Scheduling of Electric Vehicle Charging via Multi-Server Fair Queueing

Xudong Wang, *Senior Member, IEEE*, Yibo Pi, and Aimin Tang

Abstract—Charging electric vehicles (EVs) at home is attractive to EV users. However, when the penetration level of EVs becomes high, a distribution grid suffers from problems such as under-voltage and transformer overloading. EV users also experience a fairness problem, i.e., the limited capacity is unfairly shared among EVs. To solve these problems, a physical fair-queueing framework is established for EV charging. In this framework, a distribution sub-grid is first mapped to a multi-server queueing system, and then a fluid-model based queueing scheme called physical multi-server generalized processor sharing (pMGPS) is designed. pMGPS ensures perfect fairness but cannot be used practically due to its nature of fluid model. To this end, a packetized scheme called physical start-time fair queueing (pSTFQ) is developed to schedule tasks of EV charging. The fairness performance of the pSTFQ scheduling scheme is characterized by the ratio of energy difference between pSTFQ and pMGPS. This critical performance metric is studied through theoretical analysis and is also evaluated via simulations. Performance results show that the pSTFQ scheduling scheme achieves an energy difference ratio of less than 4 percent in various scenarios without causing under-voltage and transformer overloading problems.

Index Terms—Electric vehicle charging, distribution grid, multi-server fair queueing

1 Introduction

Electric vehicles (EVs) have been considered as an alternative to conventional vehicles. As predicted in [1], by 2050 the penetration level of EVs will reach 62 percent in the US. Compared to the supercharging and battery swap services at public charging stations, charging EVs at home is also very attractive. EV users can have their EVs charged immediately after they arrive home. However, unlike public charging stations that are deliberately designed for EV charging, charging stations at homes are distributed and pose threats to the existing distribution grid.

Generally, EVs consume much larger power than home appliances like air conditioners and heaters. For an individual user, charging an EV at home can be easily achieved by installing an outlet with high power rating. However, from the perspective of an entire distribution grid, if a large number of EVs are to be charged simultaneously, the peak load is expected to increase abruptly. Since the capacity of existing transformers is planned without considering the additional peak load from such EVs [2], these transformers can easily experience an *overloading problem*. EV charging also increases the current flowing through transmission feeders, which results in severe voltage drop along these feeders, especially in rural areas with long transmission feeders [3]. Voltage below the operation range leads to an *under-voltage problem*, which can damage home appliances. In this paper, we call both problems the *physical problems* of power grid.

Reinforcing the existing distribution grid by upgrading the transformers and feeders can solve the above problems, but it is costly. Thus, a preferable and practical approach is to shift EV charging in peak-load time to non-peak-load time. Following this approach, so far both offline and online schemes have been proposed. Offline schemes [4], [5], [6] usually formulate the load-shifting task as an optimization problem, subject to the constraints of EV energy requirements and grid safety. The energy requirement of an EV is said to be satisfied if its required energy can be charged before deadline. Based on a solution to the optimization problem, the on/off time of each EV can be determined. Offline methods share several common issues. First, the load profiles of homes need to be forecast, but it may be much different from the real value. Charging decisions based on erroneous forecast profiles can still cause transformer overloading or under-voltage problems. Second, offline schemes assume the grid capacity is sufficient to charge all EVs by deadlines. Third, each EV has to follow a pre-specified deadline strictly, which is not practical for EV users. For example, an EV arriving at 6 pm sets a deadline of 12 pm, its charging period may be scheduled between 10 to 12 pm. In case it departs at 10 pm, it awaits 4 hours for nothing.

So far online schemes [7], [8] have been developed to avoid forecast errors by using updated status data of power grid. However, in the case of arbitrary early departure of EVs or occurrence of the physical problems, these schemes cannot ensure EVs to be charged in a fair way. Thus, a *fairness problem* arises. The schemes in [9], [10] consider the physical problems, but the fairness problem is still not addressed.

- *X. Wang and A. Tang are with the University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai 200240, China. E-mail: wxudong@ieee.org, tangaiming@sjtu.edu.cn.*
- *Y. Pi was with the University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai 200240, China and is now with the University of Michigan, Ann Arbor, MI 48109. E-mail: piyibo@gmail.com.*

Manuscript received 9 Apr. 2016; revised 15 May 2017; accepted 18 May 2017. Date of publication 1 June 2017; date of current version 11 Oct. 2017. (Corresponding author: Xudong Wang.)
Recommended for acceptance by M. M. Hayat.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPDS.2017.2710197

Moreover, all the above online schemes are developed under the assumption that the grid capacity is sufficient for EV charging, which may not be true when the penetration level of EVs is high. The scheme of deadline scheduling with commitment [11] circumvents this issue by declining additional charging requests once the grid capacity has been reached. This approach is reasonable for public charging stations. However, for EV charging at home, declining a charging request from any EV user simply results in the fairness problem.

Proportional fairness among EVs is considered in [12], [13], so the fairness problem due to early departures of EVs does not exist. However, these schemes require the distribution grid capacity to be known, which is difficult to obtain, especially when the capacity is constrained by the voltage drop as mentioned in [13], [14]. Another drawback of the proportional sharing schemes is that an EV may be allocated with a charging rate below the power rating, but the commercial EVs must be charged at specific rate levels according to the state of charge (SOC) in batteries [15].

In this paper, a new scheduling scheme is developed for EV charging and eliminates all the above issues in both offline and online schemes. More specifically, under the physical constraints of the distribution grid, i.e., voltage range and transformer capacity, fair EV charging is achieved through scheduling on/off states of EV batteries. Moreover, the scheme does not rely on any knowledge about the topology[1] of the distribution grid.

To fulfill the above task, we leverage fair queueing theories [16], [17], [18] in information systems to develop an innovative multi-server fair queueing scheme for physical systems like EV charging in a distribution grid. In this scheme a sub-grid in the distribution grid is first mapped onto a queueing system that virtually represents the physical system (i.e., the sub-grid). Considering this physical queueing system, a service discipline called physical multi-server generalized processor sharing (pMGPS) is then proposed to ensure perfect fairness. pMGPS follows a fluid-flow model as it assumes continuous charging rates in EVs, so it cannot schedule EV charging practically. To support on/off EV charging at a specific level of charging rate, a packetized scheme called physical start-time fair queueing (pSTFQ) is designed by taking pMGPS as a reference. The performance gap between pSTFQ and pMGPS is captured by the ratio of energy difference between these two queueing systems. Theoretical analysis is conducted in this paper to derive the upper bound of energy difference ratio. Simulations further confirm that the energy different ratio is within 4 percent under various scenarios. Moreover, simulation results demonstrate that our scheduling scheme achieves fair EV charging without causing any physical problems like transformer overloading or under-voltage.

The rest of the paper is organized as follows. In Section 2.1, the system architecture of EV charging in a distribution grid is presented. In Section 3, the physical multi-server queueing system is introduced for EV charging and pMGPS is developed for such a system. The packetized scheme pSTFQ is designed in Section 4. The fairness performance is analyzed in Section 5, and simulation results are presented in Section 6. The paper is concluded in Section 7.

1. In this paper, the topology of a power grid includes both connections and parameters such as resistance, capacitance, and inductance.

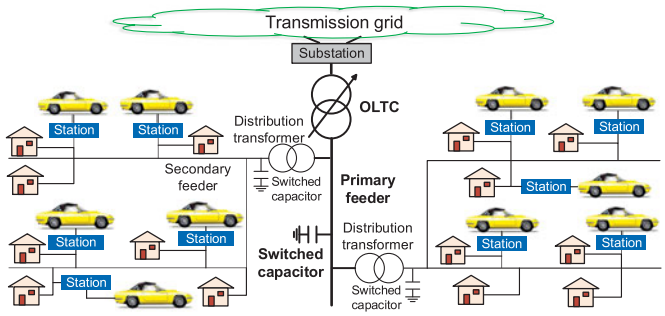


Fig. 1. A distribution grid with EVs.

