Energy Efficient Mobile Cloud Computing Powered by Wireless Energy Transfer

Changsheng You, Kaibin Huang and Hyukjin Chae

Abstract—Achieving long battery lives or even self sustainability has been a long standing challenge for designing mobile devices. This paper presents a novel solution that seamlessly integrates two technologies, *mobile cloud computing* and *microwave power transfer* (MPT), to enable computation in passive low-complexity devices such as sensors and wearable computing devices. Specifically, considering a single-user system, a base station (BS) either transfers power to or offloads computation from a mobile to the cloud; the mobile uses harvested energy to compute given data either locally or by offloading. A framework for energy efficient computing is proposed that comprises a set of policies for controlling CPU cycles for the mode of local computing, time division between MPT and offloading for the other mode of offloading, and mode selection. Given the CPU-cycle statistics information and channel state information (CSI), the policies aim at maximizing the probability of successfully computing given data, called *computing probability*, under the energy harvesting and deadline constraints. The policy optimization is translated into the equivalent problems of minimizing the mobile energy consumption for local computing and maximizing the mobile energy savings for offloading which are solved using convex optimization theory. The structures of the resultant policies are characterized in closed form. Furthermore, given *non-causal* CSI, the said analytical framework is further developed to support computation load allocation over multiple channel realizations, which further increases the computing probability. Last, simulation demonstrates the feasibility of wirelessly powered mobile cloud computing and the gain of its optimal control.

Index Terms—Wireless power transfer, energy harvesting communications, mobile cloud computing, energy efficient computing.

I. INTRODUCTION

The explosive growth of Internet of Things (IoT) and mobile communication is leading to the deployment of tens of billions of cloud-based mobile sensors and wearable computing devices in near future [1]. Prolonging their battery lives and enhancing their computing capabilities are two key design challenges. They can be tackled by several promising technologies: 1) *microwave power transfer* (MPT) for powering the mobiles using microwaves [2], 2) *mobile computation offloading* (MCO) for offloading computation-intensive tasks from the mobiles to the cloud [3], and 3) energy efficient *local computing* using the mobile CPU. These technologies are seamlessly integrated in the current work to develop a novel design framework for realizing wirelessly powered mobile cloud computing under the criterion of maximizing the probability of successfully computing given data, called

C. You and K. Huang are with the Dept. of EEE at The University of Hong Kong, Hong Kong (Email: csyou@eee.hku.hk, huangkb@eee.hku.hk). H. Chae is with LG Electronics, S. Korea (Email: hyukjin.chae@lge.com). Updated on March 14, 2016.

computing probability. The framework is feasible since MPT has been proven in various experiments for powering small devices such as sensors [4] or even small-scale airplanes and helicopters [5]. Furthermore, sensors and wearable computing devices targeted in the framework are expected to be connected by the cloud-based IoT in the future [1], providing a suitable platform for realizing MCO.

A. Prior Work

MCO has been an active research area in computer science [3] where research has focused on designing mobile-cloud systems and software architectures [6], [7], virtual machine migration design in the cloud [8] and code partitioning techniques in the mobiles [6] for reducing the energy consumption and improving the computing performance of mobiles. Nevertheless, implementation of MCO requires data transmission and message passing over wireless channels, incurring transmission power consumption [9]. The existence of such a tradeoff has motivated cross-disciplinary research on jointly designing MCO and adaptive transmission algorithms to maximize the mobile energy savings [10]–[12]. A stochastic control algorithm was proposed in [10] for adapting the offloaded components of an application to a time-varying wireless channel. Furthermore, multiuser computation offloading in a multi-cell system was explored in [11], where the radio and computational resources were jointly allocated for maximizing the energy savings under the latency constraints. In [12], the threshold-based offloading policy was derived for the system with intermittent connectivity between the mobile and cloud.

Energy-efficient mobile (local) computing is also an active field where rich and diversified techniques have been designed for reducing the mobile energy consumption [13]–[18]. The scheduling of multiple computing tasks was optimized in [13] to increase the energy savings. For the same objective, dynamic power management was proposed in [14] where components of computing tasks were dynamically reconfigured and selectively turned off. Another energy-efficient approach is to control the CPU-cycle frequencies under the deadline constraint by exploiting the fact that lowering the frequencies reduces the energy consumption [15]–[17]. Recently, the CPU-cycle frequencies are jointly controlled with MCO given a stochastic wireless channel in [18]. The framework is further developed in the current work to include the new feature of MPT. This introduces several new design challenges. Among others, the algorithmic design of local computing and offloading becomes more complex under the energy harvesting constraint due to MPT, which prevents energy consumption from exceeding the amount of harvested energy at every time

instant [19]. Another challenge is that MPT and offloading time shares the mobile antenna and the time division has to be optimized.

The MPT technology has been developed for point-to-point high power transmission in the past decades [2]. Recently, the technology is being further developed to power wireless communications. This has resulted in the emergence of an active field called *simultaneous wireless information and power transfer* (SWIPT). Various SWIPT techniques have been developed by integrating MPT with communication techniques including MIMO transmission [20], OFDMA [21], two-way communication [22] and relaying [23]. Furthermore, existing wireless networks such as cognitive radio [24] and cellular networks [25] have been redesigned to feature MPT. Recent advancements on SWIPT are surveyed in [26] and [27]. Most prior work on SWIPT aims at optimizing communication techniques to maximize the MPT efficiency and system throughput. In contrast, the current work focuses on optimizing the local computing and offloading under a different design criterion of maximum computing probability.

B. Contributions and Organization

Consider a single-user system comprising one multi-antenna *base station* (BS) using transmit/receive beamforming for transferring power to a single-antenna mobile or relaying offloaded data from the mobile to the cloud. To compute a fixed amount of data, the mobile operates in one of the two available modes: local computing and offloading. In the mode of local computing, MPT occurs simultaneously as computing based on the controllable CPU-cycle frequencies. Nevertheless, in the mode of offloading, the given computation duration is adaptively partitioned for separate MPT and offloading since they share the mobile antenna. Assume that the mobile has the knowledge of statistics information of CPU cycles and channel state information (CSI). The individual modes as well as mode selection are optimized for maximizing the computing probability under the energy harvesting and deadline constraints. For tractability, the metric is transformed into equivalent ones, namely *average mobile energy consumption* and *mobile energy savings*, for the modes of local computing and offloading, respectively. Compared with [18], the current work integrates MPT with the mobile cloud computing, which introduces new theoretical challenges. In particular, the energy harvesting constraint arising from MPT makes the optimization problem for local computing non-convex. To tackle the challenge, the convex relaxation technique is applied without compromising the optimality of the solution. It is shown in the sequel that the local computing policy of [18] is a special case of the current work where the transferred power is sufficiently high. Furthermore, the case of dynamic channel for mobile cloud computing is explored. Approximation methods are used for deriving the simple and close-to-optimal policies.

The contributions of the current work are summarized as follows.

- *Optimal local computing*: First, consider a static channel. For the mode of local computing, CPU-cycle frequencies are optimized for minimizing the average mobile energy consumption under the energy harvesting and deadline constraints. The corresponding optimization problem is non-convex but solved by the convex relaxation that is proved to maintain the optimality. The resultant policy is shown to depend on two derived thresholds on the transferred power. The first determines the feasibility of successful computing. Given feasible computing, the optimal frequency of the n-th CPU cycle is shown to be *proportional* to $p_k^{-\frac{1}{3}}$ with p_k being the probability of its occurrence and independent of the transferred power if it is above the second threshold; otherwise, the frequency is *proportional* to $(p_k + \lambda)^{-\frac{1}{3}}$ with λ being a constant determined by the transferred power.

- *Optimal computation offloading*: Next, for the mode of offloading, the partition of the computation duration for separate MPT and offloading is optimized to maximize the energy savings. As a result, a threshold is derived on the product of the BS transmission power and squared channel power gain, above which offloading is feasible. Given feasibility, the optimal offloading duration is shown to be *proportional* to the input data size and *inversely proportional* to the channel bandwidth. In other words, small data size and large bandwidth reduces time allocated for offloading and increases time for MPT and vice versa.

- *Mobile mode selection*: The above results are combined to select the mobile mode for maximizing the computing probability. Given feasible computing in both modes, the one yielding the larger energy savings is preferred and the selection criterion is derived in terms of thresholds on the BS transmission power as well as the deadline for computing.

- *Optimal data allocation for a dynamic channel*: Last, the above results are extended to the case of a dynamic channel, modeled as independent and identically distributed (i.i.d.) block fading, and *non-causal* CSI at the mobile (acquired from e.g., channel prediction). The problem of optimizing an individual mobile mode (local computing or offloading) is formulated based on the *master-and-slave model* using the same metric as the fixed-channel counterpart. The master problem concerns the optimal data allocation for computing in a fixed number of fading blocks. The slave problem targets the mode optimization in a single fading block for allocated data and is similar to the fixed-channel counterpart. By approximating the master problems, sub-optimal policies are designed for data allocation and shown by simulation to be close-to-optimal. The results can be straightforwardly combined to enable the mode selection.

The remainder of this paper is organized as follows. The system model is introduced in Section II. Section III presents the optimal policies for mobile mode optimization and selection for the case of static channel. The results are extended in Section IV to the case of dynamic channel. Simulation results are presented in Section V, followed by the conclusion in Section VI.
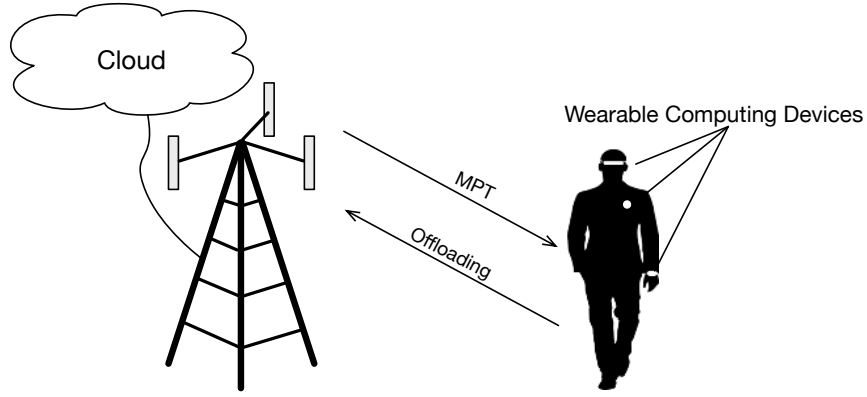
(a) Wirelessly powered mobile cloud computing system
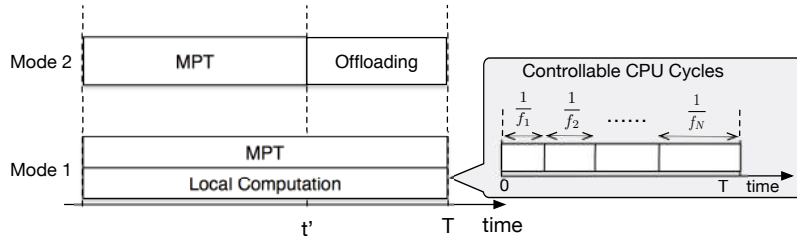
(b) Mobile operation modes

Figure 1: (a) Wirelessly powered mobile cloud computing system and (b) the mobile operation modes.

II. SYSTEM MODEL

Consider a wirelessly powered mobile cloud computing system model shown in Fig. 1(a) comprising one single-antenna mobile, e.g., mobile sensor and wearable computing device with milliwatt power consumption, and one multi-antenna BS that is a node of a cloud. The BS either transfers power wirelessly to or offloads computation from the mobile. Consider the computation of a single task that cannot be split for partial local computing and partial offloading. In other words, the mobile operation mode is either local computing or offloading but not both (see Fig. 1(b)). The local computing and MPT can be performed simultaneously while the offloading and MPT are non-overlapping in time assuming half-duplex transmission. The current work can be generalized to the case of computing a multi-task program which allows program partition and thus the simultaneous operation of these two modes. The multi-antenna BS uses beamforming to transfer power or receive the signal. Assuming channel reciprocity [20], [21], the effective scalar channel power gain is represented as h. Both cases of static and dynamic channel are considered. The dynamic channel is modeled as i.i.d. block fading where the channel power gain is fixed in each fading block and i.i.d. over different blocks. Let P_b denote the BS transmission power. It is assumed that both P_b and h are accurately estimated[1] by the mobile and used for controlling the local computing and offloading. During MPT, assuming the

mobile has infinite battery storage[2], since the energy harvested from the noise is negligible, the energy harvested by the mobile in one time unit is $\upsilon P_b h$ where the constant $0 < \upsilon \leq 1$ represents the energy conversion efficiency.

