IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 6, 2019, accepted January 25, 2019, date of publication February 4, 2019, date of current version February 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897115

Self-Sustainable Dense Cellular M2M System With Hybrid Energy Harvesting and High Sensitivity Rectenna

ZHENJIE TAN[1], HUA QU[1], JIHONG ZHAO[1,2], GONGYE REN[1], AND WENJIE WANG[1]

[1]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710061, China

Corresponding author: Zhenjie Tan (spanspan@stu.xjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61531013 and in part by the National Science and Technology Major Project under Grant 2018ZX03001016.

ABSTRACT The Internet of Things (IoT) is envisioned as an essential enabler of the smart city, which 5G small cell is going to cover. In this paper, we model the dense cellular M2M communication system using a Ginibre determinant point process and study the corresponding green communication metrics, including grid energy consumption and grid energy efficiency (EE). A dedicated millimeter-wave-based hybrid energy harvesting (EH) mechanism is proposed to maintain self-sustainable communication. Two EH modes are discussed, of which Mode 1 is suitable for long-term operation, while Mode 2 is preferred in dealing with a temporary energy shortage. We first minimize the average grid energy consumption by solving a large-scale linear programming problem. When the battery at machine device is charged with sufficient energy at the initial time, closed-form formula describing the average grid energy consumption of both the modes can be derived under moderate throughput constraints, which is proved to have clear relationship with 1) initial battery energy level at machine gateway and machine device; 2) specific hybrid EH time-division strategy; 3) system operation time; and 4) static power consumption at machine gateway and machine device. The numerical solutions are obtained using the interior-point method, and the results are illustrated and discussed. When dealing with the originally non-convex EE maximization problem, we first transform it into a series of sub-problems using the successive convex approximation method and then iteratively solve them with a trick combining Taylor expansion and square error minimization. Non-orthogonal multi-access (NOMA) transmission is investigated in our model to further improve the EE performance. The experiments show that NOMA achieves nearly three times throughput promotion and two times EE improvement.

INDEX TERMS Green communication, cellular M2M communication, millimeter wave, energy harvesting, sustainability.

I. INTRODUCTION

Cellular-based M2M communication [1] is considered as an important scenario of 5G massive machine type communication (mMTC) paradigm and has been intensively studied in literature [1]–[3]. It is one of the promising candidates for an underlying connecting backbone for the Internet of Things (IoT), distinguishing from the other paradigm termed as capillary M2M [4], which features usage of unlicensed low-power short-range radio access technologies (RATs). In [5], the authors propose a two-phase communication scheme where machine-type communication device (MTCD) harvests energy from the uplink signal of machine-type communication gate (MTCG), while performing its own data exchange with the harvested energy. However, a holistic power consumption view has not been established, since only uplink transmission is considered in their work. In [3], a similar time-division scheduling protocol is proposed, together with an energy-efficient radio resource allocation algorithm, which takes into account the latency constraints and configures user equipment as MTCG. According to [6], intermittent arrival nature of energy harvesting (EH) process is captured with a standard Kalman filter, by which data caching is well matched with the battery level.

The associate editor coordinating the review of this manuscript and approving it for publication was Valerio Freschi.

In the work of Mao *et al.* [7], performance of densely-deployed off-grid small cell base stations (SCBSs) is discussed. Here, the EH-capable SCBS is a typical realization of MTCG. In the literature, 'zero-power' self-sustainable standalone electronics are gaining constant attention due to its green communication nature and freeing from human intervention [8], which paves the way to a fully autonomous network management vision. However, far-field EH is usually considered impractical, since the path loss cost is unaffordable. According to [9], the increasing demand for self-sustaining and lifetime extended IoT system is discussed and close to one order of magnitude difference in ambient and dedicated EH efficiency (in nW/cm^2) from base station (BS) and Wi-Fi hotspot is validated. In [10], microwave power transfer (MPT) is applied in a cellular scenario where dedicated energy broadcaster called power beacon (PB) is densely deployed into the coverage area of conventional BS. To compensate the large-scale fading, feasibility and practicability of millimeter wave (mmWave) EH are proposed and discussed in [11], where sharp energy beam is verified as a critical component in achieving proper energy level. In the work of Lu *et al.* [12], Ginibre Determinant Point Process (Gin-DPP) based network model is introduced and analyzed, in which EH capability is enabled at the RF-powered devices to improve energy efficiency (EE) and provide perpetual maintenance. Determinant point process is an expansion of Poisson point process, and is capable of describing mutual repulsive correlation among random points. As we know, locations of access points in an ultra-dense cellular network are not ideally independent to each other, but subject to some specific pattern. The service area of each access point must maintain a 'healthy' condition under which end users experience good quality of service (QoS). The repulsive nature of DPP thus can help produce 'healthy' coverage through location planning of access points.

In this paper, we study the self-sustaining capability of cellular-based M2M communication system where a dedicated energy source called hybrid access point (HAP) [13] is available. In our design, HAP is responsible for mmWave energy broadcasting and control signaling, while MTCG extracts energy from aligned mmWave beam. MTCG then executes simultaneous wireless information and power transfer (SWIPT) [14] towards its serving MTCDs. Finally, MTCD uploads its sensor data to MTCG with the harvested energy during uplink transmission. In order to prolong the lifetime of MTCD and help establish an off-grid M2M network, we assume that both MTCG and MTCD are equipped with a limitless super capacitor as the battery. Similar time-division protocol as in [5], and [15] is adopted, to which we add additional EH phase. Looking carefully into the battery dynamics, we classify MTCD into *undercharged* and *abundantly-charged* states in terms of initial battery energy level, which helps us obtain a clear understanding about the relationship between grid energy consumption and battery energy consumption. Detailed settings of our model are clarified in fig.1.

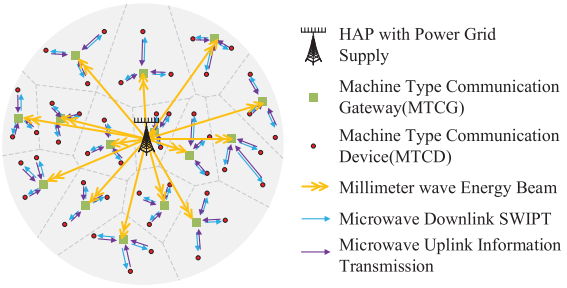


FIGURE 1. System Model (S=15,L=3).

Our main contributions can be concluded as follows: (1) we propose a hybrid EH mechanism in cellular-based M2M system where both MTCGs and MTCDs are getting rid of power lines. (2) Two EH modes are proposed to minimize average grid energy consumption, of which closed-form solutions in terms of system duration time, initial battery energy level and time slot splitting strategy can be derived when MTCDs are *abundantly-charged*. (3) Grid energy efficiency is maximized through our modified successive convex approximation (SCA) method and experimental results are analyzed. (4) Expectation of mmWave energy harvesting rate is derived in terms of gamma function.

The remainder of this paper is organized as follows: Section II introduces our network model and channel model. Section III gives the grid energy consumption problem. Section IV derives the expectation of mmWave EH rate and finds the closed-form solution of proposed problem. Section V establishes the EE maximization problem and solves it with convex approximation. Section VI demonstrates the numerical results and Section VII concludes our work.

II. SYSTEM MODEL

A. NETWORK MODEL

In our model, the HAP is responsible for energy broadcasting, while MTCG conducts both data exchange and power transferring. A total number of S MTCGs from set $\mathcal{S}$ distribute within the coverage area of HAP, which is a circle of radius r^a, according to Gin-DPP. U MTCDs from set $\mathcal{U}$ are relayed by MTCGs and a given number of L MTCDs are uniformly located in the proximity of each MTCG within a circle of radius r^0. We assume that content required by MTCDs have already been cached at its serving MTCG [16] since most of them are configuration updates changing on a relatively large time scale, therefore only control signal is backhauled between MTCG and HAP, and the corresponding throughput is negligible in our analysis.

We consider a hybrid EH process containing the following phases: (1) mmWave-based downlink energy broadcasting, (2) microwave-based downlink SWIPT, and (3) microwave-based uplink information transmission. We further split the unit time slot into the corresponding three parts, denoting as $\varepsilon_t, \Delta_t, \delta_t$, with $\varepsilon_t + \Delta_t + \delta_t = 1$. It is worth

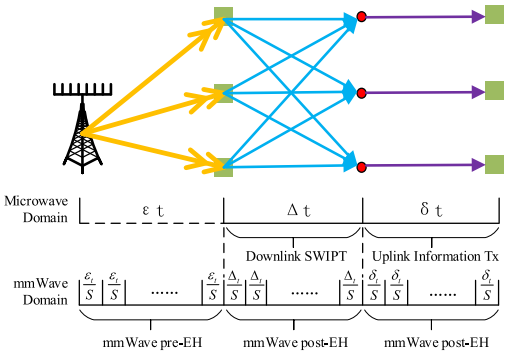


FIGURE 2. Time Division Hybrid EH Protocol.

mentioning that during phase Δ_t and δ_t, mmWave energy broadcasting is still working since it operates on a different carrier frequency, as illustrated in fig.2. Thus, ε_t is termed as 'mmWave pre-EH' phase, while Δ_t and δ_t are termed as 'mmWave post-EH/Downlink SWIPT' phase and 'mmWave post-EH/Uplink Information Tx' phase, respectively. In our model, mmWave EH is conducted in time-division manner, where both 'pre-EH' and 'post-EH' periods are equally splitted among MTCGs, namely $t_s = \frac{\varepsilon_t}{S} + \frac{\Delta_t}{S} + \frac{\delta_t}{S} \equiv \frac{1}{S}, \forall s \in \mathcal{S}$. Assume that MTCGs and MTCDs are initialized with the same battery energy $e_{s,0}$ and $e_{u,0}$, respectively. In microwave-based transmission, total spectrum of W^d is orthogonally splitted among the serving MTCDs of each MTCG in downlink, while W^u is used in uplink.

B. CHANNEL MODEL

In our model, received mmWave power level at MTCG s on the i-th time slot can be denoted as,

$$p^{(i)}_{mmW} = l\left(d_{m,s}\right) MNG_m G_s p^{(i)}_{m,s} \tag{1}$$

where $l\left(d_{m,s}\right)$ is the path loss and $d_{m,s}$ is the distance between transmit antenna and receive antenna of mmWave signal. Adopting LOS channel model in [17], where free space path loss is combined with an additional 25dB shadowing loss to account for possible obstruction, $l\left(d_{m,s}\right) = l_{free}\left(d_{m,s}\right) + 25dB$. According to [11], channel model of mmWave transmission consists of a main line of sight (LOS) channel and several non-line of sight (NLOS) channels. The LOS condition dominates the performance when a LOS channel does exist. Otherwise the NLOS channels take over. Although LOS channel is hard to maintain in mmWave transmission, to derive a compact and instructive analysis formula about the grid power consumption, we assume that the mmWave channel experiences LOS condition during the whole time (e.g. by mounting MTCG at rooftop). M and N are the dimension of Tx and Rx antenna array. G_m,G_s are the antenna gain at HAP and MTCG. $p^{(i)}_{m,s}$ is the mmWave transmit power level. Small-scale fading in mmWave transmission is neglected, due to the large antenna array effect [18].