2 Overall System Design

2.1 System Architecture

The typical architecture of a distribution grid is illustrated in Fig. 1, which consists of medium voltage (MV) and low voltage (LV) networks. In MV networks, voltage on primary feeders is regulated by the on-load tap changer at the substation and the switched capacitors and voltage regulators connected to primary feeders. In LV networks, switched capacitors on the secondary side of distribution transformers regulate the voltage on secondary feeders. When a number of EVs are attached to the distribution grid from different homes, their charging schedules need to be properly coordinated among these EVs; otherwise, the distribution transformers may be overloaded or the secondary feeders may suffer from an under-voltage problem.

As shown in Fig. 1, a distribution grid consists of multiple *sub-grids*, each of which includes a distribution transformer, several secondary feeders, and various loads. Thanks to the hierarchical architecture of a distribution grid, the sub-grids work independently. Thus, this paper is focused on a sub-grid.

A detailed look into a sub-grid is shown in Fig. 2. Each user is assumed to have a smart meter that has two capabilities: 1) measurement of voltage and power; 2) communications. Each EV is connected to the sub-grid via a smart charger. Besides its capability of turning on or off battery charging, a smart charger also needs to measure power flow of EV charging. A smart charger communicates with its associated smart meter via a wired link. It should be noted that on/off operations of battery charging do not involve the charge/discharge cycles that lead to memory effect [19], so the battery lifetime will not be impacted by our scheduling scheme. To coordinate battery charging among different EVs, a central controller located beside the distribution transformer schedules EV charging for the entire sub-grid, and a communication network is needed to connect all smart meters and the central controller. Via this communication network, the central controller collects status information from smart meters (as well as smart chargers) and dispatches the scheduling results to smart meters. The communication network can be wired or wireless; proper design of a specific network is out of the scope of this paper. This paper is focused on the design of a scheduling scheme for EV charging based on fair queueing. This scheme runs in the central controller.

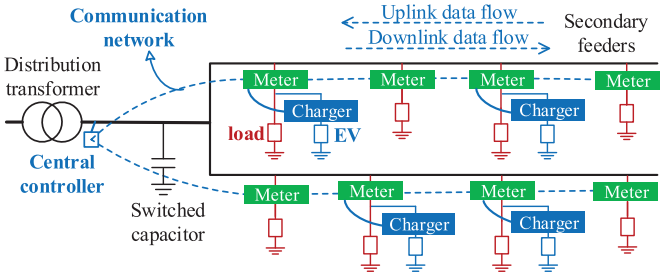


Fig. 2. The major components in a sub-grid.

2.2 Coordination of Battery Charging Among EVs

The on/off control of EV charging is conducted periodically. Before each period starts, the central controller relies on a scheduling scheme to determine the on or off state for each EV. The scheduling scheme needs to ensure fairness among EVs and avoid physical problems such as transformer overloading and under-voltage, subject to the constraint of the limited capacity of the sub-grid. As shown in Fig. 3, periodic control of EV charging based on the scheduling scheme can constantly push the voltage back to the normal range, while under-voltage can easily occur without such control.

No matter how small the charging period is, load changes can happen during this period. Unless the period is sufficiently small, the accumulated load change may cause under-voltage or transformer overloading. However, if the charging period is too small, an EV battery is switched on/off too frequently, which leads to harmonics in grid. Thus, given a periodic control scheme, some measures must be taken to address the above issues, as explained below.

With periodic control, the on/off states of EVs are controlled at the beginning of each charging period denoted as T. Unless under-voltage happens in the middle of a charging period, no other actions on EVs are made during a charging period. Once under-voltage occurs, immediate actions should be taken, i.e., switches of EVs are selected randomly and turned off until the voltage is back to normal. This scenario is illustrated in the under-voltage case (right before $5T$) in the solid line. Thanks to periodic control, under-voltage is a rare event in scheduled EV charging. Moreover, by shortening the charging period, we can reduce the probability of under-voltage to a satisfactory level.

Let V_i be the voltage of user i and $\mathcal{U}$ be the set of all users. The voltage of users should satisfy the *voltage constraint*

$$V_{\min} + \Delta V \leq V_i \leq V_{\max} - \Delta V, \quad \forall i \in \mathcal{U}, \tag{1}$$

where $V_{\min}$ and $V_{\max}$ are the lower and upper bounds of the voltage range, and ΔV is the safety margin. The safety margin is critical for the sub-grid to be resilient to load changes.

The situation of transformer overloading is much simpler. Let S_0 be the power flow through the distribution transformer, the *power flow constraint* needs to be satisfied

$$|S_0| \leq S_0^{\max}, \tag{2}$$

where $S_0^{\max}$ is the maximum allowable power flow through the transformer. Whenever this constraint is violated, the overloading issue can be handled by a new scheduling result for the next charging period, and no safety margin is reserved for the power flow S_0. The reason is two-fold: 1) A distribution transformer can tolerate a load that is three times of its nameplate rating for less than half an hour in sacrifice of lifetime [2]; 2) The accumulated load changes during a charging period are much smaller than the nameplate ratings of a transformer and thus have merely no effects on the lifetime of the transformer.

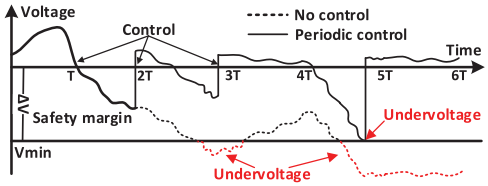


Fig. 3. Voltage variations: Periodic control versus no control on EV charging.

3 Physical Queueing System and Its Generalized Processor Sharing (GPS)

Under the periodic control, an EV may be turned on in one charging period and off in the next charging period, which implies the required energy of EVs can be charged in non-consecutive charging periods. If we consider the energy that an EV is charged during a charging period as a packet, the EV as a queue, and the charger of the EV as a server, then the charging process of a single EV is just a single-server queueing system where the service rate of the server equals the charging power of the EV. However, considering all EVs being charged in the sub-grid, the charging process is actually a multi-server queueing system where the service rates of servers are coupled through the physical constraints of the sub-grid. More specifically, the total service rate of servers (i.e., the total charging power) cannot lead to an overloading or under-voltage problem.

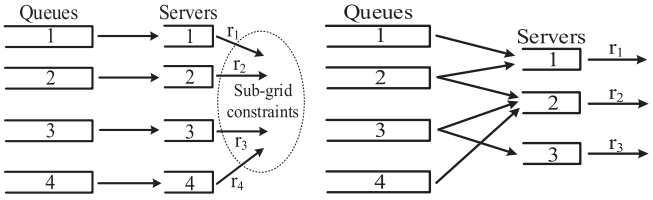
In fair queueing of data networks, servers need to determine the queues to be serviced next such that each queue gets a fair share of resources. Scheduling of EV charging in a sub-grid follows a similar process, i.e., the central controller determines proper EVs to be charged next such that EVs can get a fair share of power resources. Thus, we can leverage the design principle of fair queueing in information systems for a physical system like EV charging in a sub-grid. However, there exist significant differences between these two systems. In order to utilize fair queueing theories, the first step is to map a sub-grid into a physical queueing system.

3.1 Physical Multi-Server Queueing System

Since multiple EVs can be charged on the same sub-grid simultaneously, the physical queueing system is actually a multi-server queueing system. We call such a queueing system a physical multi-server queueing system. A model is shown in Fig. 4a, and the detailed mapping is explained below.