A. Local Computing Model

Adopting the model in [17], [18], the required number of CPU cycles for computing given data is randomly generated as follows. Let L denote the number of bits in input data for computation and T represent its deadline. Define the *CPU cycle information* (CCI) as the number of CPU cycles required for computing 1-bit of input data. Then the CCI can be modeled as a random variable denoted as X and LX gives the number of CPU cycles for computing L-bit input data [16]. Define N_0 as a positive integer such that $\Pr(X > N_0) \leq \epsilon$ where ϵ is a small real number: $0 < \epsilon \ll 1$. It follows that $\Pr(LX > N) \leq \epsilon$ where $N = LN_0$. Then given L-bit input data, N upper bounds the number of CPU cycles almost surely. Define $p_k = \Pr(LX \geq k)$ with $k = 1, 2, \cdots, N$ such that p_k is the probability that the data has not been processed completely after k CPU cycles. In other words, p_k is the probability that the k-th cycle is executed. Note that the sequence $p_1, p_2, \cdots, p_N$ is monotone decreasing, regardless of any probability distribution of the random variable X.

---

[1]Accounting for imperfect channel knowledge, robust optimization techniques can be applied to the current framework for obtaining control policies to guarantee a required computing probability.

[2]Considering finite battery storage at mobiles will complicate the corresponding structure of the optimal solution. However, it is expected to have the threshold-based structure as we derive for the case of infinite storage.

Assumption 1. The mobile only has the knowledge of the CCI distribution.[3]

The energy consumption of local computing is modeled as follows. In practice, mobile energy consumption is contributed by computation, transmission and fixed circuit power. The circuit power is omitted for simplicity but it can be accounted for by adding a constant in the problem formulation that, however, does not affect the solution method and key results. Let $E(f)$ denote the amount of energy for computation in a single CPU cycle with the frequency f (or equivalently having a CPU-cycle duration of $1/f$). Following the model in [28], [29], under the assumption of a low CPU voltage, $E(f) = \gamma f^2$ where γ is a constant determined by the switched capacitance. However, the analysis of current work also applies to a general polynomial function $E(f)$ that is monotone increasing and convex. Let $f_1, f_2, \cdots, f_N$ denote the CPU-cycle frequencies for CPU cycles $1, 2, \cdots, N$, respectively.

Last, for tractability, several assumptions are made for the case of dynamic channel.

Assumption 2. For the case of dynamic channel, the mobile has *non-causal* CSI.[4]

Non-causal CSI in Assumption 2 corresponds to the case where the mobile predicts the channel variation or has a pre-determined trajectory. The assumption allows off-line data allocation for computing in multiple fading blocks.

Assumption 3. For the case of dynamic channel, the number of channel realizations in the computation duration $[0, T]$ is much smaller than the typical number of CPU cycles required for mobile local computing in the same duration.

Assumption 3 states different time scales for the channel variation and CPU cycles, allowing energy-efficient control policies to be designed using a decomposition approach in the sequel.

Assumption 4. For the case of dynamic channel, the input data can be divided continuously into parts that can be computed separately.

The input data of several applications can be divided, such as Gzip compression and feature extraction. Assumption 4 made for tractability simplifies the data allocation for computing in separate fading blocks to reduce energy consumption. In practice, the optimal solution in the current paper can be discretized by rounding.

B. Computation Offloading Model

Consider computation offloading where the mobile transmits the data to the BS for computing in the cloud and receives the result via the BS. Given the mobile transmission power P_t,

[3]The results on CPU-cycle optimization in the current work can be straightforwardly extended to the case where the mobile has the prior knowledge of CCI instead of CCI distribution, resulting in a simpler policy structure.

[4]Considering causal CSI, if the channel gains of adjacent fading blocks are temporally correlated that can be modeled as the channel state Markov chain, dynamic programming can be used for deriving the optimal solution, which, however, is intractable to analyze and thus is not considered in the current work.

the uplink channel capacity (in bit/s), denoted as C, is given by:

$$C = B \log\left(1 + \frac{P_t h}{\sigma^2}\right)$$

where B is the channel bandwidth and σ^2 is the variance of complex white Gaussian channel noise. Moreover, it is assumed that the time for computing in the cloud and transmitting the computation result from the BS to the mobile is negligible since the cloud has practically infinite computational resources and the BS can afford high transmission power to reduce the downlink transmission delay. Last, the computation result is assumed of a small size such that demodulating the result data at the mobile results in negligible energy consumption compared with that for local computing or offloading.

C. Performance Metrics

Given the CCI X, let $E_{\text{MPT}}(t)$ represent the amount of cumulative energy harvested by the mobile over duration $[0, t]$ and $E_{\text{mob}}(t, X, \{f_k\}_{k=1}^X)$ as the amount of cumulative energy consumed by the mobile for local computing or offloading. Then the performance metric, computing probability, which is the probability of successfully computing given data, is denoted as P_c, defined as

$$P_{\text{c}} = \mathbb{E}\left[I\left(E_{\text{MPT}}(t) \geq E_{\text{mob}}(t, X, \{f_k\}_{k=1}^X), \ \forall \ t \in [0, T]\right)\right] \tag{1}$$

where the indicator function $I(\mathcal{E})$ gives 1 when the event $\mathcal{E}$ occurs and 0 otherwise. Note that the argument of the indicator function in (1) specifies the energy harvesting constraint. It can be observed from (1) that maximizing the computing probability by optimizing the CPU-cycle frequencies is equivalent to minimizing the mobile energy consumption $E_{\text{mob}}(t, X, \{f_k\}_{k=1}^X)$ for $t \in [0, T]$. Nevertheless, this is infeasible since X is unknown to the mobile based on Assumption 1. To overcome this difficulty, the current work instead focuses on minimizing the *expected* energy consumption in the duration $[0, T]$, namely $\mathbb{E}\left[E_{\text{mob}}(T, X, \{f_k\}_{k=1}^X)\right]$, but still under the same constraints, which is given by

$$\mathbb{E}\left[E_{\text{mob}}(T, X, \{f_k\}_{k=1}^X)\right] = \sum_{k=1}^N \gamma p_k f_k^2 + \sum_{k=N+1}^{LX} \gamma p_k f_k^2. \tag{2}$$

Furthermore, for a small ϵ that is close to zero (see Section II-A), the second component of (2) is negligible. Then the following approximation is used for tractability:

$$\mathbb{E}\left[E_{\text{mob}}\left(T, X, \{f_k\}_{k=1}^X\right)\right] \approx \sum_{k=1}^N \gamma p_k f_k^2. \tag{3}$$

Next, consider offloading. By slight abuse of notation, let $E_{\text{MPT}}(t')$ represent the amount of total energy harvested by the mobile over duration $[0, t']$ and $E_{\text{mob}}(t')$ as the amount of energy consumed by the mobile for offloading in the duration $[t', T]$. Then the computing probability for the current case can be defined as

$$P_{\text{c}} = \mathbb{E}\left[I\left(E_{\text{MPT}}(t') \geq E_{\text{mob}}(t'), \ \forall \ t' \in [0, T]\right)\right]. \tag{4}$$

As observed from (4), maximizing the computing probability is equivalent to maximizing the energy savings $[E_{\text{MPT}}(t') - E_{\text{mob}}(t')]$.

III. ENERGY EFFICIENT MOBILE CLOUD COMPUTING WITH A STATIC CHANNEL

In this section, given a static channel, the CPU-cycle frequencies and MPT-and-offloading time division are optimized for local computing and offloading, respectively. Then the results are combined for optimizing the mobile mode selection.

A. Energy Efficient Local Computing with a Static Channel

1) Problem Formulation: Based on the discussion in Section II-C, the problem of optimizing CPU-cycle frequencies aims at minimizing the average mobile energy consumption in (3) under two constraints. The first is the deadline constraint: $\sum_{k=1}^{N} \frac{1}{f_k} \leq T$. The second is the energy harvesting constraint comprising of N sub-constraints given as

$$\sum_{k=1}^{m} \gamma f_k^2 \leq \upsilon P_b h \sum_{k=1}^{m} \frac{1}{f_k}, \quad m = 1, 2, \cdots, N \quad (5)$$

where the left-hand side of the inequality is the total energy consumed by the first m CPU cycles and the right-hand side is the total energy harvested by the end of m-th cycle. It follows that the optimization problem is formulated as:

$$
\begin{aligned}
\min_{\{f_k\}} \quad & \sum_{k=1}^{N} \gamma p_k f_k^2 \\
\text{s.t.} \quad & \sum_{k=1}^{m} \gamma f_k^2 \leq \upsilon P_b h \sum_{k=1}^{m} \frac{1}{f_k}, \quad m = 1, 2, \cdots, N, \\
& \sum_{k=1}^{N} \frac{1}{f_k} \leq T, \\
& f_k > 0, \quad\quad\quad\quad\quad k = 1, 2, \cdots, N.
\end{aligned}
\quad \textbf{(P1)}
$$

Note that p_k is included in the objective function to formulate the average energy consumption, while the energy harvesting constraint is defined for all possible CPU-cycle realizations and thus is without p_k.

2) Solution: It can be observed that the energy harvesting sub-constraints in Problem P1 are non-convex, resulting in a non-convex optimization problem. To address this issue, first, define a set of new variables $\{y_k\}$ as $y_k = \frac{1}{f_k}$ for all k. Substituting them into Problem P1 and furthermore relaxing the equality constraint $y_k f_k = 1$ to be $y_k f_k \geq 1$ gives

$$
\begin{aligned}
\min_{\{f_k, y_k\}} \quad & \sum_{k=1}^{N} \gamma p_k f_k^2 \\
\text{s.t.} \quad & \sum_{k=1}^{m} \gamma f_k^2 \leq \upsilon P_b h \sum_{k=1}^{m} y_k, \quad m = 1, 2, \cdots, N, \\
& \sum_{k=1}^{N} y_k \leq T, \\
& f_k > 0, \quad \frac{1}{f_k} - y_k \leq 0, \quad k = 1, 2, \cdots, N.
\end{aligned}
\quad \textbf{(P2)}
$$

Observed that Problem P2 is a convex optimization problem. Nevertheless, the relaxation mentioned earlier has no effect on the solution optimality as shown in the following lemma.

Lemma 1. The solution for Problem P2 also solves P1.

Proof: See Appendix A. □

Lemma 1 is essential for solving the non-convex Problem P1 by equating it with the convex Problem P2 that yields the same solution but admits powerful convex optimization algorithms.

To characterize the structures of the optimal CPU-cycle frequencies, several useful properties of the solution for Problem P2 are given as follows.

Lemma 2. The solution for Problem P2, denoted as $\{y_k^*, f_k^*\}$, satisfies the following.
  1) The deadline constraint is active:
     $\sum_{k=1}^{N} y_k^* = \sum_{k=1}^{N} \frac{1}{f_k^*} = T$.
  2) The optimal CPU-cycle frequencies can be written as

$$f_k^* = \left[\frac{\mu - \upsilon P_b h \left(\sum_{m=k}^{N} \lambda_m \right)}{2\gamma \left(p_k + \sum_{m=k}^{N} \lambda_m \right)} \right]^{\frac{1}{3}}, \quad \forall k \quad (6)$$

  where the nonnegative variables μ, $\{\lambda_m\}$ are the Lagrange multipliers associated with the deadline and energy harvesting constraints, respectively.
  3) Furthermore, $f_1^* < f_2^* \cdots < f_N^*$.

Proof: See Appendix B. □

Another important property is stated in the following lemma.