On the i-th time slot, the received rate at MTCD u from MTCG s in downlink transmission can be estimated as,

$$r^{(i)}_{s,u} = \frac{W^d}{L}\log_2\left(1+\gamma^{(i)}_{s,u}\right), \gamma^{(i)}_{s,u} = \frac{(1-\sigma)\, p^{(i)}_{s,u} H^{(i)}_{s,u}}{\sum\limits_{c\in S, c\neq s} p^{(i)}_{c,u} H^{(i)}_{c,u} + \sigma_0{}^2} \tag{2}$$

and that in uplink transmission is,

$$r^{(i)}_{u,s} = \frac{W^u}{L}\log_2(1+\gamma^{(i)}_{u,s}), \gamma^{(i)}_{u,s} = \frac{p^{(i)}_{u,s} H^{(i)}_{u,s}}{\sum\limits_{c\in \mathcal{I}(u), c\neq u} p^{(i)}_{c,s} H^{(i)}_{c,s} + \sigma_0{}^2} \tag{3}$$

where $\gamma^{(i)}_{s,u}$ and $\gamma^{(i)}_{u,s}$ are the corresponding SINR values. $H^{(i)}_{s,u} = l\left(d_{s,u}\right) G_s G_u h^{(i)}_{s,u}$, $H^{(i)}_{u,s} = l\left(d_{u,s}\right) h^{(i)}_{u,s} G_u G_s$. $h^{(i)}_{s,u}$ and $h^{(i)}_{u,s}$ are the small-scale fading coefficients, and are random variables chosen from complex Gaussian distribution with zero mean and unit variance. $\mathcal{I}(u)$ is the set of MTCDs transmitting on the same subcarrier as MTCD u does. $p_{s,u}$ and $p_{u,s}$ are the downlink and uplink transmit power, respectively. σ_0^2 is variance of the additive white Gaussian noise (AWGN) with power spectral density of N_{noise}. Constant shadow fading of 10 dB is added to microwave channels. Free space path loss is calculated with exponent α_χ in $l_{free}(d) = d^{-\alpha_\chi}, \alpha_\chi \in \{\alpha_{mmW}, \alpha_{micro}\}$, where $\alpha_{mmW} \neq \alpha_{micro}$. Power-splitting SWIPT is applied in downlink transmission between MTCG and its serving MTCDs. A portion of $1-\sigma$ power is used by MTCG for information transmission as in [19], while the remaining σ power wirelessly charges MTCD. We assume that MTCDs have continuous data exchange with MTCG, thus a minimum transmission rate of r^d_{min} bits/s in downlink and r^u_{min} bits/s in uplink are required. In IoT applications, uplink sensor traffic generally outstrips downlink traffic, thus we have $r^d_{min} < r^u_{min}$.

III. PROBLEM FORMULATION

Our first concern is to minimize the average grid energy consumption [20] at HAP under moderate throughput constraints. Assume that HAP and the mmWave transceivers at HAP are always active, thus the circuit components and cooling system of which are consuming grid electricity the whole time [21]. P1 is then formulated as follows,

$$\min_{\mathbf{P_m},\mathbf{P_s},\mathbf{P_u},\mathbf{Q_s},\mathbf{Q_u}} \frac{1}{T}\sum_{i=1}^{T}\left(\Delta_p \sum_{s=1}^{S}\frac{1}{S}p^{(i)}_{m,s} + p_{mmW,c} + p_{m,c}\right)$$

$$s.t.\ e_{s,0} - \sum_{j=1}^{i-1} q^{(j)}_s + \frac{\eta\varepsilon_t H_{m,s}}{S} p^{(i)}_{m,s} \geq \Delta_t\left(\sum_{u\in\mathcal{U}(s)} p^{(i)}_{s,u} + p_{0,s}\right) \tag{4a}$$

$$e_{u,0} - \sum_{j=1}^{i-1} q^{(j)}_u + \sum_{s\in\mathcal{S}} \eta\Delta_t 1_\sigma p^{(i)}_{c,u} H^{(i)}_{c,u} \geq \delta_t\left(p^{(i)}_{u,s} + p_{0,u}\right) \tag{4b}$$

$$f^d\left(\mathbf{P}_{\mathbf{s},u}^{(i)}\right)=\left(2^{\frac{Lr_{\min}^d}{W^d}}-1\right)\sigma_0^2-\left(\left(1-\sigma\right)H_{s,u}^{(i)}\right)p_{s,u}^{(i)}+\sum_{c\neq s,c\in\mathcal{S}}\left(\left(2^{\frac{Lr_{\min}^d}{W^d}}-1\right)H_{c,u}^{(i)}\right)p_{c,u}^{(i)},$$

$$f^u\left(\mathbf{P}_{\mathbf{u},s}^{(i)}\right)=\left(2^{\frac{Lr_{\min}^u}{W^u}}-1\right)\sigma_0^2-H_{u,s}^{(i)}p_{u,s}^{(i)}+\sum_{c\neq u,c\in\mathcal{I}(u)}\left(\left(2^{\frac{Lr_{\min}^u}{W^u}}-1\right)H_{c,s}^{(i)}\right)p_{c,s}^{(i)}$$

$$\sum_{i=1}^{\tau}q_s^{(i)}\leq e_{s,0},\quad\forall s\in\mathcal{S},1\leq\tau\leq T \tag{4c}$$

$$\sum_{i=1}^{\tau}q_u^{(i)}\leq e_{u,0},\quad\forall u\in\mathcal{U}\left(s\right),s\in\mathcal{S},1\leq\tau\leq T \tag{4d}$$

$$q_s^{(i)}\in R,q_u^{(i)}\in R,\quad\forall u\in\mathcal{U}\left(s\right),s\in\mathcal{S},1\leq i\leq T \tag{4e}$$

$$f^d\left(\mathbf{P}_{\mathbf{s},u}^{(i)}\right)\leq 0,f^u\left(\mathbf{P}_{\mathbf{u},s}^{(i)}\right)\leq 0,\ \forall u\in\mathcal{U},s\in\mathcal{S},1\leq i\leq T \tag{4f}$$

$$p_{m,s}^{(i)}\in\left(0,p_{m,\max}\right],p_{s,u}^{(i)}\in\left(0,p_{s,\max}\right],\ \forall u\in\mathcal{U}\left(s\right),s\in\mathcal{S} \tag{4g}$$

$$p_{u,s}^{(i)}\in\left[p_{u,\min},p_{u,\max}\right],\quad\forall u\in\mathcal{U}\left(s\right),s\in\mathcal{S} \tag{4h}$$

where Δ_p is the slope of the variable power consumption at HAP, which is $\sum_{s\in\mathcal{S}}\frac{1}{S}p_{m,s}^{(i)}$, following the linear power model of BS in [22]. $H_{m,s}=MNl\left(d_{m,s}\right)G_mG_s$, $1_\sigma=\begin{cases}\sigma, & s\\ 1, & c\in\mathcal{S},c\neq s\end{cases}$. Here, we introduce,

$$q_u^{(i)}=\delta_t\left(p_{u,s}^{(i)}+p_{0,u}\right)-\eta\Delta_t\sigma p_{s,u}^{(i)}H_{s,u}^{(i)}-\sum_{\substack{c\in\mathcal{S}\\c\neq s}}\eta\Delta_t p_{c,u}^{(i)}H_{c,u}^{(i)} \tag{5}$$

$$q_s^{(i)}=\Delta_t\left(\sum_{u\in\mathcal{U}(s)}p_{s,u}^{(i)}+p_{0,s}\right)-\frac{\eta H_{m,s}}{S}p_{m,s}^{(i)} \tag{6}$$

as auxiliary variables recording battery dynamics at MTCG and MTCD respectively. $\mathbf{P_m}$ is the collection of $p_{m,s}^{(i)}$ in terms of system operation time T and MTCG set $\mathcal{S}$, $\mathbf{P_s}$ is the three-dimensional matrix containing $p_{s,u}^{(i)}$ in terms of T, $\mathcal{S}$ and MTCD set $\mathcal{U}$, $\mathbf{P_u}$ is also three-dimensional matrix with element $p_{u,s}^{(i)}.f^d\left(\mathbf{P}_{\mathbf{s},u}^{(i)}\right)$ and $f^u\left(\mathbf{P}_{\mathbf{u},s}^{(i)}\right)$ are QoS functions expanded at the top of this page, where $\mathbf{P}_{\mathbf{s},u}^{(i)}$ is downlink power vector along dimension $\mathcal{S}$ at the i-th time slot towards MTCD u and $\mathbf{P}_{\mathbf{u},s}^{(i)}$ is uplink power vector along dimension $\mathcal{U}$ at the i-th time slot towards MTCG s. $\mathbf{Q_s}$, $\mathbf{Q_u}$ are collection of $q_s^{(i)}$ and $q_u^{(i)}$, respectively. $\mathcal{U}(s)$ is the set of MTCDs allocated to MTCG s. $p_{mmW,c}$ is the static power consumption of mmWave transceiver, while $p_{m,c}$ is the static power consumption of HAP. Constraints (4b),(4c) are the energy causality constraints, which states that harvested energy at MTCG and MTCD must exceed the consumed energy [23]. (4b) holds for $\forall s\in\mathcal{S},1\leq i\leq T$, and (4c) holds for $\forall u\in\mathcal{U}\left(s\right),s\in\mathcal{S},1\leq i\leq T$. From constraint (4d) and (4e), at the end of each time slot, the battery of MTCG and MTCD is supposed to be not drained. $p_{s,c}$ and $p_{u,c}$ are circuit power consumption at MTCG and MTCD respectively. Constraints (4g) are the throughput QoS constraints. (4h),(4i) are constraints on transmitting power level at HAP, MTCG and MTCD, respectively. We assume that transceivers at MTCG and MTCD are extremely sensitive (e.g. wake-up radio receiver in [24] with sensitivity of -43 dBm, and RF energy harvester with input power levels as small as -39 dBm [25]), thus HAP and MTCG can transmit at very low power level ($\min p_{m,s}^{(i)},p_{s,u}^{(i)}\to 0$).