3.1.1 Energy Packet of EVs

The energy charged in a period for an EV is viewed as an energy packet. When a packet of an EV is serviced by a server, it means physically the EV is being charged by a charging station. Since the power rating of an EV depends on the SOC of battery that varies with charging, we simply denote the power rating of EV i at time t as $\hat{P}_i(t)$. For any EV



(a) Physical multi-server queueing (b) Information multi-server queueing with queue preference

Fig. 4. Physical versus information multi-server queueing system.

i, since a charging period is short, the energy EV i is charged during a charging period is much smaller compared to its battery capacity. Thus, the SOC of batteries within a charging period is nearly the same. The length of a packet of EV i during any charging period $[t_1, t_1 + T]$ is equal to $\int_{t_1}^{t_1+T} \hat{P}_i(t)dt = \hat{P}_i(t_1)T$, where T is the length of a charging period. The $\hat{P}_i(t)$ at the current time can be measured, but its future rating cannot be predicted due to unknown charging profile and the battery SOC. Thus, the length of a packet is unknown until the packet starts to be serviced.

3.1.2 Weights of Queues

When EV i arrives, all packets are backlogged in queue i. The weight of queue i (denoted by w_i) is equal to the weight of EV i, which is selected by its user before charging starts. The weight represents a user's need of charging power; a user who wants faster charging selects a higher weight, as will be shown in Section 3.3. However, to ensure proper selection of weights by different users, a user needs to pay a higher fee for faster charging. It is up to the service provider to determine the relationship between weight and fee and also the specific weights that are available for users to choose.

3.2 Physical versus Information Multi-Server Queueing Systems

In a physical multi-server queueing system, each queue is assigned to a specific server, implying that each queue has set its preference on a server. The service rates of servers are coupled via the sub-grid. In the literature of fair queueing, many queueing systems have been proposed, but most of them have no queue preferences. The first queueing system with queue preferences is proposed in [20] for information systems, which is shown in Fig. 4b. It is compared with the physical multi-server queueing system as follows.

3.2.1 Energy Packets versus Information Packets

Unlike information packets, energy packets of an EV differ in three aspects: 1) All energy packets of an EV arrive when the EV arrives, but information packets arrive randomly; 2) The length of an energy packet remains unknown until service starts in this packet, but the length of an information packet is known when it arrives; 3) The service of an energy packet can be interrupted and is thus preemptive, while that of an information packet is non-preemptive.

3.2.2 Coupled versus Independent Server Rates

In an information system, the service rates of different servers are independent. However, the service rates in a physical queueing system are coupled. For example, in Fig. 4a the service rates have to be considered jointly subject to the physical constraints (i.e., (1) and (2)) of voltage and power flow in the sub-grid. In contrast, in an information multi-server queueing system (like miDRR [21]), the service rates of different servers are independent. Thus, the sum rate of servers is not constrained. Due to this feature, when applying miDRR to EV charging, the capacity of a server is only determined by the charging power of its EV. As a result, the total charging power of all servers (i.e., the entire sub-grid) may exceed the allowed capacity of the distribution transformer or cause under voltage in the sub-grid. In other words, miDRR will result in the physical problems in the sub-grid. Moreover, in an information multi-server queueing system, queues are coupled when they connect to the same server. miDRR ensures fairness among queues by considering the specific connections between queues and servers. However, queues in an EV charging system are coupled through the dependency among servers (i.e., the total capacity is constrained). Since miDRR does not consider such a scenario, it is not effective to ensure fairness among EVs.

3.3 Physical Multi-Server GPS

In an information system, generalized processor sharing is defined to be a discipline that achieves ideal fairness among queues. There are two types of GPSes in the literature, including single-server and multi-server GPSes. The definition of single-server GPS is first formally given in [17], but it is only applicable to queueing systems without queue preferences. In [20], single-server GPS is generalized to multi-server GPS for queueing systems with queue preferences. Due to these preferences, the output rates of queues are coupled. In the physical multi-server queueing system, the output rates of queues are also coupled as explained before. To achieve ideal fairness among EVs, a physical multi-server GPS is designed as follows.

Under multi-server GPS, it is assumed that an EV can be allocated with a rate between zero and its power rating $\hat{P}_i(t)$. Denote the charging rate of EV i at time t as $P_i(t)$, where $0 \leq P_i(t) \leq \hat{P}_i(t)$. If the sub-grid topology and the real-time loads of users are known, we can get voltages and power flows in the sub-grid under different charging rates of EVs using power flow analysis, and then the fair charging rates for EVs at time t can be obtained. Since the charging rate of EV i is allocated according to its weight w_i, the allocated power of this EV is $P_i(t) = w_i x$, where x is a positive real number determined by rate allocation. An iterative rate allocation procedure is executed as follows. Initially, let Δx, a very small number, be the step size of x. In each step, we increase x by Δx, run the power flow analysis, and then check two conditions: 1) The allocated rate for an EV exceeds its power rating, i.e., there exists some EV i with $P_i(t) > \hat{P}_i(t)$; 2) One of the physical constraints is violated. If condition 1) is met, there must be an EV j that is allocated a rate larger than its maximum power. In this case, we fix the charging rate of EV j to be $w_j x$. Since EV j is already charged with its power rating, we exclude EV j and continue the subsequent power allocation by increasing x and checking the two conditions in each step repetitively. Rate allocation terminates when all EVs are allocated rates larger than or equal to their power ratings or when condition 2) is

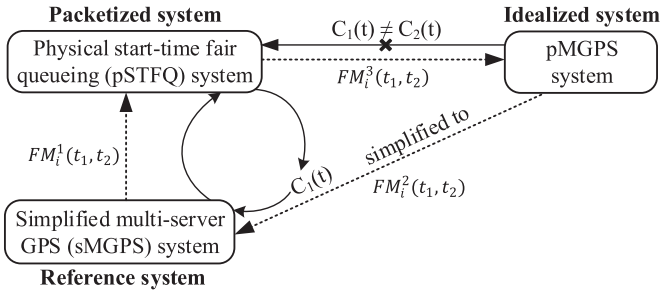


Fig. 5. Overview of the solution to approximate pMGPS.

reached. At this moment, for any EV i if $w_i x < \hat{P}_i(t)$, then $P_i(t) = w_i x$; otherwise, $P_i(t) = \hat{P}_i(t)$.

During rate allocation, the sub-grid capacity is shared among all EVs according to their weights. Since the above rate allocation procedure follows the definition of weighted max-min fairness, it actually achieves weighted max-min fairness. Although pMGPS can achieve ideal fairness among EVs, it cannot be used directly for two reasons: 1) It allocates continuous charging rates to EVs, but EVs can only be charged at their power ratings; 2) It assumes the sub-grid topology is known.

In fact, it is understandable that pMGPS cannot be used directly, as bit-by-bit GPS in an information system cannot be used directly either. In an information system, many packetized schemes (e.g., PGPS [17] and MSFQ [18]) have been designed to approximate GPS in both single-server and multi-server queueing systems. These schemes share the same key idea, i.e., serving packets in the same order as they depart from the GPS system. This key idea of using the GPS system as a reference is easy to realize in an information system, because the same server in the packetized and the GPS systems provides the same service rate (i.e., independent of queues). However, for a sub-grid, charging an EV at different locations has different impacts on user voltage and the power flow through the transformer, so a specific rate allocation is coupled with the grid capacity. For example, the sub-grid can support more EVs at the transformer's side than those at the end of the secondary feeders.