Lemma 3. The solution for Problem P2 can also be derived by solving the following Problem P3 that results from removing the first $(N-1)$ energy harvesting sub-constraints of P2:

$$
\begin{aligned}
\min_{\{f_k, y_k\}} \quad & \sum_{k=1}^{N} \gamma p_k f_k^2 \\
\text{s.t.} \quad & \sum_{k=1}^{N} \gamma f_k^2 \leq \upsilon P_b h \sum_{k=1}^{N} y_k, \\
& \sum_{k=1}^{N} y_k \leq T, \\
& f_k > 0, \quad \frac{1}{f_k} - y_k \leq 0, \quad k = 1, 2, \cdots, N.
\end{aligned}
\quad \textbf{(P3)}
$$

Proof: See Appendix C. □

Lemma 3 means the Lagrange multipliers associated with the first $(N-1)$ energy harvesting sub-constraints are equal to zero, i.e., $\lambda_k = 0$ for $k = 1, 2, \cdots N-1$. Combining Property 2) in Lemma 2 and Lemma 3 simplifies the expression for the optimal CPU-cycle frequencies as

$$f_k^* = \left[\frac{\mu - \upsilon P_b h \lambda}{2\gamma (p_k + \lambda)} \right]^{\frac{1}{3}}, \quad \forall k \quad (7)$$

where λ_N is re-denoted as λ for ease of notation. To obtain the closed-form expressions for $\{f_k^*\}$, define two positive constants a and a' as

$$a = \frac{\gamma N^3}{\upsilon T^3} \quad \text{and} \quad a' = \frac{\gamma}{\upsilon T^3} \left(\sum_{k=1}^{N} p_k^{\frac{1}{3}} \right)^2 \left(\sum_{k=1}^{N} p_k^{-\frac{2}{3}} \right). \quad (8)$$

Then the main result of this subsection is stated as follows.

Theorem 1. (Optimal CPU-cycle frequencies for local computing). The optimal CPU-cycle frequencies $\{f_1^*, f_2^*, \cdots, f_N^*\}$ that solve the optimization problem P3 satisfy the following.
1) If $P_b h < a$, $\{f_1^*, f_2^*, \cdots, f_N^*\}$ is an empty set since Problem P3 is infeasible.
2) If $a \leq P_b h < a'$,

$$f_k^* = \left[\frac{1}{T} \sum_{m=1}^{N} (p_m + \lambda)^{\frac{1}{3}}\right] (p_k + \lambda)^{-\frac{1}{3}}, \qquad \forall k \quad (9)$$

where the positive constant λ is the Lagrange multiplier with respect to (w.r.t) the energy harvesting constraint of P3 and it satisfies

$$\left[\sum_{k=1}^{N} (p_k + \lambda)^{\frac{1}{3}}\right]^2 \left[\sum_{k=1}^{N} (p_k + \lambda)^{-\frac{2}{3}}\right] = \frac{\upsilon P_b h T^3}{\gamma}. \quad (10)$$

3) If $P_b h \geq a'$, $\{f_k^*\}$ are independent of $P_b h$,

$$f_k^* = \left(\frac{1}{T} \sum_{m=1}^{N} p_m^{\frac{1}{3}}\right) p_k^{-\frac{1}{3}}, \qquad \forall k. \quad (11)$$

Proof: See Appendix D. □

The first case ($P_b h < a$) corresponds to the scenario where the transferred power $P_b h$ is so low that it is impossible to complete the local computing within the deadline T. In the second case ($a \leq P_b h < a'$), the transferred power is not large but sufficient for meeting the deadline and consequently the optimal CPU-cycle frequencies are functions of $P_b h$. The transferred power is large in the last case ($P_b h \geq a'$) where increasing P_b no longer has any effect on the optimal CPU-cycle frequencies and only increases the amount of energy savings. In this case, the optimal CPU-cycle frequencies depend only on the CCI distribution and the deadline.

Remark 1. As observed from the definition of a in (8) and Theorem 1, there exists a tradeoff between the deadline T and BS transmission power P_b. Specifically, meeting a stricter deadline requires larger P_b and vice versa. In other words, the BS can control P_b to increase the computing probability. However, jointly designing the control policies at the BS and mobile is challenging and P_b is assumed fixed for simplicity.

Remark 2 (BS transmission power and the parameter λ). Recall that λ is the Lagrange multiplier for solving Problem P3. The optimal CPU-cycle frequencies in Cases 2) and 3) of Theorem 1 correspond to $\lambda > 0$ and $\lambda = 0$, respectively. The case of $\lambda = 0$ (or equivalently $P_b h \geq a'$) implies that the energy harvesting constraint in Problem P3 is inactive at the solution point, reducing the problem and solution to be identical to those in [18] considering local computing without MPT.

Corollary 1 (Minimum average energy consumption). It follows from Theorem 1 that the minimum average energy consumption for the local computing, denoted as $\bar{E}_{\text{loc}}^*$, is given as follows.
1) If $a \leq P_b h < a'$,

$$\bar{E}_{\text{loc}}^* = \frac{\gamma}{T^2} \left[\sum_{k=1}^{N} (p_k + \lambda)^{\frac{1}{3}}\right]^2 \left[\sum_{k=1}^{N} p_k (p_k + \lambda)^{-\frac{2}{3}}\right] \quad (12)$$

and $\bar{E}_{\text{loc}}^*$ is a monotone-decreasing function of $P_b h$ satisfying

$$\frac{\gamma}{T^2} \left(\sum_{k=1}^{N} p_k^{\frac{1}{3}}\right)^3 < \bar{E}_{\text{loc}}^* \leq \frac{\gamma N^2}{T^2} \sum_{k=1}^{N} p_k.$$

2) If $P_b h \geq a'$,

$$\bar{E}_{\text{loc}}^* = \frac{\gamma}{T^2} \left(\sum_{k=1}^{N} p_k^{\frac{1}{3}}\right)^3 \quad (13)$$

that is independent of $P_b h$.

Moreover, the corresponding maximum average mobile energy savings, denoted as $\bar{S}_{\text{loc}}^*$, is given as $\bar{S}_{\text{loc}}^* = \upsilon P_b h T - \bar{E}_{\text{loc}}^*$.

Proof: See Appendix E. □

B. Energy Efficient Offloading with a Static Channel

This sub-section focuses on computation offloading. The time division between MPT and offloading is optimized for maximizing the mobile energy savings (see Section II-C).

1) Problem Formulation: The objective function, namely the mobile energy savings, is obtained as follows. As shown in Fig. 1(b), for current operation mode and $t' \in (0, T)$, the time interval $[0, T]$ is divided into two parts: $[0, t']$ and $(t', T]$, corresponding to MPT and offloading, respectively. Let the amount of energy harvested over the interval $[0, t']$ be defined as a function of t': $E_{\text{MPT}}(t') = \upsilon P_b h t'$. Next, consider offloading in the interval $(t', T]$. Fixed-rate transmission over this interval is assumed since this is the most energy-efficient data transmission policy under a deadline constraint as proved in [30]. Then the energy consumption for offloading, denoted as $E_{\text{off}}(t')$, can be written as $E_{\text{off}}(t') = [2^{\frac{L}{B(T-t')}} - 1]\frac{\sigma^2}{h}(T - t')$. The energy savings is thus given as $[E_{\text{MPT}}(t') - E_{\text{off}}(t')]$. Varying t' changes the harvested energy, offloading energy consumption as well as energy savings. Specifically, as t' increases, $E_{\text{MPT}}(t')$ grows linearly with t' but $E_{\text{off}}(t')$ monotonically increases as shown in [30]. Therefore, the energy savings may not be a monotone function and thus optimization is necessary. To simplify notation, define the offloading duration $t = T - t'$ and then $E_{\text{MPT}}(t')$ and $E_{\text{off}}(t')$ can be rewritten as $E_{\text{MPT}}(t) = \upsilon P_b h (T - t)$ and $E_{\text{off}}(t) = (2^{\frac{L}{Bt}} - 1)\frac{\sigma^2}{h}t$. Substituting the expression of $E_{\text{MPT}}(t)$ and $E_{\text{off}}(t)$ rewrites the objective function as

$$E_{\text{MPT}}(t) - E_{\text{off}}(t) = \upsilon P_b h T + \left(\frac{\sigma^2}{h} - \upsilon P_b h\right) t - \frac{\sigma^2}{h} t 2^{\frac{L}{Bt}}. \quad (14)$$

Given this objective function, the problem for the current case is formulated as

$$\begin{aligned} \max_{t} \quad & E_{\text{MPT}}(t) - E_{\text{off}}(t) \\ \text{s.t.} \quad & 0 < t < T, \\ & E_{\text{MPT}}(t) - E_{\text{off}}(t) \geq 0. \end{aligned} \qquad \textbf{(P4)}$$

2) Solution: Define

$$\rho(h) = \frac{\ln 2}{B \left[1 + W(\frac{vP_b h^2}{\sigma^2 e} - \frac{1}{e}) \right]} \qquad (15)$$

where $W(x)$ is the Lambert function defined as the solution for $W(x)e^{W(x)} = x$. Problem P4 is a convex problem as shown in the following lemma.

Lemma 4 (Convexity of P4)**.** The objective function of Problem P4 is a concave function for $t \in (0, \infty)$ and maximized at $t = \rho(h)L$ with $\rho(h)$ defined in (15).

Proof: See Appendix F. $\qquad\square$

Define a positive constant a'' as

$$a'' = \frac{\sigma^2}{v} \left\{ 1 + \left[\frac{L\ln 2}{BT} + W(-e^{-1-\frac{L\ln 2}{BT}}) \right] \right. $$
$$\left. \times \exp\left(\frac{L\ln 2}{BT} + W(-e^{-1-\frac{L\ln 2}{BT}}) + 1 \right) \right\}. \qquad (16)$$

Then optimizing the objective function of Problem P4 over the interval $(0, T)$ and investigating the feasibility condition yield the solution as shown in the following theorem.

Theorem 2. The optimal offloading duration t^* that solves Problem P4 satisfies the following.

1) If $P_b h^2 < a''$, the problem is infeasible.
2) If $P_b h^2 \geq a''$, $t^* = \rho(h)L$ with $\rho(h)$ defined in (15).

Proof: See Appendix G. $\qquad\square$

Remark 3 (Maximum energy savings)**.** It follows from Theorem 2 that the maximum mobile energy savings is

$$S_{\text{off}}^* = vP_b hT - y(h)L$$

where the function $y(h)$ is defined as

$$y(h) = \frac{\sigma^2 \ln 2}{Bh} \exp\left(W\left(\frac{vP_b h^2}{\sigma^2 e} - \frac{1}{e} \right) + 1 \right). \qquad (17)$$

Remark 4. The tradeoff between the deadline T and BS transmission power P_b as discussed in Remark 1 for the local computing also holds for the current operation mode. Moreover, increasing the channel bandwidth B allows a more stringent deadline or smaller P_b.

Remark 5 (BS transmission power vs. offloading duration)**.** It can be observed from the expression of t^* that increasing the BS transmission power P_b reduces the optimal offloading duration t^*. The reason is that higher transmission power is affordable leading to a shorter transmission duration given fixed data to be offloaded.

Remark 6 (Power beacon based MPT)**.** A power beacon can be deployed for performing MPT such that the power transfer and offloading will be served by the power beacon and BS, respectively. This will lead to different gains for the MPT and offloading channels, instead of being identical in the current model, which, however, will not cause significant changes to the key results and policy structures.

C. Offload or Not?

Since the mobile has non-causal CSI, for each channel realization, it can decide the operation mode based on whether the successful computing conditions are satisfied and which mode achieves the larger mobile energy savings.

First, if only one operation mode is feasible, i.e., $h \geq a/P_b$ for local computing or $h \geq \sqrt{a''/P_b}$ for offloading, then this mode is preferred.

Next, if both operation modes are feasible, the desirable mode is selected by comparing the amounts of their maximum energy savings. Define the difference between their maximum energy savings as $\Delta S = S_{\text{off}}^* - \bar{S}_{\text{loc}}^*$. It follows from Corollary 1 and Remark 3 that

$$\Delta S = \frac{\gamma\theta}{T^2} - y(h)L \qquad (18)$$

where the coefficient θ satisfies

$$\left(\sum_{k=1}^{N} p_k^{\frac{1}{3}} \right)^3 \leq \theta \leq N^2 \sum_{k=1}^{N} p_k$$

and $y(h)$ is given in (17). Then offloading should be performed if and only if $\Delta S \geq 0$.