IV. PROBLEM SOLUTION

According to [26], the distance of points of a Gin-DPP from the origin have the same distribution as independent gamma random variables. Let $\{x_1,\ldots,x_S\}$ be S points distributed according to Gin-DPP in region $\mathcal{A}$. Then $\{\|x_1\|,\ldots,\|x_S\|\}$ has the same distribution as $\{Y_1,\ldots,Y_S\}$, where $Y_s^2\sim gamma\left(s,1\right),1\leq s\leq S$. Then we can derive the probability distribution function (pdf) of $\|x_s\|$ as $p\left(\|x_s\|\right)=\frac{2\|x_s\|}{\Gamma(s)}\|x_s\|^{2(s-1)}e^{-\|x_s\|^2},\forall 1\leq s\leq S$, where $\Gamma\left(s\right)$ is the gamma function, $\Gamma\left(s\right)=\int_0^{+\infty}x^{s-1}e^{-x}dx,s>0$. When path loss component of the channel between HAP and MTCG s is α_s, the energy harvesting rate is defined as $e_s=\eta t_sK_sP_s\|x_s\|^{-a_s},\forall s\in\mathcal{S}$, where η is energy conversion efficiency, t_s is the energy harvesting time, $K_s=MNG_mG_s$, $p_{m,s}$ is the mmWave transmitting power at HAP towards MTCG s. We then give proposition 1,

Proposition 1: The expectation of mmWave energy harvesting rate at receiver antenna of MTCG s ($s>\alpha_s/2$) following Gin-DPP is:

$$\mathbb{E}\left[e_s\right]=\frac{\eta}{S}MNG_mG_sp_{m,s}\frac{\Gamma\left(s-\alpha_s/2\right)}{\Gamma\left(s\right)} \tag{7}$$

Proof:

According to the definition of expectation,

$$\begin{aligned}\mathbb{E}\left[e_s\right]&=\eta t_sK_sp_{m,s}\cdot\mathbb{E}\left[\|x_s\|^{-\alpha_s}\right]\overset{z_s=\|x_s\|^{-\alpha_s}}{\to}\eta t_sK_sp_{m,s}\cdot\mathbb{E}\left[z_s\right]\\&=\eta t_sK_sp_{m,s}\int_0^{+\infty}z_s\frac{2}{\alpha_s\Gamma\left(s\right)}z_s^{-\frac{2}{\alpha_s}-1}e^{-z_s^{-\frac{2}{\alpha_s}}}dz_s\\&\overset{v_s=z_s^{-\frac{2}{\alpha_s}}}{\to}\eta t_sK_sp_{m,s}\int_0^{+\infty}\frac{1}{\Gamma\left(s\right)}v_s^{s-\frac{\alpha_s}{2}-1}e^{-v_s}dv_s\\&=\frac{\eta}{S}MNG_mG_sp_{m,s}\frac{\Gamma\left(s-\alpha_s/2\right)}{\Gamma\left(s\right)}\Big|_{t_s=\frac{1}{S}}\end{aligned}$$

Here, we choose $s > \alpha_s/2$ since the gamma function is defined for all complex numbers except the non-positive integers. Superscript (i) is neglected in $p_{m,s}$ to simplify the notation. This completes the proof. ■

In our model, we propose two mmWave EH modes:

1) Mode 1: For each MTCG, HAP transmits mmWave energy with the same power level during the whole time.
2) Mode 2: For each MTCG, no equalization on transmitting power level is required at HAP.

Since there is no simple analytic expression of the solution of LP problem, to derive a closed-form representation, we consider the special case where initial battery energy at MTCD is abundant. According to the definition, $p_{m,s}^{(i)}$ is a monotonically decreasing function of $q_s^{(i)}$, while $p_{s,u}^{(i)}$ is also a monotonically decreasing function of $q_u^{(i)}$. This helps us convert the original minimization problem of $\mathbf{P_m}$,$\mathbf{P_s}$,$\mathbf{P_u}$ into maximization problem of $\mathbf{Q_s}$, $\mathbf{Q_u}$.

A. MAXIMIZATION OF $\mathbf{Q_u}$

We first give the optimum moments of $q_u^{(i)}, \forall u \in \mathcal{U}(s)$, $s \in \mathcal{S}, 1 \le i \le T$. According to the definition of $q_u^{(i)}$, we have $\min \max_i p_{s,u}^{(i)} \Leftrightarrow \min \max_i \left(\delta_t \left(p_{u,s}^{(i)} + p_{0,u}\right) - q_u^{(i)}\right) \Leftrightarrow \begin{cases} \max \min_i q_u^{(i)} \\ \min \max_i p_{u,s}^{(i)} \end{cases}$. When $r_{\min}^u$ is a *moderate* requirement ($\min \max_i p_{u,s}^{(i)}\left(r_{\min}^u\right) \le p_{u,\min}$), we can say $p_{u,s}^{(i)*} = p_{u,\min}$. From (4d), we have $\mathbf{1}_{\Delta_{lwr}}\mathbf{Q}_u \le e_{u,0}$, where $\mathbf{1}_{\Delta_{lwr}}$ is the lower triangular part of all-1 matrix. Thus by solving the following problem,

$$\max \min \ \mathbf{Q}_u \tag{8a}$$

$$s.t.\ \mathbf{1}_{\Delta_{lwr}}\mathbf{Q}_u \le e_{u,0} \tag{8b}$$

We can obtain $q_u^{(i)*} = \frac{e_{u,0}}{T}$. Substituting $q_u^{(i)*}$ into $\min \max_i \left(\delta_t \left(p_{u,s}^{(i)} + p_{0,u}\right) - q_u^{(i)}\right)$, together with $p_{s,u}^{(i)} > 0$, we have $\delta_t\left(p_{u,\min} + p_{0,u}\right) > \frac{e_{u,0}}{T}$. When $r_{\min}^d$ is also a *moderate* requirement ($\min \max_i p_{s,u}^{(i)}\left(r_{\min}^d\right) \to 0$), we finally have $e_{u,0}{}^* \approx T\delta_t\left(p_{u,\min} + p_{0,u}\right)$. The specific definition of $r_{\min}^u$ and $r_{\min}^d$ being *moderate* are described in Appendix A.

Thus, we can divide the initial battery energy level at MTCD into two stages: (1) *abundantly-charged*: when $e_{u,0} > T\delta_t\left(p_{u,\min} + p_{0,u}\right)$, (2) *undercharged*: $e_{u,0} < T\delta_t\left(p_{u,\min} + p_{0,u}\right)$.

B. MAXIMIZATION OF $\mathbf{Q_s}$

Next, we give the optimum moments of $q_s^{(i)}, \forall s \in \mathcal{S}, 1 \le i \le T$. In Mode 1, we have constraint $p_{m,s}^{(i)} = p_{m,s}^{(j)}, 1 \le i < j \le T$. According to the definition, $q_s^{(j)}$ can be represented in terms of $q_s^{(i)}$. Setting $i = 1$, we have,

$$q_s^{(j)} = q_s^{(1)} - \Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(1)}\right) + \Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(j)}\right) \tag{9}$$

Since we are aiming at minimizing $\frac{\Delta_p}{ST}\sum_{s=1}^{S}\sum_{i=1}^{T} p_{m,s}^{(i)}$, it is equaling to minimizing $p_{m,s}^{(1)}$. Thus $p_{s,u}^{(1)}$ must be approaching zero, we then have $q_s^{(j)*} \approx q_s^{(1)*} + \Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(j)}\right)$. Substituting (4d) into (4b) , condition

$$q_s^{(i)} \le \left(1 - \frac{1}{\varepsilon_t}\right)\left(\Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(i)} + p_{0,s}\right)\right) + \frac{1}{\varepsilon_t}e_{s0} - \frac{1}{\varepsilon_t}\sum_{j=1}^{i-1} q_s^{(j)}, 1 \le i \le T \tag{10}$$

holds. Together with (9), we have $q_s^{(1)*} \approx \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{T - 1 + \varepsilon_t}$. From (10), when setting $i = 1$, we can deduce another optimum moment of $q_s^{(1)}$ as $q_s^{(1)*} \approx \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{\varepsilon_t}$. According to (9) and (4d), we can also obtain the optimum value $q_s^{(1)*} = \frac{e_{s,0}}{T}$. Therefore, we can say,

$$q_s^{(1)*} = \min\left\{\frac{e_{s,0}}{T}, \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{T - 1 + \varepsilon_t}, \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{\varepsilon_t}\right\} \tag{11}$$

To obtain a compact solution, from now on, we replace '≈' with '=' to study the asymptotical performance. The following remarks can be made,

Remark 1: when there is sufficient initial battery energy at MTCG ($e_{s,0} > (1 - \varepsilon_t)\Delta_t p_{0,s}$), the optimal $q_s^{(1)}$ is,

$$q_s^{(1)*} = \begin{cases} \dfrac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{T - 1 + \varepsilon_t}, & T > \dfrac{e_{s,0}}{\Delta_t p_{0,s}} \\ \dfrac{e_{s,0}}{T}, & T < \dfrac{e_{s,0}}{\Delta_t p_{0,s}} \end{cases} \tag{12}$$

Remark 2: when there is insufficient initial battery energy at MTCG ($e_{s,0} < (1 - \varepsilon_t)\Delta_t p_{0,s}$), the optimal $q_s^{(1)}$ is,

$$q_s^{(1)*} = \begin{cases} \dfrac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s0}}{\varepsilon_t}, & T > \dfrac{\varepsilon_t e_{s,0}}{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}} \\ \dfrac{e_{s,0}}{T}, & T < \dfrac{\varepsilon_t e_{s,0}}{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s0}} \end{cases} \tag{13}$$

Since $(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0} \le 0$ and $T \ge 1$ always hold, we have $q_s^{(1)*} = \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{\varepsilon_t}, \forall T$.

We can then divide initial battery energy level at MTCG into three stages:(1) *undercharged*: $0 \le e_{s,0} < (1 - \varepsilon_t)\Delta_t p_{0,s}$, (2) *properly-charged*: $(1 - \varepsilon_t)\Delta_t p_{0,s} < e_{s,0} < T\Delta_t p_{0,s}$, (3) *overcharged*: $T\Delta_t p_{0,s} < e_{s,0}$.

$$\kappa_1 = \sum_{s\in\mathcal{S}} \frac{\Delta_p}{\eta H_{m,s}}, \quad \kappa_2 = (p_{mmW,c} + p_{m,c})\varepsilon_t, \quad \kappa_3 = \frac{T^2 - T + T\varepsilon_t}{T\Delta_t p_{0,s} - e_{s,0}} \left(p_{mmW,c} + p_{m,c}\right)$$

$$f_j\left(\mathbf{P}_s, \mathbf{Q}_s\right) = \begin{cases} F^{(j,:)}\left(\mathbf{P}_s\mathbf{I} + p_{0,s}\right) + E^{(j,:)}\mathbf{Q}_s - e_{s,0} & , 1 \le j \le T \\ B^{(j-T,:)}\mathbf{Q}_s - e_{s,0} & , T+1 < j \le 2T, \end{cases}, \quad \forall s \in \mathcal{S}.$$

$$A = \delta_t\left(p_{u,\min} + p_{0,u}\right), \quad B = \mathbf{1}_{\Delta_{lwr}}, \qquad C_{s,u} = \eta\Delta_t H_{s,u}^{(i)}, \forall s \in \mathcal{S}, u \in \mathcal{U}$$

$$D(u) = \begin{bmatrix} \sigma C_{1,\mathcal{I}(u)_1} & C_{2,\mathcal{I}(u)_1} & \cdots & C_{S,\mathcal{I}(u)_1} \\ C_{1,\mathcal{I}(u)_2} & \sigma C_{2,\mathcal{I}(u)_2} & \cdots & C_{S,\mathcal{I}(u)_2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1,\mathcal{I}(u)_S} & C_{2,\mathcal{I}(u)_S} & \cdots & \sigma C_{S,\mathcal{I}(u)_S} \end{bmatrix}_{S\times S}, E = \begin{bmatrix} \varepsilon_t & 0 & \cdots & 0 \\ 1 & \varepsilon_t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & \varepsilon_t \end{bmatrix}_{T\times T}, F = \mathbf{diag}((1-\varepsilon_t)\,\Delta_t)_{T\times T}$$

C. CLOSED-FORM RELATIONSHIP BETWEEN EH MODES

Proposition 2: When throughput requirements are modest and battery at MTCG is undercharged, the average grid energy consumption under Mode 1 has constant ratio relationship with that under Mode 2, which is $\frac{\kappa_1(\Delta_t p_{0,s} - e_{s,0}) + \kappa_2}{\kappa_1\varepsilon_t\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right) + \kappa_2}$. $\frac{\Delta_t p_{0,s} - e_{s,0}}{\varepsilon_t\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)}$ is the ratio between variable power consumption; when battery at MTCG is properly-charged, this ratio remains $\frac{\kappa_1 T + \kappa_3}{\kappa_1(T-1+\varepsilon_t)+\kappa_3}$, no matter how $e_{s,0}$ changes. $\frac{T}{T-1+\varepsilon_t}$ is the ratio between variable power consumption. $\kappa_1, \kappa_2, \kappa_3$ are expanded at the top of this page.