Let $C_1(t)$ and $C_2(t)$ be the sum rate of servers at time t in the packetized system and the pMGPS system, respectively. Since $C_1(t)$ and $C_2(t)$ are determined according to their allocated charging rates of EVs, i.e., the service rates of queues, their values actually depend on what EVs are selected for charging. Thus, both $C_1(t)$ and $C_2(t)$ are queue dependent. Moreover, since pMGPS follows a fluid model, its scheduling result may not be the same as that of the packetized scheme in each charging period. Thus, $C_1(t)$ is not necessarily equal to $C_2(t)$. The discrepancy between pMGPS and the packetized scheme poses a design challenge: the pMGPS system cannot be directly used as a reference to design the packetized system for EV charging. This challenging issue is resolved in the next section.

4 Physical Start-Time Fair Queueing

An overview of the solution to approximating pMGPS is shown in Fig. 5. Instead of using pMGPS as a reference directly, we simplify pMGPS to a simplified multi-server GPS (sMGPS) by replacing the sub-grid constraints in the pMGPS system with $C_1(t)$. In this way, the sMGPS system always has the same sum rate of servers as the packetized system and thus can be used as a reference for the packetized system. $C_1(t)$ is obtained by measuring and summing the charging rates of all EVs that are supported in the sub-grid during the period before time t.

Let $W_i^F(t_1,t_2)$ be the amount of energy EV i is charged in F system during time $[t_1,t_2]$, where F can be sMGPS, pSTFQ, or pMGPS. Since pMGPS achieves ideal fairness, the difference between $W_i^{\text{pMGPS}}(t_1,t_2)$ and $W_i^{\text{pSTFQ}}(t_1,t_2)$ indicates the gap of fairness performance between pMGPS and pSTFQ for EV i at any time t. In other words, the difference denoted by $FM_i^3 = W_i^{\text{pMGPS}}(t_1,t_2) - W_i^{\text{pSTFQ}}(t_1,t_2)$ provides a fairness measure for EV i. The smaller FM_i^3 is, the better fairness pSTFQ achieves for EV i. Considering all EVs in the entire system, the fairness measure is represented by FM^3.

In the following sections, we will study FM_i^3 by respectively analyzing the bounds of FM_i^1 and FM_i^2, where $FM_i^1(t_1,t_2) = W_i^{\text{sMGPS}}(t_1,t_2) - W_i^{\text{pSTFQ}}(t_1,t_2)$ indicates the performance gap between sMGPS and pSTFQ for EV i during $[t_1,t_2]$ and $FM_i^2(t_1,t_2) = W_i^{\text{pMGPS}}(t_1,t_2) - W_i^{\text{sMGPS}}(t_1,t_2)$ captures the performance gap between pMGPS and sMGPS for EV i during $[t_1,t_2]$. Considering all EVs in the system, the two performance gaps are represented by FM^1 and FM^2, respectively.

4.1 Simplified Multi-Server GPS

In the sMGPS system, the rate sum of servers is equal to $C_1(t)$, which is allocated to EVs as follows,

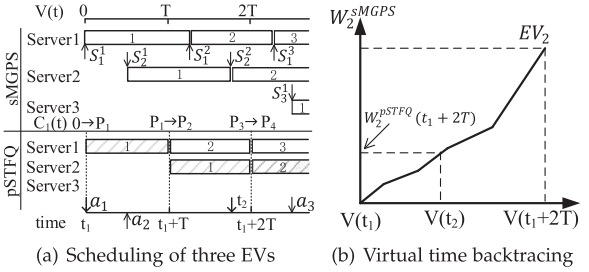
$$P_i(t) = \min(\hat{P}_i(t), w_i x) \text{ and } \sum\nolimits_{i\in\mathcal{N}} P_i(t) = C_1(t), \quad (3)$$

where $\mathcal{N}$ is the set of all EVs in pSTFQ, x is an independent variable making Eq. (3) hold, and $\hat{P}_i(t) = 0$ for all t if EV i is full of charge or does not need to be charged. For EV i with $\hat{P}_i(t) < w_i x$, $P_i(t) = \hat{P}_i(t)$ and for EV i with $\hat{P}_i(t) \geq w_i x$, $P_i(t) = w_i x$, which is weighted max-min fairness.

In Fig. 5, the relationship among pSTFQ, sMGPS, and $C_1(t)$ is represented as a circle, which shows how the pSTFQ and sMGPS systems interact with each other. After $C_1(t)$ is measured from the pSTFQ system, it is fairly allocated to queues in the sMGPS system. Each queue in the sMGPS system processes packets with the allocated rates. Meanwhile, the sMGPS system, as the reference system, determines the service order of packets in the pSTFQ system. The service order of packets, in turn, impacts the scheduling order of EVs and thus influences the value of $C_1(t)$. This circular process continues repeatedly during EV charging. It should be noted that the maximum values of $C_1(t)$ are different for pSTFQ and sMGPS. In pSTFQ, $C_1(t)$ cannot exceed a certain value that leads to the physical problems in the sub-grid. However, in sMGPS no physical problems are concerned, so $C_1(t)$ is only upper-bounded by the total charging rates of all active EVs in the system.

4.2 Finish-Time versus Start-Time Fair Queueing

Under single-server GPS, if two packets have arrived, the relative finish order of these two packets is independent of future packet arrivals. However, as proved in [21], the finish order of packets under multi-server GPS changes upon the arrivals of future packets. Thus, under multi-server GPS, the



(a) Scheduling of three EVs (b) Virtual time backtracing

Fig. 6. Two examples showing how pSTFQ works.

finish order of packets is determined only if these packets finish service. In other words, we cannot use the finish order of packets in the sMGPS system as a reference for the service of packets in the pSTFQ system. A similar problem has been mentioned in [22]. Instead of using finish times, a start-time queueing algorithm is proposed in [22], which can schedule packets based on their virtual start times before finish times are known. Under sMGPS, when an packet of an EV arrives, its virtual start time is determined, even though its finish time is unknown. For an EV, it is easy to determine the virtual start time of its first packet, but it is nontrivial to do so for the following packets. To solve this problem, a virtual start time backtracing scheme will be presented in Section 4.3.1.

Based on the virtual start times of the first packets in EVs, a proper number of EVs are selected in pSTFQ. These EVs starts charging for a period. Once a charging period starts, a new value of $C_1(t)$ is determined and will be used by sMGPS for later scheduling.

4.3 Detailed Algorithm of pSTFQ

We first give several examples on how to use sMGPS as a reference and then present the detailed algorithm of pSTFQ.

4.3.1 Using sMGPS as a Reference

Suppose a sub-grid with 3 EVs. The sub-grid can only charge 2 EVs simultaneously. Let p_i^k be the kth packet of EV i, S_i^k be the virtual start time of p_i^k and L_i^k be the length of p_i^k. The scheduling of EV charging during the first 3 periods is illustrated in Fig. 6a. Before t_1, no EVs were being charged, so the sum rate of the servers in the pSTFQ system is zero, i.e., $C_1(t) = 0$. When EV 1 arrives at $a_1 = t_1$, packets of EV 1 are backlogged into queue 1 in both the pSTFQ and sMGPS systems. We assume the virtual time of the sMGPS system, $V(t)$, is zero at t_1, i.e., $V(t_1) = 0$, so the first packet of EV 1, p_1^1, is tagged with $S_1^1 = 0$. As explained before, the schedule of EV charging is determined based on virtual start time of packets in EV queues. More specifically, EVs are sorted in the increasing order of virtual start times of their head-of-line (HOL) packets. Based on this order, the central controller selects EVs to be charged. At time t_1, only EV 1 has virtual start time, so it is selected for charging. Once EV 1 starts charging, the central controller changes $C_1(t)$ from 0 to $P_1 = \hat{P}_1(t_1)$ as in Fig. 6a. $C_1(t)$ remains to be $\hat{P}_1(t_1)$ in the first charging period, because no other EVs are charged or depart during this period.