From (18) and (17), the effects of parameters such as the computation deadline T and BS transmission power P_b on the offloading decision are characterized as follows.

1) If $T \leq \sqrt{\frac{\gamma\theta}{y(h)L}}$, offloading is preferred which implies that a more strict deadline requirement tends to select the offloading mode. It can be interpreted as follows. As the deadline increases, the growing rate of harvested energy for local computing is larger than that for offloading. Moreover, the energy consumption for local computing decreases with an increasing deadline, however, that of offloading can be proved to be invariant from (17).
2) If $P_b \leq \frac{\sigma^2}{vh^2}(1 + ea''' \ln a''')$ where $a''' = \frac{Bh\gamma\theta}{eT^2\sigma^2 L \ln 2}$, offloading is selected which indicates that offloading is preferred when the BS transmission power is insufficient. The reason for this threshold-based mode selection w.r.t the variation of BS transmission power is similar to that of deadline.

IV. ENERGY EFFICIENT MOBILE CLOUD COMPUTING WITH A DYNAMIC CHANNEL

While wireless channel is assumed fixed in the preceding section, dynamic channel is considered in this section which is modeled as M channel fading blocks with block duration T_c satisfying $MT_c = T$. The mobile is assumed to have the prior knowledge of the channel power gains in these fading blocks (see Assumption 2), enabling off-line data allocation for computing in different fading blocks. Similar to the fixed-channel counterpart, the mobile is assumed to select one of the two operation modes for a single task in these fading blocks. Local computing and offloading are optimized separately in the following sub-sections. The structures of resultant control policies are also analyzed. The results can be straightforwardly combined to optimize the mobile operation mode selection as in Section III-C with the details omitted for simplicity.

A. Energy Efficient Local Computing with a Dynamic Channel

1) Problem Formulation: Exploiting the different time scales for channel variation and computing in Assumption 3, the problem of optimizing the CPU cycle frequencies for local computing can be decomposed into a master and a slave problems as follows.

- **Slave problem:** Given a particular fading block and allocated data, the slave problem aims at minimizing the average energy consumption for computing in this block by controlling the CPU cycle frequencies in a similar way as the fixed-channel counterpart. Consider a single fading block with channel power gain h and allocated data size ℓ. Moreover, let $R \geq 0$ denote the amount of residual energy from computing in the preceding block. The slave problem for this block is formulated as

$$
\min_{\{f_k\}} \quad \sum_{k=1}^{\ell N_0} \gamma p_k f_k^2
$$
$$
\text{s.t.} \quad \sum_{k=1}^{m} \gamma f_k^2 \leq R + \upsilon P_b h \sum_{k=1}^{m} \frac{1}{f_k}, \quad m = 1, \cdots, \ell N_0,
$$
$$
\sum_{k=1}^{\ell N_0} \frac{1}{f_k} \leq T_c,
$$
$$
f_k > 0, \quad k = 1, \cdots, \ell N_0. \tag{P5}
$$

Setting $R = 0$ reduces Problem P5 to P1 for the case of static channel. Let $G_{\text{loc}}(\ell, R, h)$ denote the minimum average energy consumption for computing ℓ-bit data in a single fading block. Then $G_{\text{loc}}(\ell, R, h) = \sum_{k=1}^{\ell N_0} \gamma p_k (f_k^*)^2$ where $\{f_k^*\}$ solve the above slave problem.

- **Master problem:** The master problem divides the input data for computing in different fading blocks under the criterion of minimum total energy consumption. Let n denote the index of fading blocks. The master problem is formulated as follows.

$$
\min_{\{\ell_n\}} \quad \sum_{n=1}^{M} G_{\text{loc}}(\ell_n, R_n, h_n)
$$
$$
\text{s.t.} \quad G_{\text{loc}}(\ell_n, R_n, h_n) \leq R_n + \upsilon P_b h_n T_c, \quad n = 1, \cdots, M,
$$
$$
R_n = R_{n-1} + \upsilon P_b h_{n-1} T_c
$$
$$
- G_{\text{loc}}(\ell_{n-1}, R_{n-1}, h_{n-1}), \quad n = 2, \cdots, M,
$$
$$
R_1 = 0, \qquad \sum_{n=1}^{M} \ell_n = L,
$$
$$
\ell_n \geq 0, \quad n = 1, \cdots, M. \tag{P6}
$$

Note that the first two constraints of Problem P6 imply that $R_n \geq 0$ for $n \geq 2$.

2) CPU-cycle Control Policy: The policy resulting from solving Problem P5 can be modified from that obtained from solving Problem P1 for the case of static channel. For ease of notation, define two constants:

$$
b = \left(\frac{\upsilon P_b h T_c^3 + R T_c^2}{\gamma \theta_0} \right)^{\frac{1}{3}}, \ b' = \left(\frac{\upsilon P_b h T_c^3 + R T_c^2}{\gamma \theta_1} \right)^{\frac{1}{3}} \tag{19}
$$

and two energy consumption functions

$$
\bar{E}_0(\ell) = \frac{\gamma \varphi_0 \ell^3}{T_c^2}, \qquad \bar{E}_1(\ell) = \frac{\gamma \varphi_1 \ell^3}{T_c^2}
$$

where ℓ is the input-data size and $\theta_0, \theta_1, \varphi_0$ and φ_1 are the scaling factors of ℓ^3 for the two thresholds and energy consumption functions determined by the system model and CCI (see Appendix C of [18]), with $\theta_0 > \theta_1$ and $\varphi_0 < \varphi_1$. Following the same procedure as for deriving Theorem 1 and Corollary 1, the optimal policy for the current case is obtained as follows.

Corollary 2. Consider an arbitrary channel fading block with the corresponding input-data size ℓ and residual energy R. The optimal CPU-cycle frequencies $\{f_k^*\}$ and the minimum average energy consumption $G_{\text{loc}}(\ell, R, h)$ are described as follows.

1) If $\ell \leq b$,

$$
f_k^* = \left(\frac{1}{T_c} \sum_{m=1}^{\ell N_0} p_m^{\frac{1}{3}} \right) p_k^{-\frac{1}{3}}, \ \forall k, \text{ and } G_{\text{loc}}(\ell, R, h) = \bar{E}_0(\ell).
$$

2) If $b < \ell \leq b'$,

$$
f_k^* = \left[\frac{1}{T_c} \sum_{m=1}^{\ell N_0} (p_m + \lambda)^{\frac{1}{3}} \right] (p_k + \lambda)^{-\frac{1}{3}}, \quad \forall k
$$

where the positive constant λ satisfies

$$
\left[\sum_{k=1}^{\ell N_0} (p_k + \lambda)^{\frac{1}{3}} \right]^2 \left[\sum_{k=1}^{\ell N_0} (p_k + \lambda)^{-\frac{2}{3}} \right] = \frac{\upsilon P_b h T_c^3 + R T_c^2}{\gamma};
$$

and

$$
\bar{E}_0(\ell) < G_{\text{loc}}(\ell, R, h) \leq \bar{E}_1(\ell).
$$

Moreover, when $\ell \to b$, $G_{\text{loc}}(\ell, R, h) \to \bar{E}_0(\ell)$ and when $\ell \to b'$, $G_{\text{loc}}(\ell, R, h) \to \bar{E}_1(\ell)$.

3) If $\ell > b'$, $\{f_1^*, f_2^*, \cdots, f_{\ell N_0}^*\}$ is an empty set.

This Corollary shows that the slave problem is feasible only if $\ell \leq b'$ where b' is determined by the channel power gain h and the residual energy R. Moreover, when $\ell \leq b$, the CPU-cycle frequencies and energy consumption are independent of $P_b h$, implying that the energy harvesting constraint is inactive for small data-input size. However, for the case of $b < \ell \leq b'$, both $\{f_k^*\}$ and $G_{\text{loc}}(\ell, R, h)$ are determined by $P_b h$, which is consistent with the case of static channel.

3) Sub-optimal Data Allocation Policy: The derivation for the optimal data allocation policy is intractable due to the lack of a closed-form expression for the energy consumption function, $G_{\text{loc}}(\ell, R, h)$, which can be observed from the solution for the slave problem. In this sub-section, a sub-optimal but simple policy is derived by the approximation for $G_{\text{loc}}(\ell, R, h)$ and the amounts of residual energy in all fading blocks, denoted as $\{R_n\}$.

First, the proposed sub-optimal data allocation policy requires only the following properties of $G_{\text{loc}}(\ell, R, h)$.

Assumption 5. For an arbitrary block fading channel with data-input size ℓ, $G_{\text{loc}}(\ell, R, h)$ is a monotone-increasing, differentiable and convex function for $\ell \in [0, b']$.

The assumption can be justified as follows. First, the monotonicity of $G_{\text{loc}}(\ell, R, h)$ arises from the fact that computing more data requires additional energy. Next, finite computing energy per bit gives the rationality for the differentiability property. Last, the second derivative of $G_{\text{loc}}(\ell, R, h)$ relates to the increase in energy consumption per additional data bit. For the special case of equal CPU-cycle frequencies, the energy consumption of $\frac{\gamma N_0^3 \ell^2}{T_c^2}$ per bit grows with data-input size ℓ, supporting the assumption on the convexity of $G_{\text{loc}}(\ell, R, h)$.

Next, the residual energy variables $\{R_n\}$ are approximated. To this end, R_n can be bounded as follows.

Lemma 5. Given that local computing of input data is feasible, then $R_1 = 0$ and

$$\bar{\varphi}(\upsilon P_b h_{n-1} T_c + R_{n-1}) \leq R_n \leq \upsilon P_b h_{n-1} T_c + R_{n-1} \quad (20)$$

for $n = 2, \cdots, M$ where $\bar{\varphi} = 1 - \frac{\varphi_1}{\theta_1} \geq 0$.

Proof: See Appendix H. □

Note that the lower bound on R_n is nonzero due to the energy harvesting constraint. Since it is difficult to obtain R_n in closed form for the same reason as for deriving $G_{\text{loc}}(\ell, R, h)$ and the upper bound corresponds to the case without data computing, R_n is replaced by its lower bound in the design of the proposed sub-optimal data allocation policy and the result is represented by $\hat{R}_n$. In other words, $\hat{R}_n = \bar{\varphi}(\upsilon P_b h_{n-1} T_c + \hat{R}_{n-1})$ for $n \geq 2$ and $\hat{R}_1 = 0$. Correspondingly, the constants b and b' defined in (19) are modified as

$$\hat{b}_n = \left(\frac{\upsilon P_b h_n T_c^3 + \hat{R}_n T_c^2}{\gamma \theta_0}\right)^{\frac{1}{3}} \text{ and } \hat{b}'_n = \left(\frac{\upsilon P_b h_n T_c^3 + \hat{R}_n T_c^2}{\gamma \theta_1}\right)^{\frac{1}{3}} .$$

Using the above approximation and definitions, the minimum average energy consumption of the n-th fading block, denoted as $\hat{G}_{\text{loc}}$, follows from Corollary 2 and Assumption 5 as

$$\hat{G}_{\text{loc}}(\ell_n, \hat{R}_n, h_n) = \begin{cases} \frac{\gamma \varphi_0 \ell_n^3}{T_c^2}, & \text{if } \ell_n \leq \hat{b}_n \\ g(\ell_n), & \text{if } \hat{b}_n < \ell_n \leq \hat{b}'_n \end{cases} \quad (21)$$

where $g(\ell_n)$ is a general function such that $\hat{G}_{\text{loc}}(\ell_n, \hat{R}_n, h_n)$ has the properties in Assumption 5.

Based on the above approximations, the data-allocation problem for minimizing the total energy consumption can be readily reformulated in a simple form as follows.

(Sub-optimal Data Allocation)

$$\min_{\{\ell_n\}} \quad \sum_{n=1}^{M} \hat{G}_{\text{loc}}(\ell_n, \hat{R}_n, h_n)$$

$$\text{s.t.} \quad \sum_{n=1}^{M} \ell_n = L, \quad \textbf{(P7)}$$

$$0 \leq \ell_n \leq \hat{b}'_n, \qquad n = 1, 2, \cdots, M.$$

Problem P7 is a convex optimization problem. The structure of the resultant data allocation policy can be characterized as follows. Let $b_n(\xi)$ denote the root of equation: $\frac{\partial \hat{G}_{\text{loc}}}{\partial \ell_n}(b_n, \hat{R}_n, h_n) = \xi$ where ξ is a Lagrange multiplier. Then the main result of this sub-section is obtained as shown below.