Proof: (1) *Undercharged*: We first look into Mode 1. Since we are aiming at minimizing $p_{m,s}^{(1)}$, from the definition of $q_s^{(i)}$ and (8), we have to simultaneously minimize $\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(1)}$ and maximize $q_s^{(1)}$. Thus we can have,

$$p_{m,s}^{(1)*} = \frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - q_s^{(1)}\right) \tag{14a}$$

$$= \frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - \left(1 - \frac{1}{\varepsilon_t}\right)\Delta_t p_{0,s} - \frac{1}{\varepsilon_t} e_{s,0}\right) \tag{14b}$$

$$= \frac{1}{\varepsilon_t}\frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - e_{s,0}\right) \tag{14c}$$

and $\min \frac{1}{T}\sum_{i=1}^{T} p_{m,s}^{(i)} = \frac{1}{T}\sum_{i=1}^{T} p_{m,s}^{(1)*} = p_{m,s}^{(1)*}$.

Switch to Mode 2. Again we minimize $\frac{\Delta_p}{ST}\sum_{s=1}^{S}\sum_{i=1}^{T} p_{m,s}^{(i)}$. According to the definition of $q_s^{(i)}$, we have, $\min \frac{1}{T}\sum_{i=1}^{T} p_{m,s}^{(i)} = \min \frac{1}{T}\sum_{i=1}^{T}\left(\frac{S}{\eta H_{m,s}}\left(\Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(i)} + p_{0,s}\right) - q_s^{(i)}\right)\right)$. Since $e_{s,0} - \sum_{i=1}^{T} q_s^{(i)*} = 0$ in *undercharged* condition to utilize the battery energy to the most extent, together with $p_{s,u}^{(i)*} = \min p_{s,u}^{(i)} \approx 0$, $1 \le i \le T$, we have,

$$\min \frac{1}{T}\sum_{i=1}^{T} p_{m,s}^{(i)}$$

$$= \frac{S}{T}\frac{\left(\sum_{i=1}^{T}\left(\Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(i)*} + p_{0,s}\right)\right) - e_{s,0}\right)}{\eta H_{m,s}} \tag{15a}$$

$$= \frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right) \tag{15b}$$

Combining (14c) and (15b), together with the static power consumption of system, the ratio can be derived.

(2) *Properly-charged*: Under Mode 1, according to (10),

$$p_{m,s}^{(1)*} = \begin{cases} \frac{S\left(\Delta_t p_{0,s} - \frac{(\varepsilon_t - 1)\Delta_t p_{0,s} + e_{s,0}}{T-1+\varepsilon_t}\right)}{\eta H_{m,s}} & , T > \frac{e_{s,0}}{\Delta_t p_{0,s}} \\ 0 & , T \le \frac{e_{s,0}}{\Delta_t p_{0,s}} \end{cases} \tag{16}$$

where $0 = \max\left\{0, \frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)\right\}, T \le \frac{e_{s,0}}{\Delta_t p_{0,s}}$. Under Mode 2,

$$\min \frac{1}{T}\sum_{i=1}^{T} p_{m,s}^{(i)} = \begin{cases} \frac{S\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)}{\eta H_{m,s}} & , T > \frac{e_{s,0}}{\Delta_t p_{0,s}} \\ 0 & , T \le \frac{e_{s,0}}{\Delta_t p_{0,s}} \end{cases} \tag{17}$$

where $0 = \max\left\{0, \frac{S}{\eta H_{m,s}}\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)\right\}, T \le \frac{e_{s,0}}{\Delta_t p_{0,s}}$. Combining (16) and (17), together with the static power consumption of system, the ratio can be derived.

The proof is complete. ■

D. NUMERICAL SOLUTION USING INTERIOR-POINT METHOD

To obtain the numerical result, we apply logarithmic barrier interior-point method to solve P1. We first transform P1 into

unconstrained problem P2 $= f(\mathbf{P_s}, \mathbf{Q_s}, r)$ as,

$$\min_{\mathbf{P_s},\mathbf{Q_s},r} f_0(\mathbf{P_s},\mathbf{Q_s}) + \sum_{s=1}^{S}\sum_{j=1}^{2T} -r\log\left(-f_j(\mathbf{P}_s,\mathbf{Q}_s)\right) + \sum_{u=1}^{U}\sum_{j=1}^{T} -r\log\left(-g_j(\mathbf{P}_u)\right) \quad (18)$$

where, $f_0(\mathbf{P_s},\mathbf{Q_s}) = \frac{1}{T}\sum_{s=1}^{S}\sum_{i=1}^{T}\frac{\Delta_t\left(\sum_{u\in\mathcal{U}(s)} p_{s,u}^{(i)}+p_{0,s}\right)-q_s^{(i)}}{\eta H_{m,s}}$, and $f_j(\mathbf{P}_s,\mathbf{Q}_s)$ is expanded at the top of the previous page. $\mathbf{I}$ is the all-one column vector.

$g_j(\mathbf{P}_u) = B^{(j-2T,:)}\left(A-(D(u)\mathbf{P}_u)^{(s(u),:)}\right)^T - e_{u,0}, 1 < j \le T, \forall u \in U$. $\mathcal{I}(u)_s$ returns the s-th element in set $\mathcal{I}(u)$. $s(u)$ returns the MTCG id to which MTCD u is attached. A, B, C, D, E, F are listed at the bottom of this page. Here, $\mathbf{P}_s = \left[\mathbf{P}_{s,1}, \mathbf{P}_{s,2}, \ldots, \mathbf{P}_{s,L}\right]$, $\mathbf{P}_u = \left[\mathbf{P}_{1,u}, \mathbf{P}_{2,u}, \ldots, \mathbf{P}_{S,u}\right]^T$, $\mathbf{P}_{s,l} = \left[p_{s,l}^{(1)}, p_{s,l}^{(2)}, \ldots, p_{s,l}^{(T)}\right]^T, s \in \mathcal{S}, 1 \le l \le L$, $\mathbf{Q}_s = \left[q_s^{(1)}, q_s^{(2)}, \ldots, q_s^{(T)}\right]^T, s \in \mathcal{S}$. $(\cdot)^T$ is the transpose operator. $(\cdot)^{(a,:)}$ and $(\cdot)^{(:,a)}$ are the operators which abstract the a-th row and a-th column of a matrix, respectively. r is iteratively updated with $r^{(k)} = \mu r^{(k-1)}$, $\mu < 1$, and $\mathbf{P_s}, \mathbf{Q_s}$ are updated using Newton's method.

E. COMPUTATIONAL COMPLEXITY

Applying the barrier method, the duality gap after the initial centering step, and k additional centering steps, is $(2ST+UT)\mu^k r^{(0)}$. Therefore the desired accuracy ρ is achieved after exactly $\left\lceil \frac{\log\left((2ST+UT)/(\rho r^{(0)})\right)}{\log\mu} \right\rceil$ centering steps, plus the initial centering step. Thus the complexity is $\mathcal{O}\left(\left\lceil \frac{f\left(\mathbf{P_s}^{(0)},\mathbf{Q_s}^{(0)},r\right)-p^*}{\alpha\beta\|\nabla f\|_2^2(m/M^2)} + 6 \right\rceil \left\lceil \frac{\log\left((2ST+UT)/(\rho r^{(0)})\right)}{\log\mu} \right\rceil\right)$, where $mI \prec \nabla^2 f \prec MI$ and α, β are the parameters of backtracking line search in Newton's method.

V. ENERGY EFFICIENCY MAXIMIZATION

The other goal of our work is to maximize the grid energy efficiency of the whole system under proposed EH modes during operation time of T seconds. The EE expression adopted in our analysis is:

$$EE_{var} = \frac{\sum_{i=1}^{T}\sum_{s=1}^{S}\sum_{u\in\mathcal{U}(s)} r_{s,u}^{(i)} + \sum_{i=1}^{T}\sum_{s=1}^{S}\sum_{u\in\mathcal{U}(s)} r_{u,s}^{(i)}}{(1/S)\sum_{t=1}^{T}\sum_{s=1}^{S} p_{m,s}^{(i)}}, \quad (19)$$

which denotes the energy efficiency in terms of variable grid power consumption. Obviously, EE_{var} is not a convex function of $\mathbf{P_m}, \mathbf{P_s}, \mathbf{P_u}$, due to its fractional form and co-channel interference within 'Downlink SWIPT' and 'Uplink information Tx' phases. Therefore, we first transform EE maximization problem into its linear form by Dinkelbach's method [27] as:

$$P3: \max_{\substack{\mathbf{P_m},\mathbf{P_s},\mathbf{P_u}\\ \mathbf{Q_s},\mathbf{Q_u},\mu}} \left(\sum_{i=1}^{T}\sum_{s=1}^{S}\sum_{u\in\mathcal{U}(s)} r_{s,u}^{(i)} + \sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{u\in\mathcal{U}(s)} r_{u,s}^{(i)}\right) - \mu\left(\frac{1}{S}\sum_{t=1}^{T}\sum_{s=1}^{S} p_{m,s}^{(i)}\right) \quad (20a)$$

$$s.t. \quad (4b)-(4f), (4h), (4i) \quad (20b)$$

here, μ is the tuning parameter iteratively updated to the optimal value of EE. By solving P3, the updating rule can be written as,

$$\mu^{(k+1)} = EE_{var}\left(r_{s,u}^{(i),*}, r_{u,s}^{(i),*}, p_{m,s}^{(i),*} \middle| \mu^{(k)}\right) \quad (21)$$

where k is the iteration index and $r_{s,u}^{(i),*}, r_{u,s}^{(i),*}, p_{m,s}^{(i),*}$ are the optimum downlink throughput, uplink throughput and mmWave transmit power level when $\mu = \mu^{(k+1)}$, respectively. The iteration process terminates when condition $\left|\mu^{(k+1)} - \mu^{(k)}\right| \le \varepsilon_\mu$ is satisfied. ε_μ is the predefined convergence threshold. It is worth noting that, QoS constraints are removed in P3 to investigate the global optimum throughput at which EE is maximized, no matter what the exact service type is. This can also help us optimize service level configuration of the proposed cellular M2M system.