Before $t_1 + \frac{T}{2}$, $C_1(t)$ is totally allocated to EV 1 under sMGPS. When EV 2 arrives at $a_2 = t_1 + \frac{T}{2}$, p_2^1 is tagged with $S_2^1 = t_1 + \frac{T}{2}$. After $t_1 + \frac{T}{2}$, $C_1(t)$ is fairly allocated between EV 1 and 2 based on Eq. (3) under sMGPS until $t_1 + T$. At $t_1 + T$, p_1^1 finishes service under pSTFQ and p_1^2 becomes the HOL packet of EV 1. Similarly, at $t_1 + T$, we can have a sorted order of EVs based on the virtual start times of their HOL packets. However, since p_1^2 has not started service under sMGPS at $t_1 + T$, S_1^2 is unknown at $t_1 + T$ and only p_2^1 has known virtual start time. In this case, since the sub-grid can charge two EVs, it is necessary to select p_1^2 to maximize the utilization of sub-grid capacity. For any HOL packet p_i^j, if it has unknown virtual start time, it means the service start time of this packet in pSTFQ is ahead of that in sMGPS. In other words, for EV i we must have $W_i^{\text{pSTFQ}} > W_i^{\text{sMGPS}}$, i.e., FM_i^1 is negative.

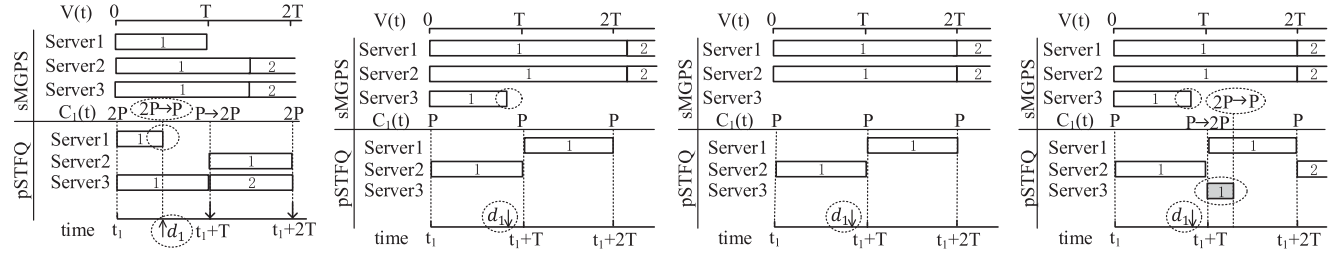
For EVs with negative FM^1s, to prevent FM^1 from further decreasing, HOL packets with larger FM^1 are selected first. In Fig. 6a, the packets with known virtual start times are filled with gray hatch. They are always selected before the packets with unknown virtual start times. After the central controller turns on EVs 1 and 2, $C_1(t_1 + T)$ becomes $P_2 = \hat{P}_1(t_1 + T) + \hat{P}_2(t_1 + T)$.

Contrary to p_1^1, p_2^1 finishes service under sMGPS before it finishes service under pSTFQ. The length of p_2^1 (i.e., L_2^1) is known at $t_1 + 2T$ when it finishes service under pSTFQ. However, when p_2^1 finishes service under sMGPS at some time $t_2 < t_1 + 2T$, L_2^1 is still unknown. To determine the virtual start time of p_2^2 (i.e., S_2^2) at $t_1 + 2T$, we need a virtual time backtracing scheme, as explained below. First, the energy charged over time in sMGPS is recorded for each EV. In this example, $W_2^{\text{sMGPS}}(t)$ for EV 2 needs to be recorded as shown in Fig. 6b where the curve of $W_2^{\text{sMGPS}}(t)$ with respect to the virtual time $V(t)$ is depicted. Second, at $t = t_1 + 2T$, after p_2^1 finishes service in pSTFQ, the energy that has been charged, i.e., $W_2^{\text{pSTFQ}}(t + 2T)$, is determined. We then find the virtual time $V(t_2)$ such that $W_2^{\text{sMGPS}}(V(t_2))) = W_2^{\text{pSTFQ}}(t + 2T)$. Finally, S_2^2 is found, which is equal to $V(t_2)$.

4.3.2 Handling Special Cases Due to EV Departure

In information queueing systems, when a packet arrives, it departs from the queueing system after it finishes service. However, for EV charging, all packets of an EV arrive when the EV arrives, but not all packets can finish service when the EV departs. For any EV i departing at time t, it departs with two cases: $FM_i^1(t_1, t) < 0$ and $FM_i^1(t_1, t) \geq 0$, where t_1 is the beginning time of EV charging. In either case, the energy charged in sMGPS becomes different from that in pSTFQ. Without careful handling of these two cases, the energy difference grows constantly, and eventually sMGPS will lose its reference role for pSTFQ. To handle these two cases, we need to achieve two goals: 1) the energy of remaining EVs needs to be balanced in sMGPS and pSTFQ; 2) the energy of the departing EV needs to be balanced in sMGPS and pSTFQ. To illustrate our solution, four examples are presented in Fig. 7 where the dotted circles highlight the changes associated with an EV's departure. In all examples, three EVs start charging at time t_1 and all have an equal charging rate P and an equal weight.

The first example is shown in Fig. 7a where the sub-grid can charge two EVs simultaneously and EV 1 departs at time $d_1 = t_1 + \frac{1}{2}T$. At d_1, EV 1 departs with $FM_1^1(0, d_1) < 0$, i.e., the energy EV 1 is charged under pSTFQ is larger than that under sMGPS. In the pSTFQ system, the service of p_1^1



(a) Case 1 (b) Case 2: basic scheme (c) Case 2: ideal scheme (d) Case 2: with virtual packets

Fig. 7. Examples showing how pSTFQ works when an EV departs.

stops and L_1^1 is set to the amount of energy that has been charged for p_1^1 in the pSTFQ system, and then $C_1(t)$ decreases from $2P$ to P. The decrement of $C_1(t)$ reduces the allocated rates for p_2^1 and p_3^1 in the sMGPS system. More specifically, at $t = d_1$, $FM_2^1(t_1, t) + FM_3^1(t_1, t) = \frac{1}{6}PT$, which implies the allocated energy to EV 2 and 3 under sMGPS is larger than that under pSTFQ, while at $t = t_1 + T$, $FM_2^1(t_1, t) + FM_3^1(t_1, t) = 0$, meaning the energy balance between sMGPS and pSTFQ is recovered. Upon the departure of EV 1, all its packets after p_1^1 are deleted from both the pSTFQ and sMGPS systems. p_1^1 in the sMGPS system continues to be served until it reaches L_1^1 at $t_1 + T$. After this moment, queue 1 in both the pSTFQ and sMGPS systems becomes empty and $FM_1^1(t_1, t_1 + T) = 0$.

In the second example in Fig. 7b, the sub-grid can only charge one EV at a time and EV 3 departs at time d_1. At d_1, part of p_3^1 has been served under sMGPS, but it has not started service under pSTFQ. EV 3 departs with $FM_3^1(0, d_1) \geq 0$. Since for EV 3, the allocated energy under sMGPS is larger than that under pSTPQ and it departs at d_1, there is no need to allocate more energy to EV 3 under sMGPS after its departure. The service of p_3^1 under sMGPS stops upon EV 3 departure and L_3^1 is set to the amount of energy that has been charged for p_3^1 under sMGPS. After EV 3 departs, the energy allocated to EVs 1 and 2 under sMGPS is less than that under pSTFQ, i.e., $FM_1^1(t_1, d_1) + FM_2^1(t_1, d_1) = -\frac{1}{3}P(d_1 - t_1)$. If the pSTFQ system continues following the basic scheme in Fig. 7b by selecting EV 2 at $t_1 + T$ without further adjustments, the energy difference between the sMGPS and pSTFQ systems for EV 1 and 2 will continue to exist and will further deteriorate due to the departure of other EVs with $FM^1 \geq 0$. Before discussing how to make adjustments, we first look at the ideal scheme in Fig. 7c.