Proposition 1. If $L \leq \sum_{n=1}^{M} \hat{b}'_n$, Problem P7 is feasible. And the data-allocation policy that solves Problem P7 is given as

$$\ell_n^* = \begin{cases} \hat{b}'_n, & h_n \leq \frac{\gamma \theta_1 b_n^3(\xi) - \hat{R}_n T_c^2}{\upsilon P_b T_c^3} \\ b_n(\xi), & h_n > \frac{\gamma \theta_1 b_n^3(\xi) - \hat{R}_n T_c^2}{\upsilon P_b T_c^3}. \end{cases} \quad (22)$$

Proof: See Appendix I. □

Note that $\{\ell_n^*\}$ are nonzero, indicating that the policy utilizes *all* fading blocks for computing since local computing and MPT can be performed simultaneously over all fading blocks.

B. Energy Efficient Offloading with a Dynamic Channel

1) Problem Formulation: Following the local computing counterpart, the problem for optimal computation offloading is formulated using the master-and-slave model as follows.

- **Slave problem:** Given fixed allocated data, residual energy and channel power gain, the slave problem aims at finding the optimal time division of a fading block for separate energy harvesting and offloading. Consider a single fading block with channel power gain h, allocated data-input size ℓ and the residual energy R which comes from offloading in the preceding block. The slave problem is formulated as follows for maximizing the energy savings in this block with the optimal time division.

$$\max_{t} \quad E_{\text{MPT}}(t, h) - E_{\text{off}}(t, h)$$

$$\text{s.t.} \quad 0 < t < T_c, \quad \textbf{(P8)}$$

$$R + E_{\text{MPT}}(t, h) - E_{\text{off}}(t, h) \geq 0$$

where $E_{\text{MPT}}(t, h) = \upsilon P_b h(T_c - t)$ and $E_{\text{off}}(t, h) = (2^{\frac{\ell}{Bt}} - 1)\frac{\sigma^2}{h} t$. Note that setting $R = 0$ reduces Problem P8 to P4. Let $G_{\text{off}}(\ell, R, h)$ denote the maximum energy savings for offloading the ℓ-bit data in this single fading block. Then $G_{\text{off}}(\ell, R, h) = E_{\text{MPT}}(t^*, h) - E_{\text{off}}(t^*, h)$ where t^* solves the above slave problem.

- **Master problem:** The master problem concerns the optimal data allocation for offloading in different fading blocks with the objective of maximizing the total energy savings. Let n denote the index of fading blocks. Given the solution for the slave problem, the master problem of the optimal data allocation over different fading blocks is formulated as follows.

$$\max_{\{\ell_n\}} \quad \sum_{n=1}^{M} G_{\text{off}}(\ell_n, R_n, h_n)$$

$$\text{s.t.} \quad R_n = \sum_{m=1}^{n-1} G_{\text{off}}(\ell_m, R_m, h_m), \quad n = 2, \cdots, M,$$

$$R_n \geq 0, \qquad n = 2, \cdots, M,$$

$$R_1 = 0, \qquad \sum_{n=1}^{M} \ell_n = L,$$

$$\ell_n \geq 0, \qquad n = 1, 2, \cdots, M. \quad \textbf{(P9)}$$

2) Optimal Time Division Policy: The slave problem, Problem P8, can be modified from the fixed-channel counterpart, Problem P5, by adding the residual energy and thus solved following a similar procedure. For this purpose, define the following constants:

$$c = \frac{T_c B \left[1 + W(\frac{\upsilon P_b h^2}{\sigma^2 e} - \frac{1}{e})\right]}{\ln 2} \text{ and } c' = BT_c \log\left(1 + \frac{Rh}{\sigma^2 T_c}\right).$$

Then the optimal time division policy for the current case with residual energy R is obtained as shown in Corollary 3, following a similar procedure as for deriving Theorem 2.

Corollary 3. Consider an arbitrary channel fading block with ℓ-bit data input and residual energy R. The optimal offloading duration t^* and the maximum energy savings $G_{\text{off}}(\ell, R, h)$ in this block are given as follows.

1) If either: (a) $R \leq \frac{BT_c y(h)}{\ln 2} - \frac{\sigma^2}{h}T_c$ and $\ell \leq \frac{\upsilon P_b h T_c + R}{y(h)}$ or (b) $R > \frac{BT_c y(h)}{\ln 2} - \frac{\sigma^2}{h}T_c$ and $\ell < c$, then

$$t^* = \rho(h)\ell \quad \text{and} \quad G_{\text{off}}(\ell, R, h) = \upsilon P_b h T_c - y(h)\ell$$

where the constants $y(h)$ and $\rho(h)$ are defined in (17) and (15), respectively.

2) If $R > \frac{BT_c y(h)}{\ln 2} - \frac{\sigma^2}{h}T_c$ and $c \leq \ell \leq c'$, then

$$t^* = T_c \quad \text{and} \quad G_{\text{off}}(\ell, R, h) = -(2^{\frac{\ell}{BT_c}} - 1)\frac{\sigma^2}{h}T_c.$$

3) For other combinations of conditions for R and ℓ, Problem P8 is infeasible.

In the above Corollary, the largest data-input size, optimal time division and maximum energy savings are determined by the channel power gain h and the residual energy R. Specifically, the time division of case 1) is the same as that of static channel case and the maximum energy savings decreases linearly with the growing of data-input size. Case 2) corresponds to the scenario where the mobile has large residual energy and data input size. In this case, spending all time on offloading is the optimal time division policy.

3) Sub-optimal Data Allocation Policy: One can observe from problem P9 that different summation terms in the objective function are coupled due to the residual energy delivered from one block to the next. The conventional approach for solving this type of optimization problem is using dynamic programming (DP). The state space for the resultant DP is continuous and has to be discretized to facilitate iterative computation for the optimal policy, bringing high complexity. However, simulation reveals the sub-optimal low-complexity policy to be presented shortly can achieve close performance as DP. More importantly, the DP approach yields little insight to the structure of optimal policy while the said sub-optimal policy allows data allocation to be derived in closed form.

The proposed sub-optimal policy is obtained by setting the residual energy variables as zero: $\{R_n\} = \{0\}$, which is observed from the energy harvesting constraint in Problem P9 to be their lower bounds. Combining the approximation of $\{R_n\}$ with Corollary 3 reduces Problem P9 as

(Sub-optimal Data Allocation)

$$\max_{\{\ell_n\}} \quad \sum_{n=1}^{M} \upsilon P_b h_n T_c - y(h_n)\ell_n$$

$$\text{s.t.} \quad \sum_{n=1}^{M} \ell_n = L, \tag{P10}$$

$$0 \leq \ell_n \leq \frac{\upsilon P_b h_n T_c}{y(h_n)}, \qquad n = 1, 2, \cdots, M.$$

Problem P10 is a convex optimization problem and solving it gives the optimal policy in closed form. To state the policy, let the channel power gains $\{h_n\}$ be rearranged and re-denoted as $\{\tilde{h}_n\}$ such that $\{y(\tilde{h}_n)\}$ are in ascending order: $y(\tilde{h}_1) \leq y(\tilde{h}_2) \leq \cdots \leq y(\tilde{h}_M)$. Moreover, let Π represent the permutation matrix with

$$[\tilde{h}_1, \tilde{h}_2, \cdots, \tilde{h}_M]^T = \Pi \times [h_1, h_2, \cdots, h_M]^T.$$

Using these definitions, the optimal policy from solving Problem P10 is given in the following proposition.

Proposition 2. If $L \leq \sum_{n=1}^{M} \frac{\upsilon P_b h_n T_c}{y(h_n)}$, Problem P10 is feasible. And it can be observed that the data-allocation policy solving Problem P10 is:

$$[\ell_1^*, \ell_2^*, \cdots, \ell_M^*]^T = \Pi^{-1} \times [\tilde{\ell}_1^*, \tilde{\ell}_2^*, \cdots, \tilde{\ell}_M^*]^T$$

with $\{\tilde{\ell}_n^*\}$ given in the following

$$\tilde{\ell}_n^* = \begin{cases} \frac{\upsilon P_b \tilde{h}_n T_c}{y(\tilde{h}_n)}, & n = 1, 2, \cdots j \\ L - \sum_{k=1}^{j} \frac{\upsilon P_b \tilde{h}_k T_c}{y(\tilde{h}_k)} & n = j + 1, \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

where j is unique and satisfies:

$$\sum_{n=1}^{j} \frac{\upsilon P_b \tilde{h}_n T_c}{y(\tilde{h}_n)} < L \leq \sum_{n=1}^{j+1} \frac{\upsilon P_b \tilde{h}_n T_c}{y(\tilde{h}_n)}.$$

The above data allocation policy is a greedy approach which allocates data to the fading blocks sequently by the ascending order of $y(h_n)$ until all the input data has been allocated.

V. SIMULATION RESULTS

In this section, the performance of wirelessly powered mobile cloud computing with static and dynamic channels is investigated by simulation. The parameters are set as follows unless specific otherwise. The data input size L is 1000-bit and the number of CPU cycles required for per bit is modeled by a Gamma distribution with $\alpha = 4$ and $\beta = 200$ as in [18], resulting in $\epsilon = 0.05$. The constant determined by the switch capacitance γ is 10^{-28}. The energy conversion efficiency υ is 0.8 [31]. Let N_t denote the number of BS antennas and set as $N_t = 2$. The $N_t \times 1$ vector channel, denoted as $\mathbf{h}$, is assumed to follow Rician fading and thus modeled as

$$\mathbf{h} = \sqrt{\frac{\Omega K}{1 + K}}\bar{\mathbf{h}} + \sqrt{\frac{\Omega}{1 + K}}\mathbf{h_w}$$

where the Rician factor $K \in \{0, 10\}$, the average fading power gain $\Omega = 5 \times 10^{-6}$, line-of-sight (LoS) component $\bar{\mathbf{h}}$ has all elements equal to one, and $\mathbf{h_w}$ is a $N_t \times 1$
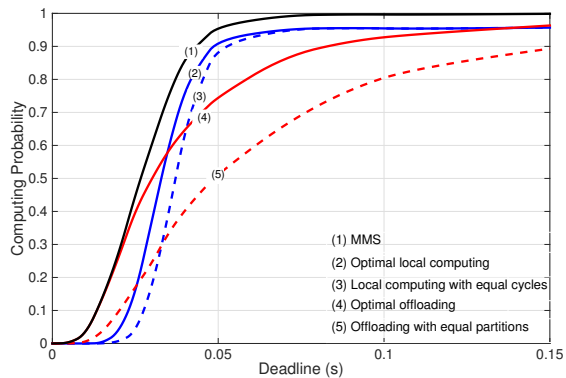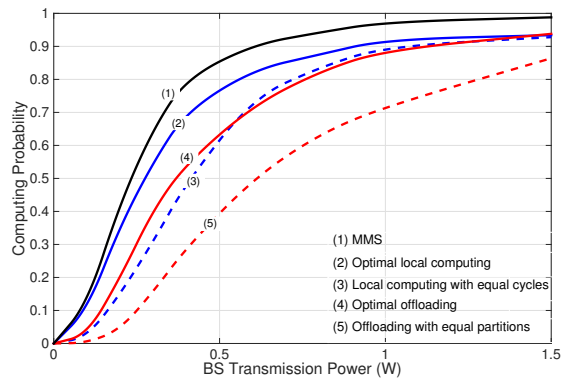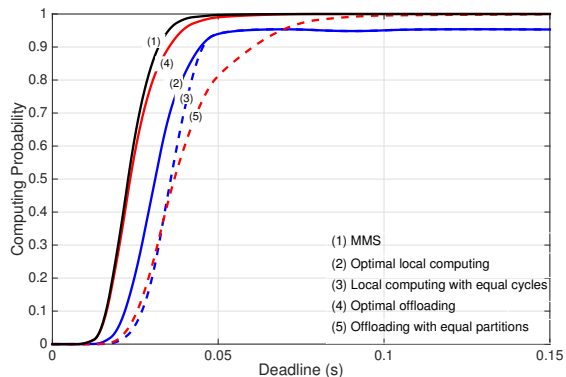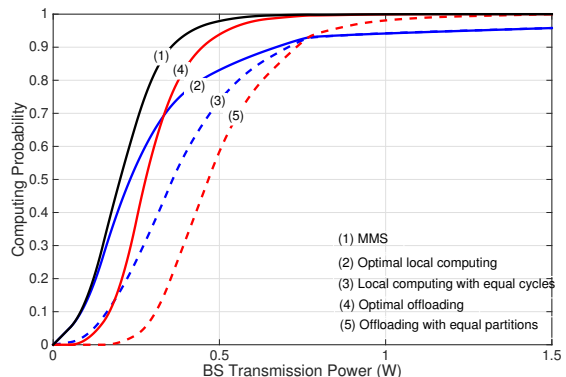
(a) Rician factor $K = 0$

(b) Rician factor $K = 10$

Figure 2: Effect of deadline on the computing probability for the case of static channel. The BS transmission power is fixed at $P_b = 0.5$ W.

