lim_{\epsilon_U \to 1 \text{ and } d_U \to \infty} b_2^{\min} = b_1^{\min}$$
$$\lim_{\epsilon_U \to 1 \text{ and } d_U \to \infty} p_{A,2}^{\min} = p_{A,1}^{\min}. \tag{29}$$

Note that $r_{X,1}^{\max}$ is equivalent to the mean channel capacity according to (21), which implies that the throughput of the stochastic traffic also converges to the mean channel capacity.

V. RESULTS

In this section, we present numerical results from the analysis and compare with simulation results to discuss the performance of the WPC system.

If not otherwise highlighted, the various involved parameters and the adopted analysis scenarios are as follows. We set the bandwidth $W = 100$kHz, the power spectral density of the background noise $N_0 = -160$dBm, the transmission power of AP $p_A = 0.01$J per time unit, the duration of a TB $T = 1$ time unit and the energy harvesting efficiency ζ=1. The energy storage capacity is assumed to be always sufficient, i.e., no energy overflow would happen. Additionally, in order to study the performance for the stochastic traffic which is served on stochastic channel, we employ Traffic 2 with mean arrival rate $\lambda = 10$ packets per time unit and constant packet size $L = 100$kbits and Rayleigh fading channel with mean channel power gain $h = 1$ to perform analysis. The analytical results are based on the tight bounds in Theorems 1 and 2.

Fig. 3 depicts the delay violation probability varying with the delay requirement for the DL and UL transmissions respectively. Static channel and stochastic channel are both considered. Fig. 3 shows that the delay performance of data transmission on static channel is better than that on stochastic channel. In addition, a longer TB block leads to a higher delay since the duration of the no-transmission phase of a TB tends to be longer. Comparing Fig. 3(a) with Fig. 3(b), it can be found that the delay performance of the DL is better than that of the UL especially for stochastic channel even though the duration of the transmission period is identical (e.g., 0.9 time units per TB) for both DL and UL. This is due to that the mean transmission capacity of the UE is lower than that of the AP for the stochastic channel, which will be further confirmed in Figs. 4 – 6 which depict the throughput performance in detail.

In addition, the accuracy of the analytical results is also checked with simulation results in Fig. 3. Fig. 3 indicates that the analytical results are accurate in both DL and UL. In addition, the accuracy increases as T decreases. Note that the phenomenon (i.e., the dash lines) seems to be opposite visually when $d = 2.5$ time units. Actually, the gap is

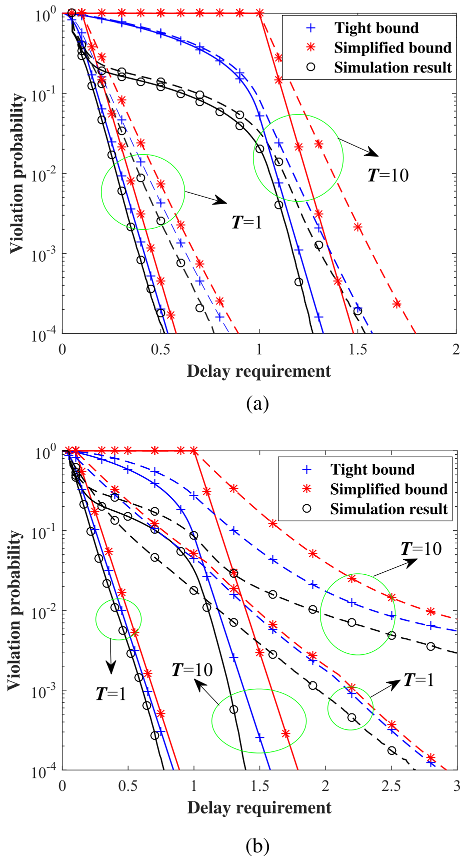


FIGURE 3. Delay violation probability varying with delay requirement (time units) for different configurations of T (time units), where the solid lines represent the static channel while the dashed lines represent the stochastic channel. (a) DL, $\tau = 0.9$. (b) UL, $\tau = 0.1$.