A. THROUGHPUT CONVEX APPROXIMATION

However, non-convexity still exists after linearization of our adopted EE expression. According to [28], the interference-limited throughput maximization problem is proved to be NP-hard, thus global optimum is difficult to be found. To transform P3 into a tractable form, we adopt the popular SCA method. Firstly, the interference-limited downlink throughput is originally written in terms of SINR as in (2). From [29], $\log(1+x)$ can be can be well approximated with $a\log(x)+b$, where $a = \frac{x_0}{1+x_0}$, $b = \log(1+x_0) - \frac{x_0}{1+x_0}\log(x_0)$. x_0 here is the point around which $\log(1+x)$ is approximated. Through this manipulation, downlink throughput can be approximated with,

$$\hat{r}_{s,u}^{j,(i)} = \frac{W^d}{L}\left(a^{j,(i)}\log_2\left(\gamma_{s,u}^{(i)}\right) + b^{j,(i)}\right) \quad (22)$$

where,

$$a_{s,u}^{j,(i)} = \frac{\sqrt{1-\sigma}\bar{p}_{s,u}^{j-1,(i)}H_{s,u}^{(i)}}{\sum_{\substack{c\in S\\ c\neq s}}\bar{p}_{c,u}^{j-1,(i)}H_{c,u}^{(i)} + {\sigma_0}^2 + \sqrt{1-\sigma}\bar{p}_{s,u}^{j-1,(i)}H_{s,u}^{(i)}} \quad (23)$$

$$b_{s,u}^{j,(i)} = \log_2\left(1 + \frac{\sqrt{1-\sigma}\bar{p}_{s,u}^{j-1,(i)}H_{s,u}^{(i)}}{\sum_{c\in S, c\neq s}\bar{p}_{c,u}^{j-1,(i)}H_{c,u}^{(i)} + {\sigma_0}^2}\right) - a_{s,u}^{j,(i)}\log_2\left(\frac{\sqrt{1-\sigma}\bar{p}_{s,u}^{j-1,(i)}H_{s,u}^{(i)}}{\sum_{c\in S, c\neq s}\bar{p}_{c,u}^{j-1,(i)}H_{c,u}^{(i)} + {\sigma_0}^2}\right) \quad (24)$$

The superscript j denotes the j-th convex approximation process and $\bar{p}_{s,u}^{j-1,(i)}$ is the optimum solution of the previous approximated convex problem.

Furthermore, we eliminate non-convexity in $\gamma_{s,u}^{(i)}$ by introducing auxiliary variables $v_{s,u}^{(i)} = log_2 p_{s,u}^{(i)}$. $\hat{r}_{s,u}^{j,(i)}$ is then written in terms of $v_{s,u}^{(i)}$ as:

$$\hat{r}_{s,u}^{j,(i)} = \frac{W^d}{L} a^{(i)} \left(\log_2 \left(\sqrt{1-\sigma} H_{s,u}^{(i)} \right) + v_{s,u}^{(i)} \right) - \frac{W^d}{L} a^{(i)} \log_2 \left(\sum_{\substack{c \in S \\ c \neq s}} H_{c,u}^{(i)} 2^{v_{c,u}^{(i)}} + {\sigma_0}^2 \right) + \frac{W^d}{L} b^{(i)} \tag{25}$$

The resulting function is a concave function of $v_{s,u}^{(i)}$, since the first term is a linear function, while the second term is logarithm of a sum of exponential functions. According to [29], the concavity is proved. Similarly, uplink throughput can also be approximated through SCA with $\hat{r}_{u,s}^{j,(i)}$ at $a_{u,s}^{j,(i)}$, $b_{u,s}^{j,(i)}$, $v_{u,s}^{(i)}$ and is not expanded here for the sake of conciseness. Since we adopt time division multiplexing, the channel state can be estimated according to reciprocity, therefore we have $\mathbb{E}\left[H_{s,u}^{(i)}\right] = \mathbb{E}\left[H_{u,s}^{(i)}\right], \forall s \in \mathcal{S}, u \in \mathcal{U}, 1 \leq i \leq T$.

B. NON-ORTHOGONAL UPLINK TRANSMISSION

To further promote the EE performance, we also investigate how non-orthogonal multi-access (NOMA) mechanism helps realize an energy-efficient cellular M2M system. In this section, we apply power-domain NOMA and see how it helps improve system EE performance through uplink transmission. Receiver antennas at MTCG s apply successive interference canceling (SIC) to decode information from superposition signals [4]. To perform high quality SIC decoding, the transmitting antenna set must be constrained at a suitable scale [5]. Usually two to four MTCDs are favorable for both energy efficiency and decoding correctness. $U_s = \sum_{s'=1}^{s} \left|\mathcal{U}(s')\right|$ is the index of the last MTCD in set $\mathcal{U}(s)$, and $|\mathcal{U}(s)|$ is the cardinality of set $\mathcal{U}(s)$. Channel state between MTCDs and the associated MTCG is estimated and sorted in ascending order as $H_{U_{s-1}+1,s}^{(i)} \leq H_{U_{s-1}+2,s}^{(i)} \leq \ldots \leq H_{U_s,s}^{(i)}$. Applying power-domain NOMA, the uplink throughput can be written as,

$$\begin{aligned}
&\sum_{i=1}^{T} \sum_{s=1}^{S} \sum_{u=U_{s-1}+1}^{U_s} r_{u,s}^{(i)} \\
&= \sum_{i=1}^{T} \sum_{s=1}^{S} \sum_{u=U_{s-1}+1}^{U_s} W \log_2 \left(1 + \frac{p_{u,s}^{(i)} H_{u,s}^{(i)}}{\sum_{u=U_{s-1}+1}^{u-1} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2} \right) \\
&= \sum_{i=1}^{T} \sum_{s=1}^{S} \sum_{=U_{s-1}+1}^{U_s} W \log_2 \left(\frac{\sum_{u=U_{s-1}+1}^{u} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2}{\sum_{u=U_{s-1}+1}^{u-1} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2} \right) \\
&= \sum_{i=1}^{T} \sum_{s=1}^{S} W \log_2 \left(\prod_{u=U_{s-1}+1}^{U_s} \left(\frac{\sum_{u=U_{s-1}+1}^{u} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2}{\sum_{u=U_{s-1}+1}^{u-1} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2} \right) \right) \\
&= \sum_{i=1}^{T} \sum_{s=1}^{S} W \log_2 \left(\frac{\sum_{u=U_{s-1}+1}^{U_s} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2}{{\sigma_0}^2} \right)
\end{aligned} \tag{26}$$

we can then transform it into a concave function of $p_{u,s}^{(i)}$,

$$W^u \sum_{i=1}^{T} \sum_{s=1}^{S} \left(\log_2 \left(\sum_{u=U_{s-1}+1}^{U_s} p_{u,s}^{(i)} H_{u,s}^{(i)} + {\sigma_0}^2 \right) - \log_2 \left({\sigma_0}^2 \right) \right) \tag{27}$$

C. THE NESTED CONVEX APPROXIMATION ALGORITHM

A model with nonlinear equality constraints is not convex [29], since the feasible region may not be convex set. Therefore, we relax equality constraints into inequality constraints to obtain the convex approximated problem of P3, which is written in P4 as follows,

$$\max_{\substack{\mathbf{P_m},\mathbf{P_s},\mathbf{V_s},\mathbf{P_u} \\ \mathbf{Q_s},\mathbf{Q_u},\mu}} \left(\sum_{i=1}^{T} \sum_{s=1}^{S} \sum_{u \in \mathcal{U}(s)} \hat{r}_{s,u}^{(i)} + \sum_{i=1}^{T} \sum_{s=1}^{S} \sum_{u \in \mathcal{U}(s)} \hat{r}_{u,s}^{(i)} \right) - \mu^{(k)} \left(p_{mmW,c} + p_{m,c} + (1/S) \sum_{i=1}^{T} \sum_{s=1}^{S} p_{m,s}^{(i)} \right) \tag{28a}$$

$$s.t.\ (4b) - (4f), (4h), (4i) \tag{28b}$$

$$2^{\mathbf{V_s}} \leq \mathbf{P_s}, 2^{\mathbf{V_u}} \leq \mathbf{P_u} \tag{28c}$$

Through the maximization process, $\mathbf{P_s}$ and $\mathbf{P_u}$ can be minimized to having a small gap against $2^{\mathbf{V_s}}$ and $2^{\mathbf{V_u}}$. To further narrow down this gap, we propose a small trick by first projecting optimum result from P4 onto $\mathbf{V_s}$ and $\mathbf{V_u}$ domain as the starting point, and then iteratively applying first order Taylor expansion to replace $p_{s,u}^{(i)}$ with $v_{s,u}^{(i)}$ around $v_{s,u}^{(i),*-}$ and replace $p_{u,s}^{(i)}$ with $v_{u,s}^{(i)}$ around $v_{u,s}^{(i),*-}$, where $v_{s,u}^{(i),*-}$ and $v_{u,s}^{(i),*-}$ are the optimum results obtained from solving P4 in the last iteration. The gap is then measured in convexified square error and iteratively minimized with lower upper bound.

Finally, the resulting problem is a concave problem of auxiliary variable $\mathbf{V_s}$ and $\mathbf{V_u}$, thus can be easily solved with conventional convex optimization tools, such as *cvxopt* [30]. According to SCA scheme, P3 is successively approximated and solved around the optimum point from the previous iteration, until the stable optimum is reached with a given $\mu^{(k)}$. When $\mu^{(k)}$ converges following (21), we can finally say that the optimal grid EE is found.

TABLE 1. Simulation parameters.

Parameter	Value	Parameter	value
T(s)	10	$p_{u,\min}(\mu\text{W})$	100
N_{noise}(dBm/Hz)	-120 [12]	W^d(KHz)	45
S	15	W^u(KHz)	45
L	3	σ	0.16
r_a(m)	40	M	16
r_0(m)	2	N	8
η	0.4	G_m(dB)	17 [21]
ε_t(s)	0.5	G_s(dB)	5 [21]
$\triangle_t$(s)	0.2	G_u(dB)	0 [21]
δ_t(s)	0.3	$e_{s,0}$(mJ)	≥ 0
α_{micro}	3.0	$e_{u,0}$(mJ)	≥ 0
α_{mmW}	2.0	r^d_{min}(Kbps)	15
$p_{0,s}$(mW)	10 [13]	r^u_{min}(Kbps)	45
$p_{0,u}(\mu\text{W})$	2.64 [12]	$p_{s,max}$(dBm)	20 [12]
$p_{m,max}$(dBm)	46 [12]	$p_{u,max}$(dBm)	10 [5]
$\triangle_p$	4.7 [22]	$p_{m,c}$(W)	130 [22]
$r^{(0)}$	1	$p_{mmW,c}$(W)	3.9 [22]
μ	0.5	ρ	10^{-3}
α	0.01	β	0.5

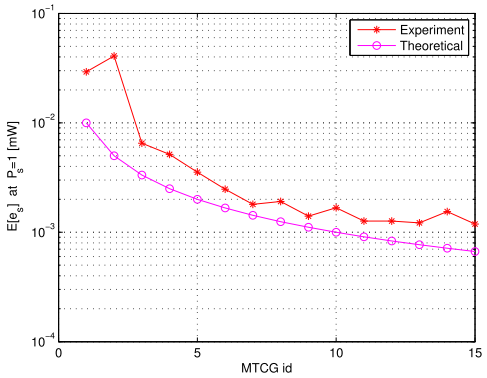


FIGURE 3. Expectation of mmWave energy harvesting rate.

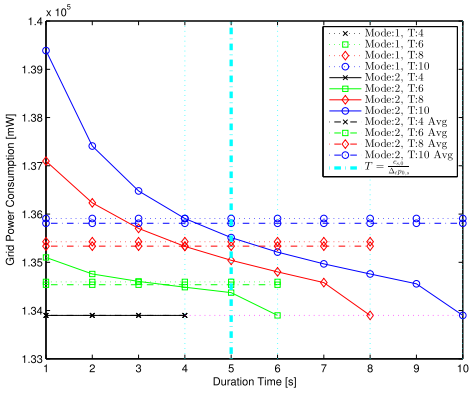


FIGURE 4. MTCG is *properly-charged* ($e_{s,0}$=10mJ).