(a) Rician factor $K = 0$

(b) Rician factor $K = 10$

Figure 3: Effect of BS transmission power on the computing probability for the static channel. The deadline is fixed at $T = 0.035$ s.

i.i.d. $\mathcal{CN}(0, 1)$ vector representing small-scale fading. The effective channel power gain $h = \|\mathbf{h}\|^2$, resulting from the transmit/receive beamforming. In addition, the variance of complex white Gaussian channel noise σ^2 is 10^{-9} W and the channel bandwidth B is 1 MHz.

A. Static Channel

The performance of three polices is evaluated, including the optimal local computing, optimal offloading and optimal mobile mode selection (MMS) integrating the preceding two polices. The baseline schemes for optimal local computing and optimal offloading are local computing with equal CPU-cycle frequencies and offloading with equal time partition (for MPT and offloading), respectively.

Fig. 2 (a) and (b) show the curves of computing probability versus deadline T for the Rician factor $K = 0$ and 10, respectively. Several observations can be made. First, computing probability is observed to be a monotone-increasing function of T. Next, for a highly random channel ($K = 0$), the crossing of the curves (2) and (4) suggests mode switching as the deadline increases. Specifically, local computing and offloading are preferred for relatively loose and strict deadlines, respectively. The reason is that compared with offloading, the computing probability for local computing grows faster as the deadline increases and also decays faster as the deadline decreases. Note
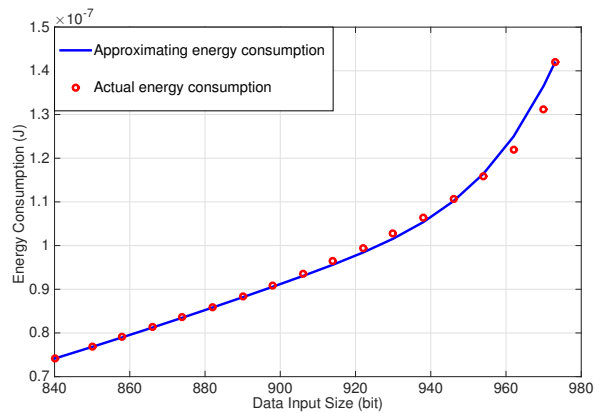

Figure 4: Approximation $\hat{G}_{\text{loc}}$ with $P_b = 1$ W and $T = 0.035s$.

that the thresholds for mode switching w.r.t the computing probability has no simple closed form. However, given LoS ($K = 10$), the optimal offloading is always preferred since the required transmission energy is small for such a channel. Last, compared with their corresponding baseline schemes, optimizing offloading shows more substantial performance gain than optimizing local computing.

The curves of computing probability versus BS transmission power P_b are plotted in Fig. 3 (a) for $K = 0$ and in Fig. 3 (b) for $K = 10$ with a fixed computing deadline
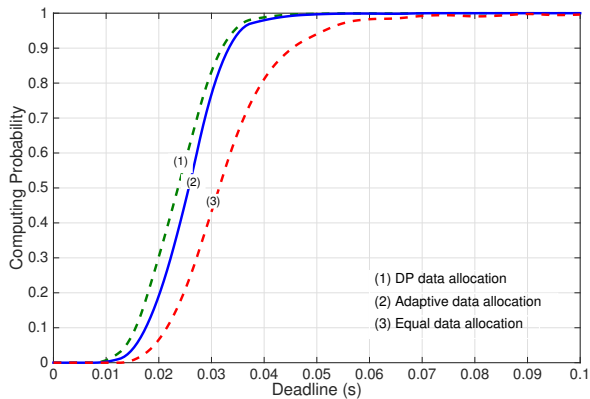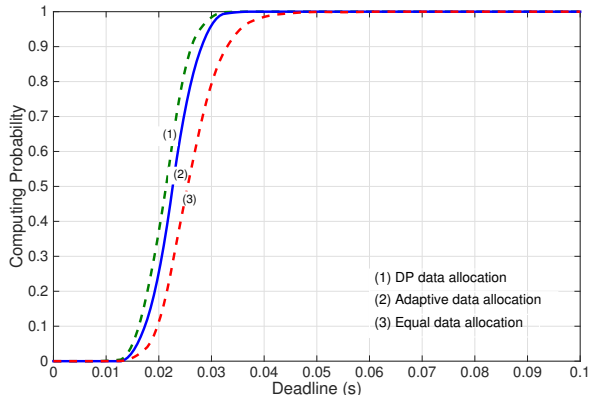
(a) Rician factor $K = 0$

(b) Rician factor $K = 10$

Figure 5: Effect of deadline on the computing probability for the dynamic channel. The BS transmission power is fixed at $P_b = 0.5$ W.
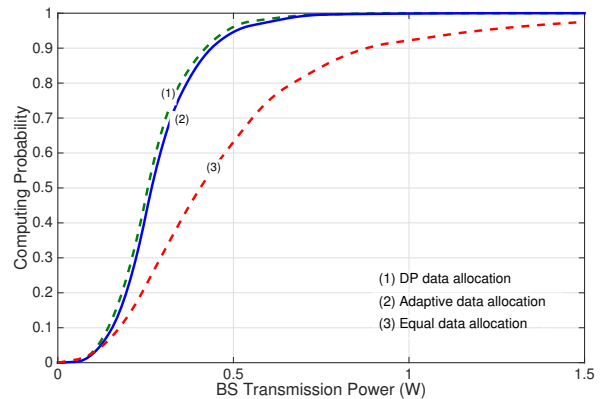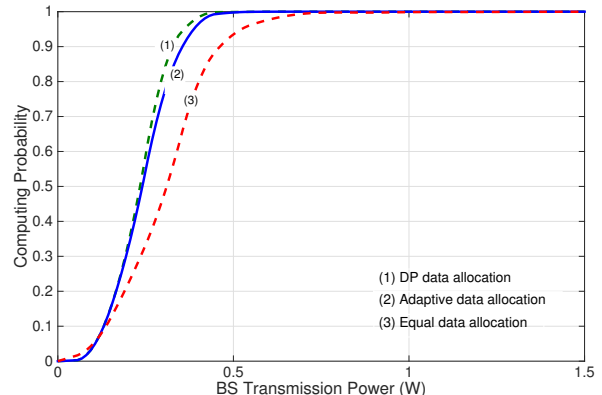

(a) Rician factor $K = 0$

(b) Rician factor $K = 10$

Figure 6: Effect of BS transmission power on the computing probability for the dynamic channel. The computing deadline is fixed at $T = 0.035$ s.

$T = 0.035$s. As observed from the figures, for a highly random channel ($K = 0$), the optimal local computing is preferred to offloading. However, for a channel with LoS ($K = 10$), the optimal local computing is preferred only when P_b is small while optimal offloading should be chosen for large P_b. Other observations are similar to those from Fig. 2.

Consider the case of dynamic channel and data allocation over 4 fading blocks. The relatively small number of fading blocks is assumed to account for the difficulties of large-range channel prediction in practice. The sub-optimal data allocation policies derived in the preceding section are compared with two baseline polices: the optimal one computed based on the derived sub-optimal data allocation policy for local computing and DP policy for offloading, referred as *DP data allocation*, as well as a simple policy based on *equal data allocation*. Specifically, the DP policy for offloading is obtained by discretizing the state space and solving the Bellman equation backward recursively [32]. Each policy is feasible if one of its two operation modes is possible. For local computing, the minimum average energy consumption for the n-th fading block $\hat{G}_{\text{loc}}(\ell_n, \hat{R}_n, h_n)$ has no closed form. To allow simulation, the function with the properties in Assumption 5 is approximated

as

$$g(\ell_n) \approx \frac{\gamma \hat{\varphi}(\ell_n) \ell_n^3}{T_c^2} \qquad (24)$$

where $\hat{\varphi}(\ell_n) = \frac{(\varphi_1 - \varphi_0)}{(\hat{b}'_n - \hat{b}_n)^4}(\ell - \hat{b}_n)^4 + \varphi_0$ is a polynomial monotone-increasing and convex function satisfying $\hat{\varphi}(\hat{b}_n) = \varphi_0$ and $\hat{\varphi}(\hat{b}'_n) = \varphi_1$. The approximation is verified by simulation to be accurate as shown in Fig. 4.

Fig. 5 (a) and (b) show the curves of computing probability versus deadline T for the Rician factor $K = 0$ and 10, respectively. The proposed sub-optimal data allocation policy is found to have close-to-optimal performance and substantial performance gain over the equal allocation policy. Moreover, the gain for a highly random channel is larger than that for a LoS channel. This shows that adaptive data allocation is an effective way for coping with the effect of fading on mobile cloud computing.

Last, the curves of computing probability versus BS transmission power P_b are plotted in Fig. 6 (a) for $K = 0$ and in Fig. 6 (b) for $K = 10$ with a fixed computing deadline $T = 0.035$ s. Large performance gain is observed for data allocation for the case of highly random channel. Other observations are similar to those from Fig. 5.

VI. CONCLUSION

A novel framework of wirelessly powered mobile cloud computing has been proposed in this paper. Applying optimization theory, a set of policies have been derived for optimizing the computing performance of two mobile operation modes, namely local computing and computation offloading, under the energy harvesting and deadline constraints. Furthermore, given non-causal CSI, a sub-optimal policy for adaptive data allocation in time has been proposed to cope with the effect of fading on the system computing performance and shown to be close-to-optimal. This set of policies constitute a promising framework for realizing wirelessly powered and cloud-based mobile devices.

This work can be extended to several interesting directions. First, full-duplex transmission can be implemented in the proposed system to support simultaneous MPT and computation offloading to improve the power transfer efficiency. Second, the current work focusing on a single-computing task can be generalized to the scenario of computing a multi-task program, which involves program partitioning and simultaneous local computing and offloading. Last, it is interesting to extend the current design for single-user mobile cloud computing system to the multiuser system that requires joint design of radio and computational resource allocation for mobile cloud computing.

APPENDIX

A. Proof of Lemma 1

Define the Lagrangian function for Problem P2 as

$$L = \sum_{k=1}^{N} \gamma p_k f_k^2 + \sum_{m=1}^{N} \lambda_m \left(\sum_{k=1}^{m} \gamma f_k^2 - \upsilon P_b h \sum_{k=1}^{m} y_k \right)$$
$$+ \mu \left(\sum_{k=1}^{N} y_k - T \right) + \sum_{k=1}^{N} \eta_k \left(\frac{1}{f_k} - y_k \right). \quad (25)$$

Applying the Karush-Kuhn-Tucker (KKT) conditions gives:

$$\frac{\partial L}{\partial f_k^*} = 2\gamma p_k f_k^* + 2\gamma f_k^* \left(\sum_{m=k}^{N} \lambda_m \right) - \eta_k \frac{1}{(f_k^*)^2} = 0, \forall k, \quad (26)$$

$$\frac{\partial L}{\partial y_k^*} = -\upsilon P_b h \left(\sum_{m=k}^{N} \lambda_m \right) + \mu - \eta_k = 0, \quad \forall k, \quad (27)$$

$$\lambda_m \left[\sum_{k=1}^{m} \gamma (f_k^*)^2 - \upsilon P_b h \sum_{k=1}^{m} y_k^* \right] = 0, \quad \forall m, \quad (28)$$

$$\mu \left(\sum_{k=1}^{N} y_k^* - T \right) = 0, \quad (29)$$

$$\eta_k \left(\frac{1}{f_k^*} - y_k^* \right) = 0, \quad \forall k, \quad (30)$$

$$\lambda_k \geq 0, \mu \geq 0, \eta_k \geq 0, f_k^* > 0, y_k^* > 0, \quad \forall k.$$

Then it is derived from (26) that

$$(f_k^*)^3 = \frac{\eta_k}{2\gamma \left(p_k + \sum_{m=k}^{N} \lambda_m \right)}, \quad \forall k. \quad (31)$$

To ensure $f_k^* > 0$, since γ and p_k are positive and $\{\lambda_m\}$ are nonnegative, it needs to satisfy that: $\eta_k > 0$ for all k.