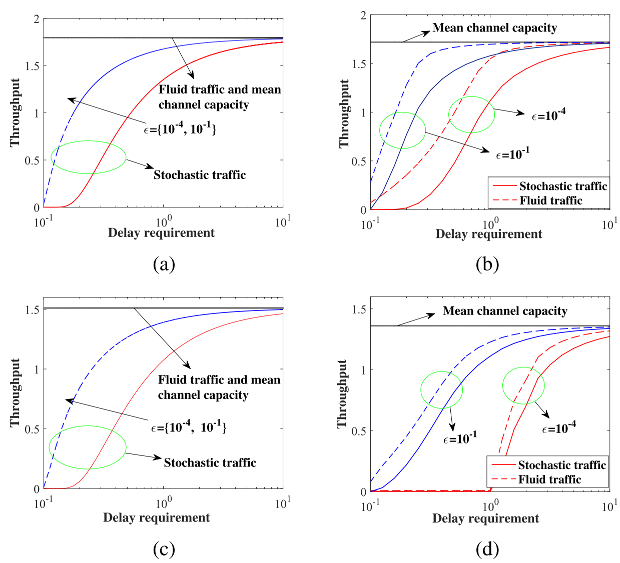


FIGURE 4. Delay-constrained throughput (Mbits per time unit) varying with delay requirement (time units) for the DL and UL transmissions. (a) DL, $\tau = 0.9$, static channel. (b) DL, $\tau = 0.9$, stochastic channel. (c) UL, $\tau = 0.1$, static channel. (d) UL, $\tau = 0.1$, stochastic channel.

still smaller for the case where $T = 1$, since the order of magnitude for $T = 1$ is 10^{-3} while that for $T = 10$ is 10^{-2}. As expected, the simplified analytical bound is looser than the corresponding tight analytical bound. However, the gap between the tight bound and the simplified bound decreases as T decreases. Therefore, when T is small enough (i.e., $T = 1$ time unit), using the simplified bound may be a good compromise between accuracy and computation complexity.

In Fig. 4, the delay-constrained throughput performance is depicted. Fig. 4 shows that the maximum sustained throughput of the stochastic traffic is different from that of the fluid traffic. Specifically, in Fig. 4(a) and Fig. 4(c), the delay-constrained throughput of the fluid traffic is always equal to the static channel capacity. Differently, the maximum sustained throughput of the stochastic traffic is highly affected by the delay constraint (i.e., the delay requirement and violation probability) and is always lower than that of the fluid traffic. Similar phenomenon can also be observed in Fig. 4(b) and Fig. 4(d) under the stochastic channel scenario. The only difference is that the maximum throughput of the fluid traffic also depends on the delay constraint under the stochastic channel scenario. Implied by Fig. 4, if delay constraint is required in data transmission, using the mean channel capacity or the capacity (i.e., throughput) found from fluid traffic model would both easily lead to overestimation due to obvious gaps among these capacities. Nevertheless, when the delay constraint is loosened sufficiently (e.g., $d = 10$ time units, $\epsilon = 0.1$), the delay-constrained throughput even with stochastic traffic converges to the mean channel capacity, as having been discussed in (29).

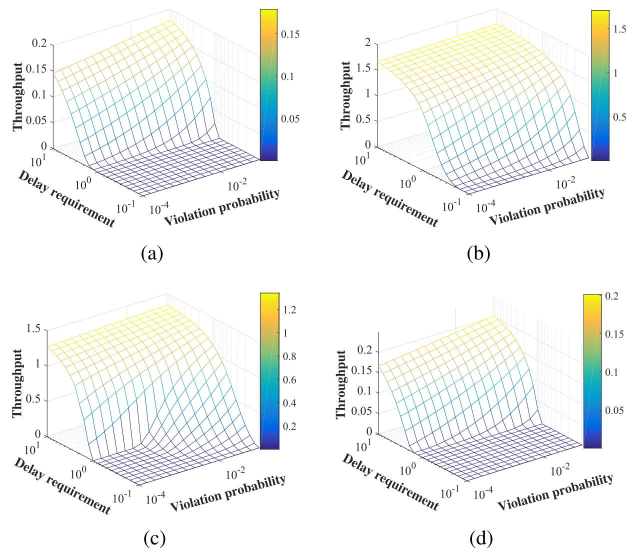


FIGURE 5. Delay-constrained throughput (Mbits per time unit) for stochastic traffic varying with delay requirement (time units) and violation probability. (a) DL, $\tau = 0.1$. (b) DL, $\tau = 0.9$. (c) UL, $\tau = 0.1$. (d) UL, $\tau = 0.9$.