In the ideal scheme, we assume that the departure time of EV 3 is known in advance. In this case, no rate should be allocated to serve p_3^1 after EV 3 arrives and $C_1(t)$ is evenly allocated to EV 1 and 2. At both $t = t_1 + T$ and $t_1 + 2T$, we have $FM_1^1(t_1, t) + FM_2^1(t_1, t) = 0$, and $FM_3^1(t_1, t) = 0$ too as EV 3 has never been serviced in pSTFQ. Since it is impossible to know the departure time of EV 3 in advance, we cannot take this approach to ensure energy balance between sMGPS and pSTFQ. However, an effective scheme that can recover the energy balance is to assign higher charging rates to EVs 1 and 2 such that $FM_1^1(t_1, t) + FM_2^1(t_1, t) = 0$ at $t = t_1 + 2T$. In other words, we need to increases capacity for $C_1(t)$ without changing the actual charging process in the physical system. To this end, we introduce a new packet called *virtual packet* into the queueing system. This virtual packet is created for the departing EV (i.e., EV 3 in this example) and has the same length as p_3^1 in the sMGPS system. When it is serviced, it does not actually consume energy in the physical system (i.e. pSFTQ), but increases capacity $C_1(t)$ for sMGPS.

The scheduling scheme based on a virtual packet is shown in Fig. 7d where a virtual packet (in gray color) is created for EV 3 at $t_1 + T$. When the virtual packet is to be serviced, it is selected after all other packets, i.e., p_1^1 and p_2^2. At $t_1 + T$, p_1^1 has smaller virtual start time than p_2^2, so p_1^1 is selected first. After the selection, since the sub-grid can only support one EV at a time, p_2^2 cannot be selected. However, since the virtual packet has no impact on the sub-grid, it can still be selected. Thus, EV 3 starts charging virtually until the virtual packet finishes its service before $t_1 + 2T$. Comparing the energy allocated to EVs 1 and 2 under sMGPS with that of the ideal scheme, we can see the result is the same. Thus, $FM_1^1(t_1, t) + FM_2^1(t_1, t) = 0$ is achieved. Furthermore, since the virtual packet of EV 3 has been serviced in pSTFQ and has the same length as p_3^1 in sMGPS, we also have $FM_3^1(t_1, t) = 0$.

The above examples show that our solutions to handling the departure of EVs are effective.

4.3.3 pSTFQ

Let $\mathcal{Q}^F(t)$ be the set of nonempty queues in F system at time t, where F can be sMGPS, pSTFQ, or pMGPS. A complete algorithm of pSTFQ consists of three parts.

Part 1) When EV i is connected to the sub-grid at time t, if it is not fully charged, EV i is considered as an active EV. All packets of EV i are backlogged in queue i in the pSTFQ and sMGPS systems, and queue i is added into $\mathcal{Q}^{\text{sMGPS}}(t)$ and $\mathcal{Q}^{\text{pSTFQ}}(t)$. Since the required energy of EV i is unknown, the total length of packets in queue i is initially set to be infinite. $\hat{P}_i(t)$ can be measured when EV i is connected to the grid.

Part 2) At the beginning of each charging period, the central controller first uses the virtual time tagging method to find the virtual start times for all HOL packets, if the virtual start times are available. Then, the central controller determines the service order for EVs with the following steps:

1) For each queue, check if its HOL packet is a virtual packet. If queue i is in $\mathcal{Q}^{\text{pSTFQ}}(t)$, but not in $\mathcal{Q}^{\text{sMGPS}}(t)$, then the HOL packet in queue i must be virtual and queue i is put into a set $\mathcal{V}(t)$, which is defined as the set of queues with a virtual HOL packet.
2) For queues that are not inside $\mathcal{V}(t)$ but have known virtual start times in their HOL packets (i.e., defined as type A queues), they are sorted in the increasing order of the virtual start times of HOL packets.

3) For queues that are not inside $\mathcal{V}(t)$ and have unknown virtual start times in their HOL packets (i.e., defined as type B queues), their FM^1 are calculated and then they are sorted in the decreasing order of the FM^1s of their queues.
4) For queues inside $\mathcal{V}(t)$ (i.e., defined as type C queues), they are sorted in the increasing order of the virtual start times of virtual HOL packets.

After the service order of all queues is determined, the service order of EVs is determined accordingly. Based on the service order in each type of queues, the central controller considers EVs from type A queues first, then type B, and finally type C. We need to turn on as many EVs as possible until some conditions are reached. These conditions include: 1) all EVs are turned on; 2) turning on one more EV will violate the physical constraints (1) and (2); 3) turning on one more EV will cause the total charging rate of EVs to be larger than that of active EVs in $\mathcal{Q}^{\text{sMGPS}}(t)$. Condition 3) ensures the rate sums of servers in the pSTFQ and sMGPS systems to be always equal. For type A and type B queues, all these conditions should be checked. However, for type C queues, condition 3) is not applicable. Since the sub-grid topology is unknown, in order to check condition 2), sensitivity analysis is adopted. Let α_{ij} be the voltage sensitivity of user i with respect to user j and β_i be the loading sensitivity of transformer to user i, which can be obtained with smart meters using the method in [14]. Let $\mathcal{S}$ be the set of EVs that can be turned on. If EVs in $\mathcal{S}$ are turned on, the voltage change at user i is $\Delta V_i = \sum_{j\in\mathcal{S}} \alpha_{ij}\hat{P}_j(t)$ and the power flow change at the transformer is $\Delta S_0 = \sum_{i\in\mathcal{S}} \beta_i \hat{P}_i(t)$. If either ΔV_i exceeds the safety margin or ΔS_0 causes S_0 to be larger than its maximum value, then Condition 2) is reached. To obtain the largest set of $\mathcal{S}$, the central controller first collects users' voltages and the power flow through the transformer, and add EVs into this set until any of the above three conditions is reached. After this step, EVs in $\mathcal{S}$ are turned on, and then for each EV $i \in \mathcal{S}$ a new $\hat{P}_i(t)$ can be measured and updated. However, even if all EVs in S are turned on, it is possible that some capacity of the sub-grid is not utilized. There are two reasons for such under-utilization. The first one is that sensitivity analysis does not really provide fully accurate decision for condition 2). To reduce such under-utilization, we iteratively further consider EVs that are not turned on until no additional EVs can be added to $\mathcal{S}$. Suppose this iterative process ends upon checking condition 2), the un-utilized capacity of the sub-grid is at most the charging rate of an EV. However, this amount of under-utilization can be further reduced. As shown in [23], sensitivity analysis is adequate to approximate voltage and power flow changes in distribution grid. Thus, with sensitivity analysis based on [14], [23], the error of checking if an additional EV can be turned on is very small.

It should be noted that, even if the largest set of $\mathcal{S}$ is formed, the sub-grid capacity may still be under-utilized. This is the second reason for capacity under-utilization, which is explained as follows. We know that the power rating of an EV is discrete, so it is possible that adding one more EV into $\mathcal{S}$ will violate the physical constraints, but without adding this EV, the sub-grid capacity is under-utilized. Such an EV that cannot be turned on due to the discrete charging rate is called an *unlucky EV*.