Since 3-tier of loops, in which both the middle and inner loops are SCA process, are applied to solve the highly non-convex problem, we name it the 'nested' convex approximation method and the details are concluded in Algorithm 1. The initial EE value ($\mu^{(1)} = 1e+03$) is chosen through trial and error, although it has slight impact on convergence speed. The outer loop also converges after 3 iterations at $\mu^{(1)} = 1e+02$

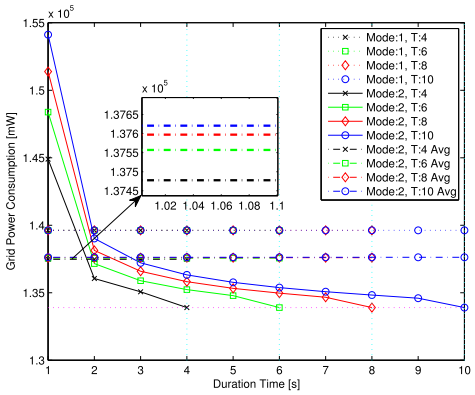


FIGURE 5. MTCG is *undercharged* ($e_{s,0}$=0.5mJ).

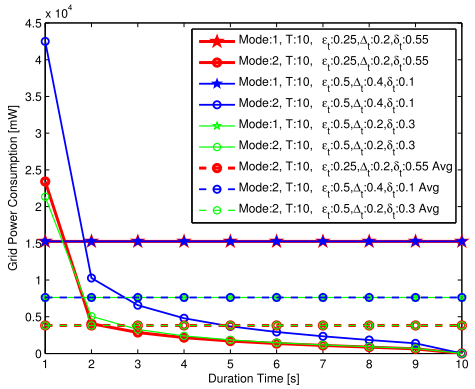


FIGURE 6. MTCG is *undercharged* ($e_{s,0}$=0mJ).

and $\mu^{(1)} = 2e+03$. $\mathbf{P_s}^{(0)}$ and $\mathbf{P_u}^{(0)}$ are initialized to random values from $(0, 1e+01]$ mW and $(0, 1]$ mW respectively in simulation experiment.

D. COMPUTATIONAL COMPLEXITY

Algorithm 1 comprises 3 loops, where the outer Dinkelbach's process converges quickly with about 3 iterations as demonstrated in figure 8, and the inner SCA process converges under sum of square error (SSE) constraint 10^{-6} with exactly 4 iterations. The complexity in solving P4 with *cvxopt* is $\mathcal{O}\left(\left(3S^4U^3+2S^3U^4\right)T^4\right)$ according to [31]. Thus the total complexity is $\mathcal{O}\left(12J\cdot\left(3S^4U^3+2S^3U^4\right)T^4\right)$.

VI. SIMULATION RESULT

Experiment parameters are configured according to previous literature and industry practice, as listed in Table 1. In this paper, we apply the rejection sampling method in [32] to produce Gin-DPP, through which MTCGs are placed on the coverage area of HAP one by one with correlated probability density function, instead of being generated at the same time. Simulations are conducted 20 times to average the result.

A. EXPECTATION OF MMWAVE ENERGY HARVESTING RATE

As can be seen in fig.3, there is a gap between our experiment result and theoretical solution of the expectation of mmWave EH rate, when transmitting power level at HAP is set to 1 mW.

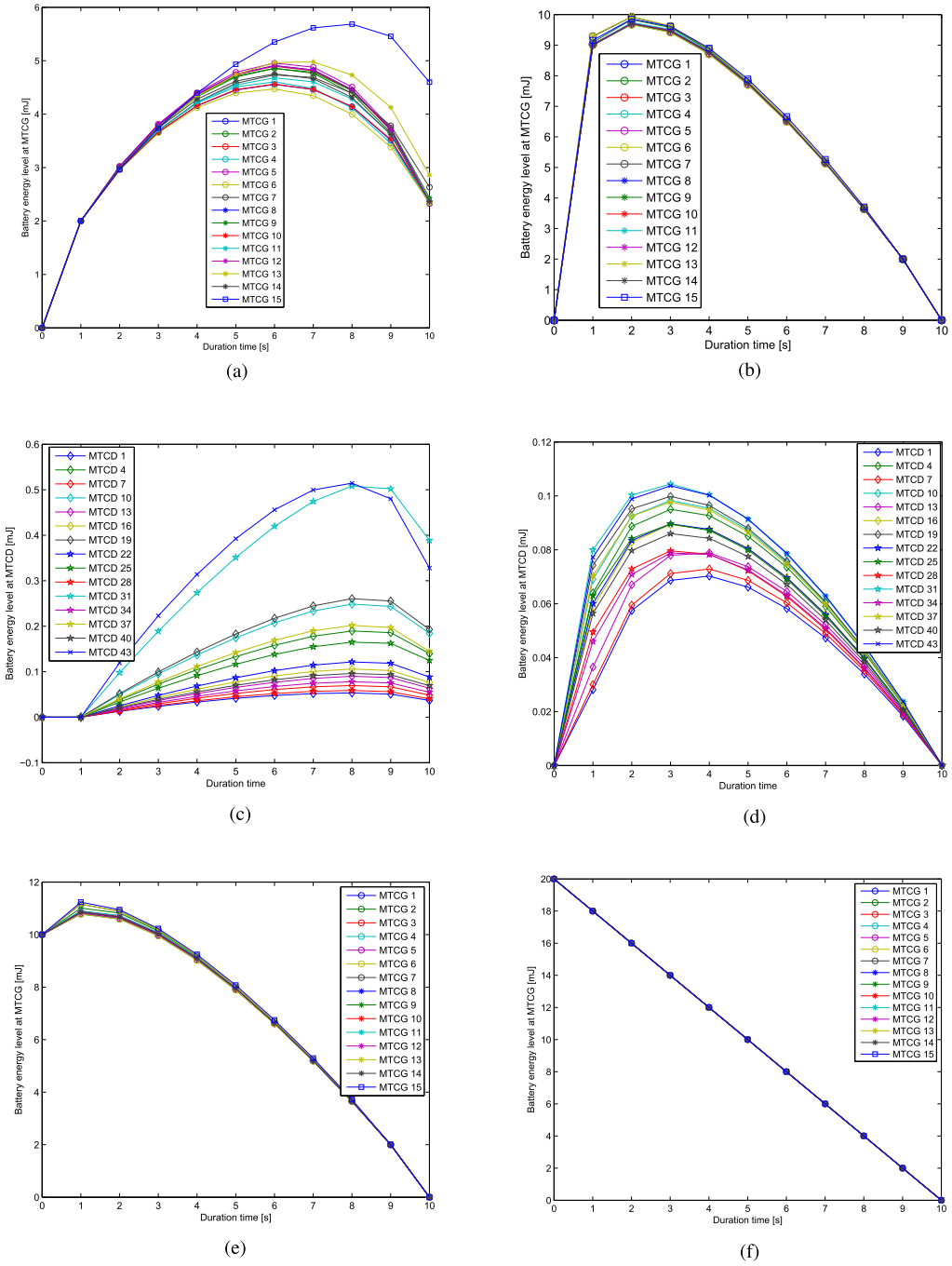


FIGURE 7. Battery dynamics at MTCG and MTCD. (a) Mode 1 ($e_{s,0}$=0mJ). (b) Mode 2 ($e_{s,0}$=0mJ). (c) Mode 1 ($e_{s,0}$=0mJ). (d) Mode 2 ($e_{s,0}$=0mJ). (e) Mode 2 ($e_{s,0}$=10mJ). (f) Mode 2 ($e_{s,0}$=20mJ).

According to [32], this gap vanishes as the number of MTCG approaches infinity. Therefore, with a denser configuration of MTCGs, a more accurate estimation of EH performance can be obtained.

B. GRID ENERGY CONSUMPTION VS. SYSTEM OPERATION TIME WHEN MTCD IS ABUNDANTLY-CHARGED

According to fig.4, we can see that before threshold $T = \frac{e_{s,0}}{\Delta_t p_{0,s}}$, variable grid power consumption under both modes approaches zero (T = 4), which is in consistency with our analysis in (13),(14). Here, the magenta dotted line represents static power consumption of HAP, which is $1.339e+05$ mW. When T passes $\frac{e_{s,0}}{\Delta_t p_{0,s}}$, the ratio between average grid power consumption under Mode 1 and Mode 2 is $1.0004 = \frac{1.3459e+05}{1.3454e+05}$, $1.00067 = \frac{1.3543e+05}{1.3534e+05}$, $1.00073 = \frac{1.3591e+05}{1.3581e+05}$ for T = 6, 8 and 10, which is the same result calculated from $\frac{\kappa_1 T+\kappa_3}{\kappa_1(T-1+\varepsilon_t)+\kappa_3}$, with $\varepsilon_t = 0.5$, $\kappa_1 = 1.9081e+03$,

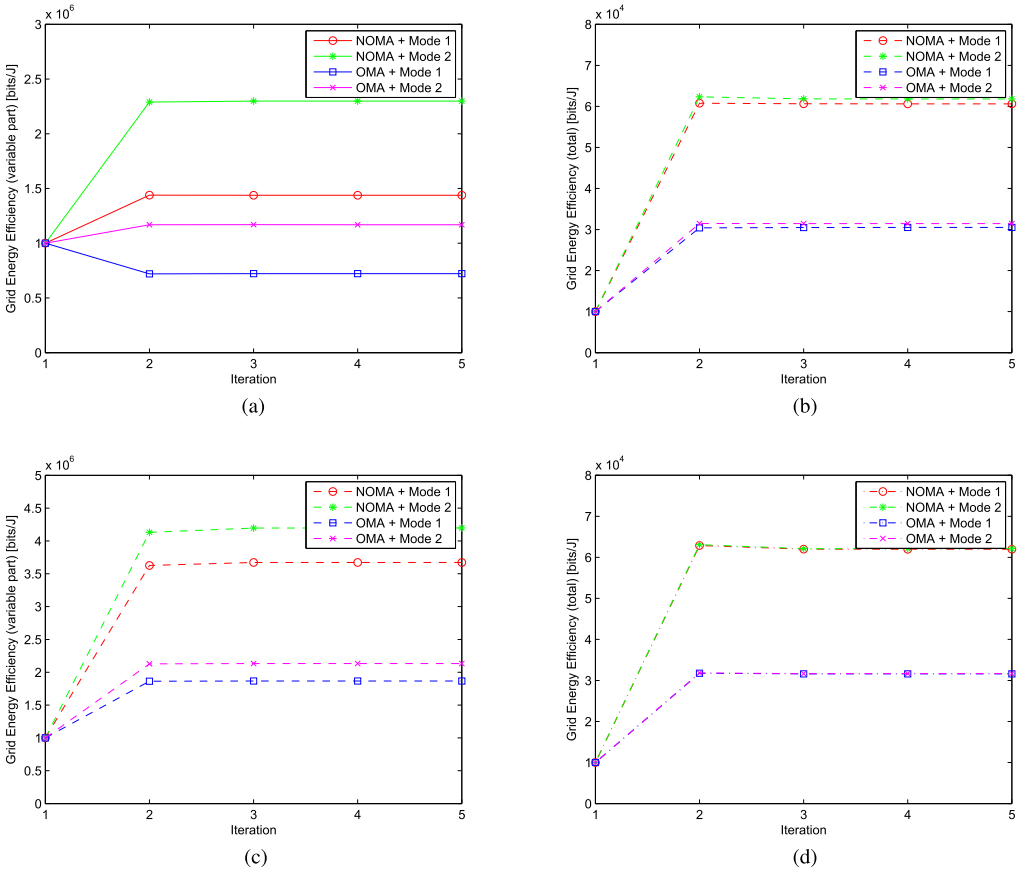


FIGURE 8. EE performance under different access mode and EH mode. (a) EE vs. variable($e_{s,0}$=0.5mJ). (b) EE vs. total($e_{s,0}$=0.5mJ). (c) EE vs. variable($e_{s,0}$=4mJ). (d) EE vs. total($e_{s,0}$=4mJ).