To further show the impact of delay constraint on the maximum sustained throughput, Fig. 5 is presented. Fig. 5 shows that the looser the delay constraint, which is the higher delay or the larger delay violation probability the system can tolerate, the higher the throughput that the traffic can acquire without violating the delay constraint. In addition, the time allocation parameter τ also can have significant effect on

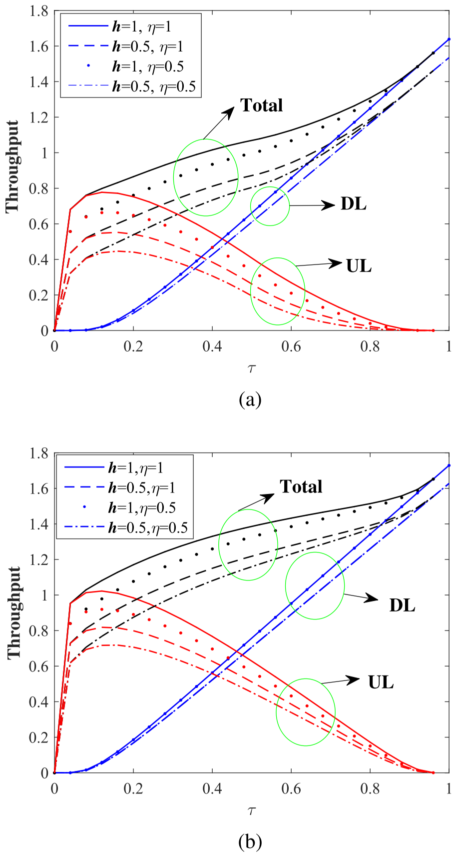


FIGURE 6. Delay-constrained throughput (Mbits per time unit) for stochastic traffic with different τ, where d_X=1 time unit. (a) $\epsilon_D = 10^{-3}$. (b) $\epsilon_D = 10^{-2}$.

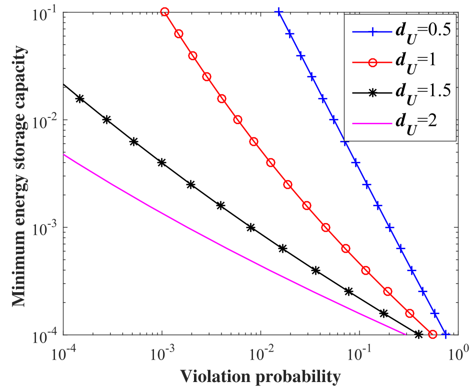


FIGURE 7. Minimum energy storage capacity (J) varying with delay violation probability and delay requirement (time units).

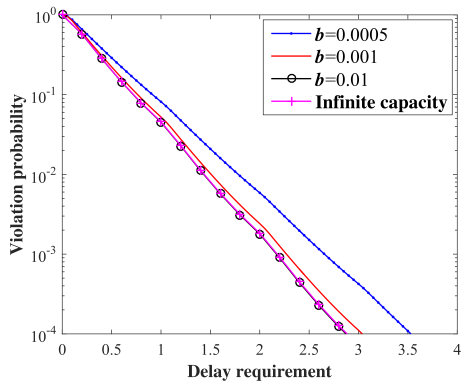


FIGURE 8. Delay violation probability varying with delay requirement (time units) and energy storage capacity (J), where $p_A = 0.01$J per time unit.

the delay-constrained throughput in both DL and UL. This is further shown in Fig. 6.