Part 3) When EV i departs or is fully charged at time t, two cases need to be considered, as mentioned in Section 4.3.2. If $FM_i^1(0,t) > 0$, queue i is deleted from $\mathcal{Q}^{\text{sMGPS}}(t)$ and the total length of packets backlogged in queue i in the pSTFQ system is set to $FM_i^1(0,t)$. EV i becomes virtual and has no physical effects on the sub-grid. Queue i is deleted from $\mathcal{Q}^{\text{pSTFQ}}$ at some time t' where $FM_i^1(0,t') = 0$. If $FM_i^1(0,t) \le 0$, queue i is deleted from $\mathcal{Q}^{\text{pSTFQ}}(t)$ and the total length of packets backlogged in queue i in the sMGPS system is set to $-FM_i^1(0,t)$. Queue i is deleted from $\mathcal{Q}^{\text{sMGPS}}$ when all the packets in queue i finish service in the sMGPS system.

5 Performance Analysis of pSTFQ

Unlike information systems, the energy difference between pSTFQ and pMGPS cannot properly characterize fairness, because the energy difference varies with the weights and charging rates of EVs and even depends on the safety margin. In other words, FM^3 (i.e., the energy gap between pSFTQ and pMGPS) itself cannot provide an appropriate measure for long-term fairness. Usually FM^3 increases with time, which is reasonable since the total charged energy of an EV also increases with time. To provide an appropriate measure of long-term fairness, FM^3 needs to be normalized with the total charged energy. Thus, a new metric called energy difference ratio (EDR) is defined as follows:

$$EDR = \max_{i\in\mathcal{N}} \left\{ \frac{FM_i^3(t_1,t_2)}{W_i^{\text{pSTFQ}}(t_1,t_2)} \right\}, \tag{4}$$

where FM^3 of EV i is normalized over its total charged energy W_i^{pSTFQ} and then the maximum value among all EVs is selected. As a result, a small EDR indicates that, even for the worst-case EV, the energy gap between pSTFQ and pMGPS is small as compared to the total charged energy in this EV.

It should be noted that $t_2 - t_1$ in Eq. (4) needs to be as large as days or months instead of just one time of charging, which is much different from the case in an information system. In an information system, packets depart only if they finish service. However, packets in an EV charging system may depart any time before they get service. Thus, short-term fairness measure cannot properly reflect the fairness performance of EV charging.

To analyze EDR, the parameter $FM_i^3(t_1,t_2)$ needs to be investigated, but it is difficult to derive the bounds for $FM_i^3(t_1,t_2)$ directly. However, as explained in Section 4, $FM_i^3(t_1,t_2)$ can be analyzed by studying $FM_i^1(t_1,t_2)$ and $FM_i^2(t_1,t_2)$, respectively. In the the rest of this section, the bounds of $FM_i^1(t_1,t_2)$ and $FM_i^2(t_1,t_2)$ are derived. The actual bound of $FM_i^3(t_1,t_2)$ will be studied further via simulations in Section 6.

Let $\mathcal{Q}_-^F(t)$ be the set of queues in $\mathcal{Q}^F(t)$ with $FM^1 < 0$ and let $\mathcal{Q}_+^F(t)$ be the complement set of $\mathcal{Q}_-^F(t)$ with respect to $\mathcal{Q}^F(t)$, where F can be pSTFQ or sMGPS.

Lemma 1. *The pSTFQ and sMGPS systems have the same busy periods, so it suffices to bound the fairness measures of EVs for each busy period of the pSTFQ system. In particular, for any time t,*

$$\sum_{i\in\mathcal{N}} W_i^{\text{pSTFQ}}(0,t) = \sum_{i\in\mathcal{N}} W_i^{\text{sMGPS}}(0,t). \quad (5)$$

Proof. The proof is in Appendix A. □

Lemma 2. *For any EV i, $FM_i^1(0,t)$ reaches the minimum when a packet in queue i finishes service in the pSTFQ system and reaches the maximum when a packet in queue i begins service in the pSTFQ system, which implies that $FM_i^1(0,t)$ reaches the maximum and minimum only at the beginning of charging periods.*

Proof. The proof is in Appendix B. □

Lemma 3. *Let $[t_1,t_2]$ be a time interval, during which $\hat{P}_i(t)$ and $\hat{P}_j(t)$ are nearly constant, where $\hat{P}_i(t)=\hat{P}_i$ and $\hat{P}_j(t)=\hat{P}_j$. For any queues i and j that are continuously backlogged during $[t_1,t_2]$ in the sMGPS system,*

$$\frac{W_i^{\text{sMGPS}}(t_1,t_2)}{W_j^{\text{sMGPS}}(t_1,t_2)} \le \max\left(\frac{\hat{P}_i}{\hat{P}_j}, \frac{w_i}{w_j}\right). \quad (6)$$

Proof. The proof is in Appendix C. □

Lemma 4. *For any time t,*

$$\sum_{i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)} FM_i^1(0,t) + \sum_{i\in\mathcal{Q}_-^{\text{sMGPS}}(t)} FM_i^1(0,t) = 0.$$

Proof. The proof is in Appendix D. □

Theorem 1. *For any time t and any queue i,*

$$-W^+ \max_{i\in\mathcal{N}}\left\{\frac{L_i^{\max}}{w_i}\right\} \le FM_i^1(0,t) \le w_i \max_{i\in\mathcal{N}}\left\{\frac{L_i^{\max}}{w_i}\right\},$$

where $W^+ = \sum_{i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)} w_i$.

Proof. The proof is in Appendix E. □

Theorem 1 shows that for an EV with a fixed weight, its FM^1 upper bound is fixed, but there is no fixed lower bound. In fact, all the fair queueing schemes using GPS as a reference have no fixed lower fairness bound, as mentioned in [17], [21]. The lower bound is not a concern, because when the FM^1 of an EV is negative, it indicates that pSTFQ outperforms sMGPS for the EV. Nonetheless, based on Lemma 4, we can see that large negative FM^1s of some EVs will push the FM^1s of other EVs close to their upper bounds, so we still want to study more on the lower bound. Note that the lower bound in Theorem 1 can be reached only under certain EV departure patterns, which is of low probability. To capture the common cases, the expected lower bound is derived in Theorem 2 given as follows.

Theorem 2. *For any time t and any queue i,*

$$\mathbb{E}[FM_i^1(0,t)] > -(\lfloor \log_2|\mathcal{N}| \rfloor + 1)L_i^{\max}, \quad (7)$$

where $|\mathcal{N}|$ is the number of EVs in $\mathcal{N}$ and $\lfloor x \rfloor$ is the largest integer less than or equal to x.

Proof. The proof is in Appendix F. □

Theorem 3. *During $[t_1,t_2]$, if for all active EVs i and j, $\frac{\hat{P}_i(t)}{w_i} = \frac{\hat{P}_j(t)}{w_j}$, we have*

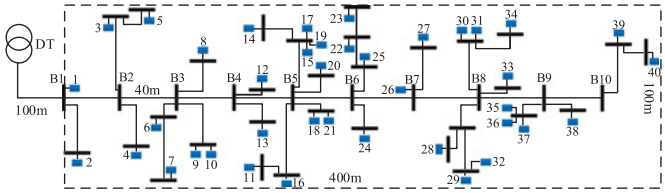


Fig. 8. Sub-grid topology used in experiments.

$$\mathbb{E}[FM_i^2(t_1,t_2)] \le \int_{t_1}^{t_2}\left(\frac{w_i P'(t)}{\sum_{i\in\mathcal{N}(t)} w_i} + \frac{w_i \Delta V_{\min}}{\sum_{i\in\mathcal{N}(t)} \beta_{ij} w_j}\right)\delta(t)dt,$$

where $\mathcal{N}(t)$ is the set of active EVs at time t, $P'(t)$ is the power rating of the unlucky EV at time t and $\delta(t)=1$ when the sub-grid capacity cannot support the charging of all active EVs at time t; otherwise, $\delta(t)=0$.