$\kappa_2 = 6.695e+04$, $\kappa_3(6) = 2.20935e+06$, $\kappa_3(8) = 1.339e+06$, $\kappa_3(10) = 1.27205e+06$.

When the MTCG is undercharged, condition $T > \frac{e_{s,0}}{\Delta_t p_{0,s}}$ is always satisfied, since $\frac{e_{s0}}{\Delta_t p_{0,s}} \leq (1-\varepsilon_t) < 1$ and $T > 1$. According to fig. 5, the ratio between average grid power consumption under Mode 1 and Mode 2 varies from $1.0156 = \frac{1.3962e+05}{1.3748e+05}$ to $1.0146 = \frac{1.3962e+05}{1.3762e+05}$ through $1.0150 = \frac{1.3962e+05}{1.3756e+05}$ and $1.0147 = \frac{1.3962e+05}{1.3760e+05}$ for $T = \{4, 6, 8, 10\}$, which is the same result calculated from $\frac{\kappa_1(\Delta_t p_{0,s} - e_{s,0}) + \kappa_2}{\kappa_1 \varepsilon_t \left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right) + \kappa_2}$. According to (11), the average grid energy consumption under Mode 1 remains when T changes, thus the dotted lines overlap with each other at power level $1.3962e+05$ mW.

As can be seen from fig.4 and fig.5, Mode 2 consumes much higher grid power at the initial time slot than Mode 1, while having a averagely lower consumption performance. This prevents Mode 2 from functioning for a relatively long time, since the power amplifiers (PA) at HAP may not work within an ideal region, and transmitting power constraints (4h-4i) would also be violated. Thus, comparing with Mode 1, Mode 2 is suitable for short-term hybrid energy harvesting. On the other hand, Mode 1 alleviates the heavy burden on PA components by transmitting at the same but lower power level, which makes Mode 1 an acceptable choice when long-term EH process is required.

C. VARIABLE GRID ENERGY CONSUMPTION VS. TIME SLOT SPLITTING STRATEGY

According to our analysis in Sec.III. C, closed-form relationship between average variable grid power consumption of Mode 1 and Mode 2 can always be found as $\frac{\Delta_t p_{0,s} - e_{s,0}}{\varepsilon_t \left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)}$, when MTCD is *abundantly-charged* and MTCG is *under-charged*. Obviously, this ratio is proportional to reciprocal of ε_t, which is verified when observing the red and green solid line in fig.6 when MTCG is undercharged with no energy, with $\varepsilon_t = 0.25$ and $\varepsilon_t = 0.5$. The red solid line has twice the power level as green solid line. Also, when $e_{s,0} = 0$ mJ, the average variable grid power consumption under Mode 1 is proportional to $\Delta_t p_{0,s}$, which is why the red solid line and blue solid line overlaps. The fixed ratio between average variable grid power consumption under Mode 1 and Mode 2 is observed as being 2 when $\varepsilon_t = 0.5$ in green and blue lines, and 4 when $\varepsilon_t = 0.25$ in red lines.

D. BATTERY DYNAMICS AT MTCG AND MTCD

When both MTCG and MTCD are charged with no energy, from Fig 7.(a)(c), we can see that battery energy would

Algorithm 1 The Nested Convex Approximation Based Energy Efficiency Maximization

Set $\mu^{(1)}$=1e+03, ε_p= 1e-02, ε_μ=1e-02, K=5, J=5, M=4, $\chi \in \{\mathbf{s}, \mathbf{u}\}$. Set $\mathbf{P}_\chi{}^{(0)}$ to any feasible value.
for $k = 1 : K$ **do**
  for $j = 1 : J$ **do**
    Calculate $\hat{r}_{s,u}^{j,(i)}$, $\forall 1 \le i \le T$ with $\mathbf{P_s}^{(j-1)}$ according to (23),(24). Calculate $\hat{r}_{u,s}^{j,(i)}$ with $\mathbf{V_s}^{(j-1)}$.
    Solve P4 at $\mu^{(k)}$, $\hat{r}_{s,u}^{j,(i)}$, $\hat{r}_{u,s}^{j,(i)}$.
    Project solution of step 5 onto $\mathbf{V}_\chi$ domain.
    for $m = 1 : M$ **do**
      Solve P4 with additional constraints,

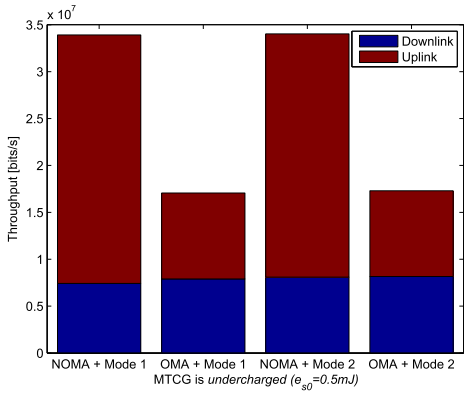
$$\left\| 2^{\mathbf{V}_\chi^{*-}} + 2^{\mathbf{V}_\chi^{*-}} (\mathbf{V}_\chi - \mathbf{V}_\chi^{*-}) - \mathbf{P}_\chi \right\|^2 \le 10^{2-2m}$$

      Update $\mathbf{V}_\chi^{*-}$ with solution of step 8.
    end for
    Update $\mathbf{P}_\chi{}^{(j)}$ with solution of step 8.
    if $\left\|\mathbf{P}_\chi{}^{(j)} - \mathbf{P}_\chi{}^{(j-1)}\right\| \le \varepsilon_p$ **then**
      Break.
    end if
  end for
  Update tuning parameter $\mu^{(k+1)}$ with converged $\mathbf{P}_\chi{}^{(j)}$.
  if $\mu^{(k+1)} - \mu^{(k)} \le \varepsilon_\mu$ **then**
    Break.
  end if
end for
Output $\mu^{(k+1)}$ as the optimum energy efficiency.

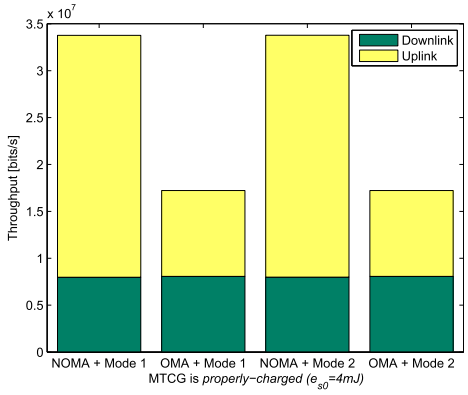
accumulate at MTCG and MTCD under Mode 1, due to a fixed and excessive transmitting power level at HAP. However, Mode 2 always finds the minimum power required and never wastes energy in charging batteries at MTCG and MTCD, as illustrated in fig 7.(b)(d). When MTCG is *properly-charged* in fig.7(e), Mode 2 completely uses initial energy at MTCG and drains battery at the end of operation. It is also observed that both MTCG and MTCD would accumulate enough energy at their battery before consuming it, resulting in a concave curve of the corresponding battery dynamics. When MTCG is initially charged with sufficient energy, it will consume the same amount of energy per time slot, until the battery is drained, as shown in fig.7(f).

E. GRID ENERGY EFFICIENCY VS. MTCG INITIAL BATTERY ENERGY LEVEL

In this section, we give the EE performance of proposed cellular M2M system applying both OMA and NOMA uplink transmission. System duration time is set to $T = 4$s to keep the problem at a tractable scale. As can be seen in fig.8(a)(c), $\mu^{(k)}$ converges after averagely 3 iterations, which is quite efficient performance. NOMA transmission achieves much higher EE value than OMA (nearly 2x), since the uplink interference are alleviated through SIC decoding.



(a)



(b)

FIGURE 9. Throughput comparison under different access mode and EH mode. (a) System throughput when MTCG is *undercharged* ($e_{s,0}$=0.5mJ). (b) System throughput when MTCG is *properly-charged* ($e_{s,0}$=4mJ).

From fig.(b)(d), due to the large amount static power consumption, EE improvement from Mode 2 against Mode 1 is around 1.97%(OMA) to 3.07%(NOMA) when MTCG is *undercharged* and 0.21%(OMA) to 0.24%(NOMA) when MTCG is *properly-charged*, which is not comparable with that from access mode shift. Mode 1 yields a lower EE performance than Mode 2 since system consumes more grid energy under Mode 1 as analysed in section IV.C. Although the exact ratio between optimum EE_{var} under two EH modes is not $\frac{\Delta_t p_{0,s} - e_{s,0}}{\varepsilon_t\left(\Delta_t p_{0,s} - \frac{e_{s,0}}{T}\right)} = 1.6$ when MTCG is *undercharged* with $e_{s,0}$=0.5mJ, it takes a rather close value within [1.5981, 1.6195], as can be observed from fig.8(a). Similar phenomenon when MTCG is *properly-charged* with $e_{s,0}$=4mJ can be observed in fig.8(b), where the ratio approaches $\frac{T}{T-1+\varepsilon_t} = 1.1429$ with [1.1428, 1.1432]. Comparing fig.9(a) and fig.9(b), we find that there's slight difference between optimum system throughput when initial battery energy level at MTCG changes. What's more, NOMA transmission helps producing nearly 3x (average:2.9329) uplink throughput than OMA scheme, which helps double the total throughput. When system works in OMA mode, both the downlink and uplink throughput are optimized to

their most energy efficient level, at which the ratio between them approaches $1-\sigma$ (0.84 in our case). This stems from the channel reciprocity property and the difference in SINR calculation (2)-(3).

VII. CONCLUSION

With the paradigm shift from Internet of Things (IoT) to Internet of Things Services (IoTS) in 5G, more attention are paid to how to innovate IoT applications and how to improve the intelligence, autonomy and visualization of IoT based on existing access network facilities. In this paper, we discuss the grid energy consumption and efficiency of cellular-based M2M system, for which a dedicated hybrid EH protocol is proposed. Millimeter wave EH is applied here to play the role of energy supplier from 'external' environment, regarding the cellular M2M system as a black box in power domain. This is a novel work on both smart grid application and edge computing, since we achieve decoupling of power grid energy production and consumption as well as modularization design of cellular M2M network and propose autonomous power management mechanism deployed at ingress/egress routers of cloud-based cellular network. Two mmWave EH modes are analysed in our model. According to our theory, EH Mode 1 is better choice when the system is supposed to work for a relatively long time, while Mode 2 is preferred when energy shortage is temporary due to its averagely lower energy consumption. However, static grid power consumption at HAP is still the overwhelming factor determining energy performance of this system. Closed-form expression of average grid energy consumption under both modes are derived when MTCD is *abundantly-charged*, while numerical solution is obtained using interior-point method when MTCD is *under-charged*. The proposed time division protocol can flexibly adapt to service QoS requirements and battery conditions, which facilitates support of a variety of IoT applications. Uplink NOMA transmission remarkably improves throughput while keeping the total energy consumption at a stable level, which helps producing a much better EE performance.