Fig. 6 illustrates the impact of the time allocation parameter τ on the delay-constrained throughput, where the power gain and energy utilization efficiency are also considered. The figure shows that the power gain has a positive effect on the delay-constrained throughput in both DL and UL. Differently, the energy utilization efficiency only affects the delay-constrained throughput in the UL. Interestingly, the sustained throughput is approximately directly proportional to τ in the DL. This is because a larger τ implies more time being allocated to the AP to transmit the DL data during a TB. Differently, there exits an optimal τ to maximize the delay-constrained throughput in the UL. This is because the delay-constrained throughput is related to both UL transmission period and UL transmission rate according to (5) and (26). As τ increases, the transmission period decreases but the the transmission rate increases. Therefore, there is a tradeoff between the transmission period and the transmission rate to maximize the UL throughput. Furthermore, the optimal τ depends on the delay constraint while observing Fig. 6(a) and Fig. 6(b). In general, increasing τ can ensure higher total throughput for the system, which may sacrifice the UL transmission performance. Thus, care should be taken in choosing proper τ in order to balance the performance requirements from the AP, the UE and the overall system.

The minimum energy storage capacity required to ensure the UL transmission performance is depicted in Fig. 7. The figure shows that the required energy storage capacity is sensitive to the delay constraint. In particular, the stricter delay constraint required by the traffic, the larger energy storage capacity is needed. This is because stricter delay constraint implies higher delay-constrained throughput needed, which further demands larger amount of minimum energy stored before the UL transmission phase. When loosening the delay constraint enough (e.g., $\epsilon_U \geq 10^{-3}$ and $d_U \geq 1$ time unit), the required energy storage capacity may turn to a low value that could be easily equipped into a UE.

In Fig. 8, the delay performance is depicted with different configurations of energy storage capacity. In general, a UE with larger energy storage capacity can guarantee better delay performance. However, the delay performance under finite energy storage capacity converges to the infinite capacity case quickly even when b is as small as 0.01J.

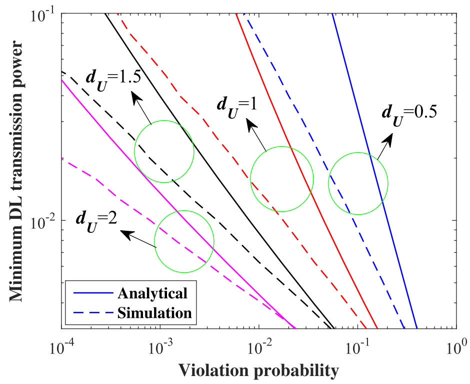


FIGURE 9. Required minimum DL transmission power (J per time unit) varying with delay violation probability and delay requirement (time units).

Finally in Fig. 9, the analytical transmission power is illustrated, which is calculated according to (18) under the given delay constraint. Fig. 9 shows that in low power region (e.g., $p_A \leq 0.02$J per time unit), the power evaluation is reasonably accurate. However, the power evaluation seems to be conservative as p_A increases. This is because the analytical p_A does not consider the impact of parameter E which also increases with p_A according to (18). Nevertheless, the analytical p_A shows the same trend as the simulation results. In addition, though seemly a large gap visually, the analytical conservative estimation is indeed in the same order in magnitude as the simulation results, which as a first try for such analysis satisfies the intention.

VI. CONCLUSION

In this paper we presented an analytical approach to study the delay and delay-constrained throughput performance of a wireless powered communication system with consideration of finite energy storage, stochastic channel and stochastic traffic arrivals. Specifically, cumulative data transmission capacity was first derived for the DL and UL transmissions respectively. Then, based on the cumulative capacity analysis, two upper bounds on the delay distribution were derived for both DL and UL. In addition, the delay-constrained throughput was obtained. Furthermore, the minimum energy storage capacity and minimum DL transmission power were studied to ensure the UL delay constraint. To elaborate the analytical results, two specific types of traffic were considered and more explicit expressions of the analytical results were provided. The simulation results conform the validity and accuracy of the analytical results. We believe, the analysis and the results shed new insights on the performance of WPC systems.