Combining it with (30) yields that $y_k^* = \frac{1}{f_k^*}$, leading to the desired result. ∎

B. Proof of Lemma 2

First, due to the positivity of η_k (see Lemma 1), it follows from (27) that: $\mu = \eta_k + \upsilon P_b h \left(\sum_{m=k}^{N} \lambda_k \right) > 0$. Combining it with (29) gives: $\sum_{k=1}^{N} y_k^* = \sum_{k=1}^{N} \frac{1}{f_k^*} = T$.

Second, the optimal CPU-cycle frequencies (6) can be obtained by combining (27) and (31).

Third, compare f_k^* and f_{k+1}^* based on (6). Since $p_{k+1} < p_k$ and $\sum_{m=k+1}^{N} \lambda_m \leq \sum_{m=k}^{N} \lambda_m$, it follows that $f_{k+1}^* > f_k^*$, completing the proof. ∎

C. Proof of Lemma 3

It is proved by contradiction as follows. Assume there exists an integer m where $2 \leq m \leq N-1$ such that $\lambda_m > 0$. It follows from (28) that

$$\sum_{k=1}^{m} \left[\gamma (f_k^*)^2 - \upsilon P_b h \frac{1}{f_k^*} \right]$$
$$= \sum_{k=1}^{m-1} \left[\gamma (f_k^*)^2 - \upsilon P_b h \frac{1}{f_k^*} \right] + \left[\gamma (f_m^*)^2 - \upsilon P_b h \frac{1}{f_m^*} \right] = 0.$$

Since $\sum_{k=1}^{m-1} \left[\gamma (f_k^*)^2 - \upsilon P_b h \frac{1}{f_k^*} \right] \leq 0$, it can be obtained that $\gamma (f_m^*)^2 - \upsilon P_b h \frac{1}{f_m^*} \geq 0$. Combining it with the monotonicity of $\{f_k^*\}$ (see Lemma 2), for the $(m+1)$-th cycle, it has $\gamma (f_{m+1}^*)^2 - \upsilon P_b h \frac{1}{f_{m+1}^*} > \gamma (f_m^*)^2 - \upsilon P_b h \frac{1}{f_m^*} \geq 0$, which results in

$$\sum_{k=1}^{m+1} \left[\gamma (f_k^*)^2 - \upsilon P_b h \frac{1}{f_k^*} \right]$$
$$= \sum_{k=1}^{m} \left[\gamma (f_k^*)^2 - \upsilon P_b h \frac{1}{f_k^*} \right] + \left[\gamma (f_{m+1}^*)^2 - \upsilon P_b h \frac{1}{f_{m+1}^*} \right] > 0,$$

contradicting the energy harvesting constraint. Similar proof by contradiction also applies to λ_1. Therefore, the Lagrange multipliers $\{\lambda_m\}$ for Problem P2 satisfy $\lambda_1 = \lambda_2 = \cdots = \lambda_{N-1} = 0$ and $\lambda_N \geq 0$, yielding the desired result. ∎

D. Proof of Theorem 1

Considering Problem P3, the conditions for feasible cases are derived as follows.

1) Case 1: $\lambda > 0$. First, substituting (7) into the deadline constraint leads to

$$\left(\frac{2\gamma}{\mu - \upsilon P_b h \lambda} \right)^{\frac{1}{3}} \left[\sum_{k=1}^{N} (p_k + \lambda)^{\frac{1}{3}} \right] = T. \quad (32)$$

Combining (7) and (32) gives (9). Next, substituting (7) into the energy harvesting constraint yields

$$\left(\frac{\mu - \upsilon P_b h \lambda}{2\gamma} \right)^{\frac{2}{3}} \left[\sum_{k=1}^{N} (p_k + \lambda)^{-\frac{2}{3}} \right] = \frac{\upsilon P_b h T}{\gamma}. \quad (33)$$

Combining (32) and (33) results in (10). Then, consider the monotone property of $P_b h$. From (10), the first derivative of $P_b h$ w.r.t λ is

$$\frac{\partial (P_b h)}{\partial \lambda} = \frac{2\gamma}{3\upsilon T^3} \left(\sum_{k=1}^{N} u_k \right)$$
$$\times \left[\left(\sum_{k=1}^{N} v_k \right)^2 - \left(\sum_{k=1}^{N} v_k w_k \right) \left(\sum_{k=1}^{N} \frac{v_k}{w_k} \right) \right]$$

where $u_k = (p_k + \lambda)^{\frac{1}{3}}$, $v_k = (p_k + \lambda)^{\frac{-2}{3}}$ and $w_k = p_k + \lambda$. Applying Cauchy inequality gives that: $\left(\sum_{k=1}^{N} v_k w_k \right) \left(\sum_{k=1}^{N} \frac{v_k}{w_k} \right) \overset{(a)}{\geq} \left(\sum_{k=1}^{N} v_k \right)^2$. The equality in (a) holds only when $\{\sqrt{v_k w_k} / \sqrt{v_k / w_k}\}$ are equal for all k, which cannot be satisfied in this problem. Consequently, $\frac{\partial P_b}{\partial \lambda} < 0$. Furthermore, the asymptotic properties of $P_b h$ w.r.t λ are characterized: when $\lambda \to 0$, it has $P_b h \to a'$; when $\lambda \to \infty$, it has $P_b h \to a$ where a and a' are defined in (8). Combining them with the monotone-decreasing property of $P_b h$ gives that: if $0 < \lambda < \infty$, it leads to $a < P_b h < a'$. In addition, when $P_b h = a$, there is only one feasible solution: $f_k^* = \frac{N}{T}$ for all k, which can also be expressed as (9) by letting $\lambda = \infty$.

2) *Case 2*: $\lambda = 0$. Since (9) is derived only using the deadline constraint, (11) can be obtained by letting $\lambda = 0$ in (9). Substituting it into the energy harvesting constraint $(\sum_{k=1}^{N} \gamma(f_k^*)^2 \leq \upsilon P_b h T)$ gives: $P_b h \geq a'$.

Thus, it can be concluded that if $P_b h < a$, Problem P3 is infeasible, completing the proof. ∎

E. Proof of Corollary 1

The results of (12), (13) and asymptotic properties can be derived straightforwardly following Theorem 1. The monotone property of $\bar{E}_{\text{loc}}^*$ for the case of $a \leq P_b h < a'$ is proved as follows.

For notation simplicity, define $a_k = (p_k + \lambda)^{\frac{-1}{3}}$ such that $p_k = a_k^{-3} - \lambda$. From (12), the first derivative of $\bar{E}_{\text{loc}}^*$ w.r.t λ is:

$$\frac{\partial \bar{E}_{\text{loc}}^*}{\partial \lambda} = \frac{2\gamma}{3T^2} \left(\sum_{k=1}^{N} a_k^{-1} \right) \left[\left(\sum_{k=1}^{N} a_k^2 \right) \left(\sum_{k=1}^{N} (a_k^{-3} - \lambda) a_k^2 \right) \right.$$
$$\left. - \left(\sum_{k=1}^{N} a_k^{-1} \right) \left(\sum_{k=1}^{N} (a_k^{-3} - \lambda) a_k^5 \right) \right].$$

By algebraic calculation, the part in the square bracket is

$$\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \frac{1}{a_i a_j} \left[a_i^3 - a_j^3 + \lambda(a_j^6 - a_i^3 a_j^3) \right]$$
$$= \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{a_i a_j} \left[(\lambda(a_i^3 - a_j^3)^2 \right] > 0,$$

leading to $\frac{\partial \bar{E}_{\text{loc}}^*}{\partial \lambda} > 0$. Combing it with $\frac{\partial (P_b h)}{\partial \lambda} < 0$ and one-one mapping between $P_b h$ and λ results in that $\bar{E}_{\text{loc}}^*$ is a monotone-decreasing function of $P_b h$. ∎

F. Proof of Lemma 4

Define two constants for Problem P4: $d = \frac{\sigma^2}{h} - \upsilon P_b h$ and $d' = -\frac{\sigma^2}{h}$. Then the first derivative of S_{off} w.r.t t for $t \in (0, \infty)$ is given as

$$\frac{\partial S_{\text{off}}}{\partial t} = d + \left(d' - \frac{d' L \ln 2}{Bt} \right) 2^{\frac{L}{Bt}}. \tag{34}$$

The second derivative follows:

$$\frac{\partial^2 S_{\text{off}}}{\partial t^2} = \frac{d' L^2 (\ln 2)^2}{B^2 t^3} 2^{\frac{L}{Bt}} < 0, \tag{35}$$

since $d' < 0$, verifying the concavity of S_{off}. Then letting the first derivative (34) be zero gives:

$$\frac{d}{d'} = \left(\frac{L \ln 2}{Bt} - 1 \right) 2^{\frac{L}{Bt}}. \tag{36}$$

Using the Lambert function, the solution for (36) is: $t = \rho(h)L$ with $\rho(h)$ defined in (15). Furthermore, it can be observed from (34) that if $t \to 0$, then $\frac{\partial S_{\text{off}}}{\partial t} \to \infty$. Therefore, S_{off} is maximized at $t = \rho(h)L$, leading to the desired result. ∎

G. Proof of Theorem 2

Based on Lemma 4, if $\rho(h)L \geq T$, S_{off}^* is maximized at $t = T$ and $S_{\text{off}}^* < 0$ such that Problem P4 is infeasible. Therefore, to guarantee the feasibility of Problem P4, two conditions should be satisfied: 1) $t^* = \rho(h)L < T$; 2) $S_{\text{off}}^*(t^*) \geq 0$.

First, since $\frac{\partial S_{\text{off}}}{\partial t} \to \infty$ when $t \to 0$, it only needs to satisfy that when $t = T$, it has $\frac{\partial S_{\text{off}}}{\partial t} < 0$. From (34), it can be obtained that

$$P_b h^2 > \frac{\sigma^2}{\upsilon} \left[\left(\frac{L \ln 2}{BT} - 1 \right) 2^{\frac{L}{BT}} + 1 \right]. \tag{37}$$

Next, substituting t^* satisfying (36) into (14) and letting $S_{\text{off}}^* \geq 0$ gives

$$\frac{\upsilon P_b h^2}{\sigma^2 e} \geq \frac{L \ln 2}{BT} \exp \left(W \left(\frac{\upsilon P_b h^2}{\sigma^2 e} - \frac{1}{e} \right) \right). \tag{38}$$

Denote $d'' = \frac{\upsilon P_b h^2}{\sigma^2 e} - \frac{1}{e}$ and $d''' = \frac{L \ln 2}{BT}$. Then (38) is rewritten as $d'' + \frac{1}{e} \geq d''' e^{W(d'')}$.

Applying $d'' = W(d'') e^{W(d'')}$ to the above inequality and multiplying $e^{-d'''}$ on both sides gives: $[W(d'') - d''']e^{W(d'')-d'''} \geq -e^{-1-d'''}$. Then, it follows that

$$d'' \geq \left[d''' + W(-e^{-1-d'''}) \right] e^{d''' + W(-e^{-1-d'''})}. \tag{39}$$

Substituting the expression of d'' and d''' to (39) gives the solution for (38) as $P_b h^2 \geq a''$ where a'' is defined in (16).