In our future work, we plan to investigate the self-organizing behavior between MTCG and MTCD using machine learning tools, trying to capture the effect cross-tier cooperation casts on energy performance of cellular-based M2M system.

APPENDIX A

Looking carefully into the definition of $f^d\left(\mathbf{P}_{\mathbf{s},u}^{(i)}\right)$, we can formulate problem for $\forall u \in \mathcal{U}$, $1 \le i \le T$ as :

$$\min\max \ \mathbf{P}_{\mathbf{s},u}^{(i)} \tag{29a}$$

$$s.t.\ G_s G_u \Phi(u) \mathbf{P}_{\mathbf{s},u}^{(i)} \ge \vartheta {\sigma_0}^2 \mathbf{I} \tag{29b}$$

where, $\Phi(u) =$

$$\begin{bmatrix} (1-\sigma)\, l\left(d_{s,u}\right) & \vartheta\, l\left(d_{c,u}\right) & \dots & \vartheta\, l\left(d_{c,u}\right) \\ \vartheta\, l\left(d_{c,u}\right) & (1-\sigma)\, l\left(d_{s,u}\right) & \dots & \vartheta\, l\left(d_{c,u}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \vartheta\, l\left(d_{c,u}\right) & \vartheta\, l\left(d_{c,u}\right) & \dots & (1-\sigma)\, l\left(d_{s,u}\right) \end{bmatrix},$$

$\vartheta = \left(2^{\frac{Lr_{\min}^d}{W^d}} - 1\right)$. Thus we have $\mathbf{P}_{\mathbf{s},u}^{(i)} \ge (G_s G_u)^{-1} \Phi(u)^{-1} \vartheta {\sigma_0}^2 \mathbf{I}$, $\mathbf{I}$ is the all-one column vector and $(\cdot)^{-1}$ is the matrix inverse operator. Since Φ can be estimated and pre-calculated, we can obtain the lower bound of the maximum transmitting power of MTCG in satisfying QoS requirement $r_{\min}^d$ as

$$\min\max \quad \mathbf{P}_{\mathbf{s},u}^{(i)} \Leftrightarrow \max \quad (G_s G_u)^{-1} \Phi(u)^{-1} \vartheta {\sigma_0}^2 \mathbf{I} \tag{30}$$

Similarly, we can derive the lower bound of the maximum transmitting power of MTCD in satisfying QoS requirement $r_{\min}^u$ as,

$$\min\max \quad \mathbf{P}_{\mathbf{u},s}^{(i)} \Leftrightarrow \max \quad (G_s G_u)^{-1} \Upsilon(\mathcal{I}(u), s)^{-1} \psi {\sigma_0}^2 \mathbf{I} \tag{31}$$

where,

$$\Upsilon(\mathcal{I}(u), s) = \begin{bmatrix} l\left(d_{u,s}\right) & \psi l\left(d_{c,s}\right) & \dots & \psi l\left(d_{c,s}\right) \\ \psi l\left(d_{c,s}\right) & l\left(d_{u,s}\right) & \dots & \psi l\left(d_{c,s}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \psi l\left(d_{c,s}\right) & \psi l\left(d_{c,s}\right) & \dots & l\left(d_{u,s}\right) \end{bmatrix},$$

$\psi = \left(2^{\frac{Lr_{\min}^u}{W^u}} - 1\right)$. Then we can say $\forall s \in \mathcal{S}$, $u \in \mathcal{U}$,

$$r_{\min}^{u*} = \arg\max_{r_{\min}^u} \left[(G_s G_u)^{-1} \Upsilon(\mathcal{I}(u), s)^{-1} \psi {\sigma_0}^2 \mathbf{I} \le p_{u,\min}\right] \tag{32}$$

and $\forall u \in \mathcal{U}$

$$r_{\min}^{d*} = \arg\max_{r_{\min}^d} \left[(G_s G_u)^{-1} \Phi(u)^{-1} \vartheta {\sigma_0}^2 \mathbf{I} \le \upsilon\right] \tag{33}$$

where υ is a small positive value, e.g.10^{-5}. When $r_{\min}^u \le r_{\min}^{u*}$ and $r_{\min}^d \le r_{\min}^{d*}$, we can say that the throughput QoS requirements are moderate.

REFERENCES

[1] M. R. Palattella *et al.*, "Internet of Things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Mar. 2016.

[2] T. Lv, Y. Ma, J. Zeng, and P. T. Mathiopoulos, "Millimeter-wave noma transmission in cellular M2M communications for Internet of Things," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1989–2000, Jun. 2018.

[3] G. Zhang, A. Li, K. Yang, L. Zhao, Y. Du, and D. Cheng, "Energy-efficient power and time-slot allocation for cellular-enabled machine type communications," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 368–371, Feb. 2016.

[4] N. Xia, H.-H. Chen, and C.-S. Yang, "Radio resource management in machine-to-machine communications—A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 791–828, 1st Quart., 2018.

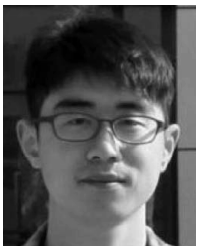
[5] Z. Yang, W. Xu, Y. Pan, R. Guan, and M. Chen, "Energy minimization in machine-to-machine systems with energy harvesting," in *Proc. IEEE 85th Veh. Technol. Conf.*, Jun. 2017, pp. 1–5.

[6] T. D. Nguyen and J. Y. D. T. Khan Ngo, "A self-sustainable RF energy harvesting algorithm for WSN-based IoT applications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[7] Y. Mao, Y. Luo, J. Zhang, and K. B. Letaief, "Energy harvesting small cell networks: Feasibility, deployment, and operation," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 94–101, Jun. 2015.

[8] 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.

[9] M. A. Andersson, A. Özçelikkale, M. Johansson, U. Engström, A. Vorobiev, and J. Stake, "Feasibility of ambient RF energy harvesting for self-sustainable M2M communications using transparent and flexible graphene antennas," *IEEE Access*, vol. 4, pp. 5850–5857, 2016.
[10] K. Huang and V. K. N. Lau, "Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 902–912, Feb. 2014.
[11] T. A. Khan, A. Alkhateeb, and R. W. Heath, "Millimeter wave energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6048–6062, Sep. 2016.
[12] X. Lu, I. Flint, D. Niyato, N. Privault, and P. Wang, "Self-sustainable communications with RF energy harvesting: Ginibre point process modeling and analysis," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1518–1535, May 2016.
[13] L. Pei *et al.*, "Energy-efficient D2D communications underlaying NOMA-based networks with energy harvesting," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 914–917, May 2018.
[14] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.
[15] Z. Yang, W. Xu, Y. Pan, C. Pan, and M. Chen, "Energy efficient resource allocation in machine-to-machine communications with multiple access and energy harvesting for IoT," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 229–245, Feb. 2018.
[16] L. Wang, S. Chen, and M. Pedram, "Power management of cache-enabled cooperative base stations towards zero grid energy," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–6.
[17] J. Du, E. Onaran, D. Chizhik, S. Venkatesan, and R. A. Valenzuela, "Gbps user rates using mmWave relayed backhaul with high-gain antennas," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1363–1372, Jun. 2017.
[18] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
[19] L. Wang, M. Elkashlan, R. W. Heath, Jr., M. Di Renzo, and K.-K. Wong, "Millimeter wave power transfer and information transmission," in *Proc. IEEE Global Commun. Conf.*, Dec. 2015, pp. 1–6.
[20] Y. Mao, J. Zhang, and K. B. Letaief, "Grid energy consumption and QoS tradeoff in hybrid energy supply wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3573–3586, May 2016.
[21] Z. Tan, H. Qu, J. Zhao, G. Ren, and W. Wang, "Low-complexity networking based on joint energy efficiency in ultradense mmWave backhaul networks," *Trans. Emerg. Telecommun. Technol.*, vol. 30, no. 1, p. e3508, 2018.
[22] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
[23] D. W. K. Ng and R. Schober, "Energy-efficient power allocation for M2M communications with energy harvesting transmitter," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2012, pp. 1644–1649.
[24] M. Zgaren and M. Sawan, "A high-sensitivity battery-less wake-up receiver for 915 MHz ISM band applications," in *Proc. IEEE Int. Conf. Electron., Circuits, Syst. (ICECS)*, Cairo, Egypt, Dec. 2015, pp. 336–339.
[25] X. Wang and A. Mortazawi, "High sensitivity RF energy harvesting from AM broadcasting stations for civilian infrastructure degradation monitoring," in *Proc. IEEE Int. Wireless Symp. (IWS)*, Beijing, China, Apr. 2013, pp. 1–3.
[26] E. Kostlan, "On the spectra of Gaussian matrices," *Linear Algebra Appl.*, vols. 162-164, no. 2, pp. 385–388, 1992.
[27] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
[28] H. Ju, B. Liang, J. Li, and X. Yang, "Dynamic power allocation for throughput utility maximization in interference-limited networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, pp. 22–25, Feb. 2013.
[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[30] L. Vandenberghe. (2010). *The CVXOPT Linear and Quadratic Cone Program Solvers*. [Online]. Available: http://www.seas.ucla.edu/~vandenbe/publications/coneprog.pdf
[31] A. Brighente and S. Tomasin, "Power allocation for non-orthogonal millimeter wave systems with mixed traffic," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 432–443, Jan. 2019.
[32] L. Decreusefond, I. Flint, and A. Vergne, "Efficient simulation of the Ginibre point process," *J. Appl. Probab., Appl. Probab. Trust*, vol. 52, no. 4, pp. 1003–1012, 2015.



ZHENJIE TAN received the B.E. degree in information engineering from Xi'an Jiaotong University, Xi'an, China, in 2013, where he is currently pursuing the Ph.D. degree. He has been with the Institute of Advanced Information and Communication Technologies, Xi'an Jiaotong University, since 2014, where he works in the projects in cooperation with ZTE Communication. His research interests include heterogeneous access networks, radio resource allocation, and mobile user profiling.

HUA QU received the B.E. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, and the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, where he is currently a Professor. His current research interests include the mobile Internet, network protocol design and control strategies for supporting emerging applications in ubiquitous networks, and radio resource management in 5G radio communication systems.

JIHONG ZHAO received the B.E. degree from the Huazhong University of Science and Technology, Wuhan, China, and the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China. She is currently a Professor with Xi'an Jiaotong University and the Xi'an University of Posts and Telecommunications. Her current research interests include broadband communications, network function virtualization, and software-defined networking.

GONGYE REN is currently pursuing the Ph.D. degree with the School of Electronic and Information Engineering, Xi'an Jiaotong University, China. His research interest includes the effects of user behaviors on the wireless networks, which includes handoff decision, green wireless networks, and content-centric networks.

WENJIE WANG received the B.S., M.S., and Ph.D. degrees in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, 1998, and 2001, respectively, where he is currently a Professor. From 2009 to 2010, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA. His main research interests include information theory, broadband wireless communications, signal processing with application to communication systems, array signal processing, and cooperative communications in distributed networks.