APPENDIX A
PROOF OF LEMMA 1

Proof: Since event $\{D_X(t) > d_X\}$ implies event $\{A_X(t) > A_X^*(t+d_X)\}$, we have $\{D_X(t) > d_X\} \subseteq \{A_X(t) > A_X^*(t+d_X)\}$. Therefore, we have

$$\begin{aligned} &Pr\{D_X(t) > d_X\} \\ &\quad \leq Pr\{A_X(t) - A_X^*(t+d_X) > 0\} \\ &\quad \overset{(a)}{=} Pr\{A_X(t) - \inf_{0\leq s\leq t+d_X}\{A_X(s) + C_X(s,t+d_X)\} > 0\} \\ &\quad \overset{(b)}{=} Pr\{\sup_{0\leq s\leq t}\{A_X(s,t) - C_X(s,t+d_X)\} > 0\} \\ &\quad = Pr\{\sup_{0\leq s\leq t}\{A_X(s,t) - C_X(s,t)\} > C_X(d_X)\} \end{aligned}$$

Here, step (a) is according to (6). In step (b), we adopt the time range $0 \leq s \leq t$ instead of $0 \leq s \leq t+d_X$ since if $s > t$, there will hold as $A_X(s,t) - C_X(s,t+d_X) < 0$.

Let $V_s = e^{\theta_X(A_X(t-s,t)-C_X(t-s,t))}$, $Y_k = A_X(k-1,k)$ and $Z_k = C_X(k-1,k)$. There holds

$$\begin{aligned} V_{s+1} &= e^{\theta_X(A_X(t-s-1,t)-C_X(t-s-1,t))} \\ &= e^{\theta_X\sum_{k=t-s}^{t}(Y_k-Z_k)} \\ &= V_s e^{\theta_X(Y_{t-s}-Z_{t-s})}. \end{aligned}$$

Since $A_X(t)$ and $C_X(t)$ have independent and stationary increments, we have

$$\begin{aligned} &\mathbb{E}[V_{s+1}|V_1, V_2, \ldots V_s] \\ &= \mathbb{E}[V_{s+1}|Y_t, Y_{t-1}, \ldots Y_{t-s+1}, Z_t, Z_{t-1}, \ldots Z_{t-s+1}] \\ &= \mathbb{E}[V_s e^{\theta_X(Y_{t-s}-Z_{t-s})}|Y_t, \ldots Y_{t-s+1}, Z_t, \ldots Z_{t-s+1}] \\ &\overset{(a)}{=} \mathbb{E}[V_s|Y_t, \ldots Y_{t-s+1}, Z_t, \ldots Z_{t-s+1}]\mathbb{E}[e^{\theta_X Y_{t-s}}]\mathbb{E}[e^{-\theta_X Z_{t-s}}] \\ &\overset{(b)}{=} V_s\mathbb{E}[e^{\theta_X A_X(1)}]\mathbb{E}[e^{-\theta_X C_X(1)}] \\ &\leq V_s. \end{aligned}$$

Here, step (a) is due to Y_{t-s} and Z_{t-s} are independent of each other and also independent of $\{Y_t, Y_{t-1}, \ldots Y_{t-s+1}, Z_t, Z_{t-1}, \ldots Z_{t-s+1}\}$. Step (b) holds since processes $A_X(t)$ and $C_X(t)$ both have stationary increments, i.e.,

$$\begin{aligned} \mathbb{E}[e^{\theta_X Y_{t-s}}] &= \mathbb{E}[e^{\theta_X A_X(t-s-1,t-s)}] = \mathbb{E}[e^{\theta_X A_X(1)}], \\ \mathbb{E}[e^{-\theta_X Z_{t-s}}] &= \mathbb{E}[e^{-\theta_X C_X(t-s-1,t-s)}] = \mathbb{E}[e^{-\theta_X C_X(1)}]. \end{aligned}$$

Hence, $V_1, V_2, \ldots, V_t$ form a non-negative supermartingale [18]. Then, according to the property of the supermartingale. the delay distribution holds as [18], [22]

$$\begin{aligned} &Pr\{D_X(t) > d_X\} \\ &\quad = Pr\{\sup_{0\leq s\leq t}\{e^{A_X(s,t)-C_X(s,t)}\} > e^{C_X(d_X)}\} \\ &\quad \leq Pr\{\sup_{1\leq s\leq t}\{V_{t-s}\} > e^{C_X(d_X)}\} \\ &\quad = Pr\{\sup_{1\leq m\leq t}\{V_m\} > e^{C_X(d_X)}\} \\ &\quad \leq Pr\{V_1 > e^{C_X(d_X)}\} \\ &\quad \overset{(a)}{\leq} \mathbb{E}[e^{-\theta_X C_X(d_X)}]\mathbb{E}[e^{\theta_X A_X(1)}]\mathbb{E}[e^{-\theta_X C_X(1)}] \\ &\quad \leq \mathbb{E}[e^{-\theta_X C_X(d_X)}]. \end{aligned}$$