Last, combine the two conditions (37) and $P_b h^2 \geq a''$. Since $-1 < W(-e^{-1-d'''}) < 0$, it has $a'' > \frac{\sigma^2}{\upsilon} \left[\left(\frac{L \ln 2}{BT} - 1 \right) 2^{\frac{L}{BT}} + 1 \right]$. In conclusion, $P_b h^2 \geq a''$, completing the proof. ∎

H. Proof of Lemma 5

According to Corollary 2, the mobile obtains the minimum average energy savings in this block if $\ell_n = b'_n$. In this case, the residual energy for the next fading block is

$$R_{n+1} = R_n + \upsilon P_b h_n T_c - G_{\text{loc}}(b'_n, R_n, h_n) \tag{40}$$

$$= R_n + \upsilon P_b h_n T_c - \frac{\gamma \varphi_1 b'^3_n}{T_c^2}. \tag{41}$$

Then substituting b'_n given in (19) into (41) gives the lower bound. The upper bound is achieved when $\ell_n = 0$. ∎

I. Proof of Proposition 1

Define the Lagrangian function for Problem P7:

$$L = \sum_{n=1}^{M} \hat{G}(\ell_n, \hat{R}_n, h_n) + \xi \left(L - \sum_{n=1}^{M} \ell_n \right) \tag{42}$$
$$+ \sum_{n=1}^{M} \varpi_n (-\ell_n) + \sum_{n=1}^{M} \zeta_n \left(\ell_n - \hat{b}'_n \right).$$

Applying the KKT conditions leads to

$$\frac{\partial L}{\partial \ell_n} = \frac{\partial \hat{G}}{\partial \ell_n}(\ell_n^*, \hat{R}_n, h_n) - \xi - \varpi_n + \zeta_n = 0, \qquad \forall n, \tag{43}$$

$$\varpi_n \ell_n^* = 0, \ \zeta_n \left(\ell_n^* - \hat{b}'_n \right) = 0, \ \varpi_n \geq 0, \ \zeta_n \geq 0, \quad \forall n, \tag{44}$$

$$\sum_{n=1}^{M} \ell_n^* = L. $$

First, it can be proved that $\ell_n^* > 0$ and $\varpi_n = 0$ for all n by the following steps:

1) Observe from (21) that $\frac{\partial \hat{G}}{\partial \ell_n}(\ell_n^*, \hat{R}_n, h_n) \geq 0$ and equals to 0 only when $\ell_n^* = 0$.
2) Suppose there exists a n such that $\ell_n^* = 0$. It leads to $\zeta_n = 0$ and $\frac{\partial \hat{G}}{\partial \ell_n}(\ell_n^*, \hat{R}_n, h_n) = 0$. From (43), it gives $\xi = -\varpi_n \leq 0$.
3) There always exists one j where $j \neq n$ such that $\ell_j^* > 0$. Therefore, $\varpi_j = 0$ and $\frac{\partial \hat{G}}{\partial \ell_j}(\ell_j^*, \hat{R}_j, h_j) > 0$. From (43), it can be derived that $\zeta_j = \xi - \frac{\partial \hat{G}}{\partial \ell_j}(\ell_j^*, \hat{R}_j, h_j) < 0$ which contradicts to the condition $\zeta_j \geq 0$ and leads to the conclusion.

Then, the data allocation follows:

1) If $\ell_n^* < \hat{b}'_n$, then $\zeta_n = 0$ and $\frac{\partial \hat{G}}{\partial \ell_n}(\ell_n^*, \hat{R}_n, h_n) = \xi$.
2) If $\ell_n^* = \hat{b}'_n$, then

$$\frac{\partial \hat{G}}{\partial \ell_n}(\ell_n^*, \hat{R}_n, h_n) + \zeta_n = \frac{\partial \hat{G}}{\partial \ell_n}(\hat{b}'_n, \hat{R}_n, h_n) + \zeta_n = \xi$$

.

In conclusion, $\ell_n^* = \min\{b_n(\xi), \hat{b}'_n\}$ where $b_n(\xi)$ is the root of function $\frac{\partial \hat{G}}{\partial \ell_n}(b_n, \hat{R}_n, h_n) = \xi$ and ξ satisfies $\sum_{n=1}^{M} \ell_n^* = L$. Specifically, when $b_n(\xi) \geq \hat{b}'_n = \left(\frac{\upsilon P_b h_n T_c^3 + \hat{R}_n T_c^2}{\gamma \theta_1} \right)^{\frac{1}{3}}$, it has $h_n \leq \frac{\gamma \theta_1 b_n^3(\xi) - \hat{R}_n T_c^2}{\upsilon P_b T_c^3}$, completing the proof. ∎

REFERENCES

[1] M. Swan, "Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0," *J. Sens. Actuator Networks*, vol. 1, pp. 217–253, 2012.

[2] W. C. Brown, "The history of power transmission by radio waves," *IEEE Trans. Microwave Theory and Techniques*, vol. 32, no. 9, pp. 1230–1242, Sep. 1984.

[3] K. Kumar, J. Liu, Y.-H. Lu, and B. Bhargava, "A survey of computation offloading for mobile systems," *Mobile Networks and Applications*, vol. 18, no. 1, pp. 129–140, 2013.

[4] T. Le, K. Mayaram, and T. Fiez, "Efficient far-field radio frequency energy harvesting for passively powered sensor networks," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 5, pp. 1287–1302, 2008.

[5] N. Shinohara, *Wireless power transfer via radiowaves*. John Wiley & Sons, 2014.

[6] E. Cuervo, A. Balasubramanian, D.-k. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "MAUI: Making smartphones last longer with code offload," in *Proc. ACM MobiSys*, pp. 49–62, Jun. 2010.

[7] S. Kosta, A. Aucinas, P. Hui, R. Mortier, and X. Zhang, "ThinkAir: Dynamic resource allocation and parallel execution in the cloud for mobile code offloading," in *Proc. IEEE INFOCOM*, pp. 945–953, 2012.

[8] L. Gkatzikis and I. Koutsopoulos, "Migrate or not? Exploiting dynamic task migration in mobile cloud computing systems," *IEEE Trans. Wireless Commun.*, vol. 20, pp. 24–32, Jun. 2013.

[9] K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?," *IEEE Computer*, no. 4, pp. 51–56, 2010.

[10] D. Huang, P. Wang, and D. Niyato, "A dynamic offloading algorithm for mobile computing," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 1991–1995, Jun. 2012.

[11] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," *IEEE Trans. Signal and Info. Processing over Nectworks*, Jun. 2015.

[12] Y. Zhang, D. Niyato, and P. Wang, "Offloading in mobile cloudlet systems with intermittent connectivity," *IEEE Trans. Mobile Computing*, vol. 14, no. 12, pp. 2516–2529, Feb. 2015.

[13] F. Yao, A. Demers, and S. Shenker, "A scheduling model for reduced CPU energy," in *Proc. IEEE Ann. Symp. Foundations of Computer Science*, pp. 374–382, 1995.

[14] L. Benini, A. Bogliolo, G. Paleologo, and G. De Micheli, "Policy optimization for dynamic power management," *IEEE Trans. Computer-Aided Design of Integrated Cir. and Sys.*, vol. 18, no. 6, pp. 813–833, 1999.

[15] P. Pillai and K. G. Shin, "Real-time dynamic voltage scaling for low-power embedded operating systems," in *Proc. ACM SIGOPS Operating Sys. Review*, vol. 35, pp. 89–102, 2001.

[16] J. R. Lorch and A. J. Smith, "Improving dynamic voltage scaling algorithms with pace," in *Proc. ACM SIGMETRICS*, vol. 29, pp. 50–61, Jun. 2001.

[17] W. Yuan and K. Nahrstedt, "Energy-efficient soft real-time CPU scheduling for mobile multimedia systems," *ACM Trans. Computer Sys.*, vol. 37, no. 5, pp. 149–163, 2003.

[18] W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. O. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, Sep. 2013.

[19] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE J. Select. Areas Commun.*, vol. 29, pp. 1732–1743, Aug. 2011.

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

[21] D. W. Ng, E. S. Lo, and R. Schober, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 6352–6370, Dec. 2013.

[22] P. Popovski, A. M. Fouladgar, and O. Simeone, "Interactive joint transfer of energy and information," *IEEE Trans. Commun.*, vol. 61, pp. 2086–2097, May. 2013.

[23] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 3622–3636, Jul. 2013.

[24] S. Lee, R. Zhang, and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 4788–4799, Sep. 2013.

[25] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 902–912, Feb. 2014.

[26] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, pp. 117–125, Apr. 2014.

[27] K. Huang and X. Zhou, "Cutting last wires for mobile communication by microwave power transfer," *IEEE Commun. Mag.*, vol. 53, pp. 86–93, Jun. 2015.

[28] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design," *IEICE Trans. Electron.*, vol. 75, no. 4, pp. 371–382, 1992.

[29] T. D. Burd and R. W. Brodersen, "Processor design for portable systems," *J. VLSI Signal Process*, pp. 203–221, Aug. 1996.

[30] B. Prabhakar, E. Uysal Biyikoglu, and A. El Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. IEEE INFOCOM*, vol. 1, pp. 386–394, 2001.

[31] H. Sun, Y.-x. Guo, M. He, and Z. Zhong, "Design of a high-efficiency 2.45-GHz rectenna for low-input-power energy harvesting," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 929–932, 2012.

[32] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 1. Athena Scientific Belmont, MA, 1995.

Hyukjin Chae received the B.S and Ph.D degree in electrical and electronic engineering from Yonsei University, Seoul, Korea. He joined LG Electronics, Korea, as a Senior Research Engineer in 2012. His research interests include interference channels, multiuser MIMO, D2D, V2X, and full duplex radio. From Sep. 2012, he has contributed and participated as a delegate in 3GPP RAN1 with interests in ePDCCH, eIMTA, FD MIMO, Indoor positioning, D2D, and V2X communications. He is an inventor of more than 100 patents.

Changsheng You (S'15) received the B.S. degree in electronic engineering and information science from the University of Science and Technology of China (USTC) in 2014. He is currently working towards the Ph.D. degree in electrical and electronic engineering at The University of Hong Kong (HKU). His research interests include mobile cloud computing, wireless power transfer, energy harvesting system and convex optimization.

Kaibin Huang (M'08-SM'13) received the B.Eng. (first-class hons.) and the M.Eng. from the National University of Singapore, respectively, and the Ph.D. degree from The University of Texas at Austin (UT Austin), all in electrical engineering.

Since Jan. 2014, he has been an assistant professor in the Dept. of Electrical and Electronic Engineering (EEE) at The University of Hong Kong. He is an adjunct professor in the School of EEE at Yonsei University in S. Korea. He used to be a faculty member in the Dept. of Applied Mathematics (AMA) at the Hong Kong Polytechnic University (PolyU) and the Dept. of EEE at Yonsei University. He had been a Postdoctoral Research Fellow in the Department of Electrical and Computer Engineering at the Hong Kong University of Science and Technology from Jun. 2008 to Feb. 2009 and an Associate Scientist at the Institute for Infocomm Research in Singapore from Nov. 1999 to Jul. 2004. His research interests focus on the analysis and design of wireless networks using stochastic geometry and multi-antenna techniques.

He frequently serves on the technical program committees of major IEEE conferences in wireless communications. He has been the technical chair/co-chair for the IEEE CTW 2013, the Comm. Theory Symp. of IEEE GLOBECOM 2014, and the Adv. Topics in Wireless Comm. Symp. of IEEE/CIC ICCC 2014 and has been the track chair/co-chair for IEEE PIMRC 2015, IEE VTC Spring 2013, Asilomar 2011 and IEEE WCNC 2011. Currently, he is an editor for IEEE Journal on Selected Areas in Communications (JSAC) series on Green Communications and Networking, IEEE Transactions on Wireless Communications, IEEE Wireless Communications Letters. He was also a guest editor for the JSAC special issues on communications powered by energy harvesting and an editor for IEEE/KICS Journal of Communication and Networks (2009-2015). He is an elected member of the SPCOM Technical Committee of the IEEE Signal Processing Society. Dr. Huang received the 2015 IEEE ComSoc Asia Pacific Outstanding Paper Award, Outstanding Teaching Award from Yonsei, Motorola Partnerships in Research Grant, the University Continuing Fellowship from UT Austin, and a Best Paper Award from IEEE GLOBECOM 2006 and PolyU AMA in 2013.