Step (a) is based on the Chernoff bound and the independence between $A_X(t)$ and $C_X(t)$. Therefore, Lemma 1 is proved. ■

APPENDIX B
PROOF OF THEOREM 1

Proof: According to Lemma 1, we have

$$\begin{aligned}
&Pr\{D_D(t) > d_D\}\\
&\le \mathbb{E}[e^{-\theta_D C_D(d_D)}]\\
&= \mathbb{E}[exp\{-\theta_D \sum_{i=1}^{\lfloor \frac{d_D}{T} \rfloor} R_{D,i}\tau T\\
&\quad + R_{D,\lceil \frac{d_D}{T} \rceil} \min\{(d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+, \tau T\}\}]\\
&\overset{(a)}{=} \mathbb{E}[e^{-\theta_D R_D (\lfloor \frac{d_D}{T} \rfloor \tau T + \min\{(d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+, \tau T\})}]\\
&\overset{(b)}{\le} \mathbb{E}[e^{-\theta_D R_D \tau (d_D - \frac{\sigma_D}{n}(1-\tau)T)^+}]\\
&\overset{(c)}{=} \lim_{n\to\infty} \frac{1}{n+1} \sum_{\sigma_D=0}^{n} \mathbb{E}[e^{-\theta_D R_D \tau (d_D - \frac{\sigma_D}{n}(1-\tau)T)^+}]\\
&\le \mathbb{E}[e^{-\theta_D R_D \tau (d_D - (1-\tau)T)^+}].
\end{aligned}$$

Here, step (a) holds due to the reason that R_D is i.i.d over time and independent of σ_D. Step (b) is in terms of Lemma 2 in Appendix D. Besides, in step (c), it is easy to prove that for two independent variables A and B,

$$\mathbb{E}[e^{-\theta AB}] = \mathbb{E}[\mathbb{E}[e^{-\theta Ab}|B = b]].$$

■

APPENDIX C
PROOF OF THEOREM 2

Proof: According to Lemma 1, we have

$$\begin{aligned}
&Pr\{D_U(t) > d_U\}\\
&\le \mathbb{E}[e^{-\theta_U C_D(d_U)}]\\
&= \mathbb{E}[exp\{-\theta_U \sum_{i=1}^{\lfloor \frac{d_U}{T} \rfloor} R_{U,i}(1-\tau) T\\
&\quad + R_{U,\lceil \frac{d_U}{T} \rceil} \min\{(d_U - \lfloor \frac{d_U}{T} \rfloor T - \frac{\sigma_U}{n}\tau T)^+, (1-\tau) T\}\}]\\
&\overset{(a)}{=} \mathbb{E}[e^{-\theta_U R_U (\lfloor \frac{d_U}{T} \rfloor (1-\tau) T + \min\{(d_U - \lfloor \frac{d_U}{T} \rfloor T - \frac{\sigma_U}{n}\tau T)^+, (1-\tau) T\})}]\\
&\overset{(b)}{\le} \mathbb{E}[e^{-\theta_U R_U (1-\tau)(d_U - \frac{\sigma_U}{n}\tau T)^+}]\\
&\overset{(c)}{=} \lim_{n\to\infty} \frac{1}{n+1} \sum_{\sigma_U=0}^{n} \mathbb{E}[e^{-\theta_U R_U (1-\tau)(d_U - \frac{\sigma_U}{n}\tau T)^+}]\\
&\le \mathbb{E}[e^{-\theta_U R_U (1-\tau)(d_U - \tau T)^+}].
\end{aligned}$$

Here, step (a) holds due to the reason that R_U is i.i.d over time and independent of σ_D. Step (b) is in terms of Lemma 3 in Appendix E. ■

APPENDIX D

Lemma 2: For non-negative variables, there always holds

$$\lfloor \frac{d_D}{T} \rfloor T\tau + \min\{(d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+, \tau T\} \ge \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)^+.$$

Proof: Since the left hand side term of the above inequality is always non-negative, we just need to prove that

$$\lfloor \frac{d_D}{T} \rfloor T\tau + \min\{(d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+, \tau T\} \ge \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T).$$

Firstly, it is apparent that

$$\begin{aligned}
&\lfloor \frac{d_D}{T} \rfloor T\tau + \tau T - \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)\\
&\ge (\frac{d_D}{T} - 1)T\tau + \tau T - \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)\\
&\ge \frac{\sigma_D}{n}(1-\tau)T \ge 0.
\end{aligned}$$

Therefore, $\lfloor \frac{d_D}{T} \rfloor T\tau + \tau T \ge \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)^+$

Secondly, if $d_D - \lfloor \frac{d_D}{T} \rfloor T \ge \frac{\sigma_D}{n}(1-\tau)T$, we have

$$\begin{aligned}
&\lfloor \frac{d_D}{T} \rfloor T\tau + (d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+\\
&\quad - \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)\\
&= (d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)(1-\tau) \ge 0.
\end{aligned}$$

On the other hand, if $d_D - \lfloor \frac{d_D}{T} \rfloor T \le \frac{\sigma_D}{n}(1-\tau)T$, we have

$$\begin{aligned}
&\lfloor \frac{d_D}{T} \rfloor T\tau + (d_D - \lfloor \frac{d_D}{T} \rfloor T - \frac{\sigma_D}{n}(1-\tau)T)^+\\
&\quad - \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)\\
&= \lfloor \frac{d_D}{T} \rfloor T\tau - \tau(d_D - \frac{\sigma_D}{n}(1-\tau)T)\\
&= (\frac{\sigma_D}{n}(1-\tau)T - d_D - \lfloor \frac{d_D}{T} \rfloor T)\tau \ge 0.
\end{aligned}$$

Hence, the Lemma 2 is proved. ■

APPENDIX E

Lemma 3: For non-negative variables, there always holds

$$\lfloor \frac{d_U}{T} \rfloor T(1-\tau) + \min\{(d_U - \lfloor \frac{d_U}{T} \rfloor T - \frac{\sigma_U}{n} T)^+, (1-\tau)T\} \ge (1-\tau)(d_U - \frac{\sigma_U}{n}\tau T)^+.$$

Proof: The proof of Lemma 3 is similar with that of Lemma 2, which is consequently omitted for saving space. ■

REFERENCES

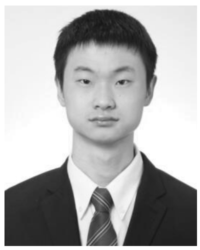
[1] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.
[2] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.
[3] P. Schulz *et al.*, "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 70–78, Feb. 2017.
[4] A. Orsino *et al.*, "Effects of heterogeneous mobility on D2D- and drone-assisted mission-critical MTC in 5G," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 79–87, Feb. 2017.

[5] M. Condoluci, G. Araniti, T. Mahmoodi, and M. Dohler, "Enabling the IoT machine age with 5G: Machine-type multicast services for innovative real-time applications," *IEEE Access*, vol. 4, pp. 5555–5569, May 2016.
[6] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.
[7] Y. L. Che, L. Duan, and R. Zhang, "Spatial throughput maximization of wireless powered communication networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1534–1548, Aug. 2015.
[8] D. Hwang, D. I. Kim, and T. J. Lee, "Throughput maximization for multiuser MIMO wireless powered communication networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5743–5748, Jul. 2016.
[9] Q. Wu, D. W. K. Ng, W. Chen, and R. Schober, "Energy-efficient resource allocation for wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2312–2327, Mar. 2016.
[10] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.
[11] K. Xiong, B. Wang, and K. J. R. Liu, "Rate-energy region of SWIPT for MIMO broadcasting under nonlinear energy harvesting model," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5147–5161, Aug. 2017.
[12] J. Yan and Y. Liu, "A dynamic SWIPT approach for cooperative cognitive radio networks," *IEEE Trans. Veh. Technol.*, to be published.
[13] S. Mahama, D. K. P. Asiedu, and K. J. Lee, "Simultaneous wireless information and power transfer for cooperative relay networks with battery," *IEEE Access*, vol. 5, pp. 13171–13178, Jul. 2017.
[14] J. Yang and Q. Yang, "Power-delay tradeoff in wireless powered communication networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3280–3292, Apr. 2017.
[15] Q. Yao, A. Huang, H. Shan, T. Q. S. Quek, and W. Wang, "Delay-aware wireless powered communication networks—Energy balancing and optimization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5272–5286, Aug. 2016.
[16] S. Kim *et al.*, "Ambient RF energy-harvesting technologies for self-sustainable standalone wireless sensor platforms," *Proc. IEEE*, vol. 102, no. 11, pp. 1649–1666, Nov. 2014.
[17] Z. Li, Y. Jiang, Y. Gao, L. Sang, and D. Yang, "Delay and delay-constrained throughput analysis of a wireless powered communication system," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.
[18] J. L. Doob, *Stochastic processes*. New York, NY, USA: Wiley, 1953.
[19] Y. Jiang and Y. Liu, *Stochastic Network Calculus*. London, U.K.: Springer, 2008.
[20] Y. Gao and Y. Jiang, "Analysis on the capacity of a cognitive radio network under delay constraints," *IEICE Trans. Commun.*, vol. E95-B, no. 4, pp. 1180–1189, 2012.
[21] C. Li, A. Burchard, and J. Liebeherr, "A network calculus with effective bandwidth," *IEEE/ACM Trans. Netw.*, vol. 15, no. 6, pp. 1442–1453, Dec. 2007.
[22] J. F. C. Kingman, "A martingale inequality in the theory of queues," *Math. Proc. Cambridge Philos. Soc.*, vol. 60, no. 2, pp. 359–361, 1964.

ZHIDU LI (S'15) received the B.Eng. degree in communication engineering from the Beijing University of Posts and Telecommunications in 2012, where he is currently pursuing the Ph.D. degree in information and communication engineering. His research interests include performance evaluation for communication systems, especially for energy harvesting communication systems. He is also interested in network calculus theory and its application in typical wireless networks.

YUMING JIANG (SM'14) received the Ph.D. degree from the National University of Singapore in 2001. From 1996 to 1997, he was with Motorola, Beijing, China, and from 2001 to 2003, he was with the Institute for Infocomm Research, Singapore. He has been a Professor with the Norwegian University of Science and Technology, Trondheim, Norway, since 2005. His research interests are the provision, analysis, and management of quality of service guarantees in communication networks, with a particular focus on stochastic network calculus.

YUEHONG GAO received the B.Eng. and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUPT) in 2004 and 2010, respectively, and the Ph.D. degree from the Norwegian University of Science and Technology (NTNU) in 2012. She was a joint Ph.D. Student with the Centre for Quantifiable Quality of Service in Communication Systems, NTNU, under the support of China Scholarship Council from 2008 to 2011. She is currently an Associate Professor with the School of Information and Communication Engineering, BUPT. Her research interests include quality of service guarantees in communication networks, performance evaluation of wireless communication systems and system level simulation methodologies, with special emphasis on 3G/4G/5G networks and heterogeneous networks related topics. She also focuses on the application of network calculus theory in typical wireless scenarios, such as priority network and energy consumption problem.



PENGXIANG LI received the B.Eng. degree from the Beijing University of Posts and Telecommunications, Beijing, in 2014, where he is currently pursuing the Ph.D. degree with the Wireless Theories and Technologies Laboratory. His research interests include channel estimation, precoding, and signal design in massive MIMO.

LIN SANG received the M.S. degree in communication and information system from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1988. From 2004 to 2014, she was the Assistant Dean of the School of Information and Communication Engineering, BUPT. She is currently a Professor with BUPT, where she is involved in communication principle teaching.

DACHENG YANG received the Ph.D. degree in communication and information system from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1988. Since 1988, he has been a Faculty Member of the School of Information and Communications Engineering, BUPT, where he is currently a Professor with the Wireless Theories and Technologies Laboratory. His research interests include mobile communication technologies, deployment and application of mobile communication systems, and standardization of mobile communication systems.