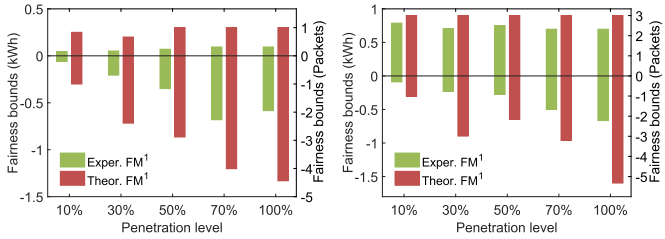
Proof. The proof is in Appendix G. □

In Theorem 3, the FM^2 bound is closely related to the discrete EV charging rates and the safety margin. If no safety margin is used in pSTFQ just as that in pMGPS, the second item of the FM^2 bound can be eliminated. The first item of the FM^2 bound cannot be eliminated since the EV charging rates are not necessarily zero. However, this term is not significant for the following reason. First, when $\delta(t)=1$, the number of active EVs is large, since a sub-grid that can only charge a small portion of EVs should be reinforced. Second, the weight of an EV is small as compared to the sum of the weights of all active EVs. It should be noted that Theorem 3 only includes the case when EV power ratings are allocated proportionally to their weights. The FM^2 bounds under other cases are verified via simulations.

6 Performance Evaluation

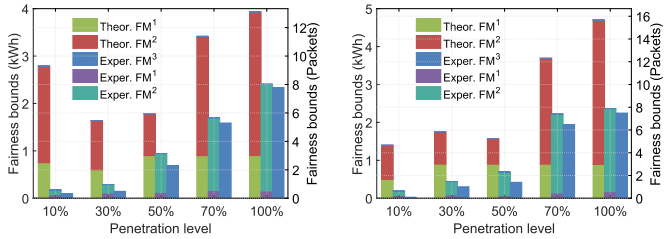
6.1 Experiment Setup

Experiments are conducted under several charging scenarios using MATPOWER 4.1 [24]. The sub-grid is a single-phase 220 V network having a topology illustrated in Fig. 8. There are totally 40 users randomly located in a 400 m×100 m rectangular area. A distribution transformer is 100 m away from bus 1 and the distances between adjacent buses from bus 1 to 10 are all equal to 40 m. The layout of transmission lines and other buses is formed by connecting users to the nearest one of the ten buses. All transmission lines have the same impedance, which is $(0.3+j\,0.08)\,\Omega$/km. Each user has one EV with random power rating P kW and the required energy, E, is randomly distributed in $[0,8P]$ kWh, where $P\in\{3,4,5,6\}$. The variable charging rate of an EV is modelled as $P\times 4^{(\text{SOC}-1)}$, where $\text{SOC}=\frac{E}{8P}$. For any EV i, w_i is selected from $\{1,2,3\}$. The arrival patterns of EVs follow the same distribution as provided in Fig. 2 in [4], and and EV users depart randomly in the range of $[0,8]$ hours since arrival. Besides the EV, each user has 15 types of appliances with an average power factor of 0.9, where 5 appliances have rated power between 1 and 2 kW, and the rest have rated power less than 1 kW. The load changes of each user are modelled as a Poisson process with an average frequency equal to 7 times/hour. All home appliances are prescribed to operate within the voltage range from 0.92 to 1.042 p.u., where p.u.



(a) Case 1: fixed weights (b) Case 2: random weights

Fig. 9. FM^1 bounds of the worst-case EVs under pSTFQ.



(a) Case 1: fixed weights (b) Case 2: random weights

Fig. 10. Performance gap among pSTFQ, sMGPS and pMGPS.

is the normalized voltage with respect to the base voltage 220 V. The safety margin of voltage is $\Delta V = 0.01$ p.u. and the charging period, T, is 3 minutes. The voltage and loading sensitivities are obtained using linear approximation in [14].

6.2 Fairness Performance of pSTFQ

We first verify if the experimental bounds of FM^1 are within its theoretical bounds. The experimental bounds of FM^1 are its maximum or minimum values in one charging scenario. For each charging scenario, experimental results are collected for 30 days of EV charging. Experiments are conducted under different charging scenarios. There are two cases for EV weights: 1) EVs have fixed weights; 2) EVs randomly select a weight for each new charging. For each case, experiments are conducted under 5 penetration levels. Moreover, two capacities are considered for the distribution transformer, i.e., 60 kVA and 120 kVA representing the cases when the sub-grid capacity is constrained by transformer capacity or by voltage, respectively.

As EVs have different FM^1 bounds, we verify the bounds by checking the worst-case EVs. EV i is defined to be a worst-case EV if it has the largest value of its experimental upper bound minus $w_i \max_{i\in\mathcal{N}}\{\frac{L_i^{\max}}{w_i}\}$ or the smallest value of its experimental lower bound minus $-(\lfloor \log_2|\mathcal{N}| \rfloor + 1)L_i^{\max}$. Thus, if the experimental bounds of the worst-case EVs are within their theoretical bounds, the experimental bounds of all other EVs follow the same.

The FM^1 bounds of the worst-case EVs are shown in Fig. 9 where the results under two transformer capacities are merged by taking the worse result. The bounds are shown in two units, and the number of packets is obtained by normalizing the energy by the maximum packet length of all EVs, i.e., 0.3 kWh in this simulation. In all cases, the experimental bounds of FM^1 are bounded by the theoretical bounds. Further, the experimental upper bounds are very close to the theoretical upper bounds when EVs have random weights, which implies the theoretical upper bounds are tight.

With the above setup of EV charging, we also evaluate the bounds of FM^2 and FM^3. Unlike FM^1, the upper bound of FM^2 increases with time, as shown in Theorem 3. Thus, the upper bound of FM^3 also increases with time. To this end, the upper bounds of FM^3 are obtained for each time of EV charging instead of the entire period of 30 days. More specifically, we take two steps to get the upper bound of FM^3: 1) For each EV, considering all times of charging within 30 days, the maximum upper bound is selected, and this particular time of charging (e.g., on a certain day) is recorded; 2) The maximum upper bound is then chosen among all EVs, and the selected EV is recorded. Given the selected EV and its particular time of charging, the corresponding upper bounds (both theoretical and experimental) of FM^1 and FM^2 are then determined. The results of these upper bounds are shown in Fig. 10. Since two scenarios of transformer capacity(60 and 120 kVA) are considered in our experiments, we merge two results of each upper bound into one by taking the worse case. From Fig. 10, it can be seen that the largest gap between pSTFQ and pMGPS is about eight packets. However, such a large gap occurs with extremely low probability (see column 4 in Table 1).

To show more details of FM^3, the distribution of its experimental upper bound for each time of EV charging is shown in Table 1. For all penetration levels, larger than 99.7 percent of these upper bounds of FM^3 are less than 5 packets. For 30 days of EV charging, the EDR is 3.45 percent under various scenarios, implying pSTFQ closely approximates pMGPS.

In Fig. 10, all the experimental bounds of FM^2 are properly bounded by the theoretical bounds, even though the theoretical bounds are not very tight. The sum of the experimental bounds of FM^1 and FM^2 is always slightly larger than FM^3, which implies combining the bounds of FM^1 and FM^2 is effective to approximate the bounds of FM^3. Further, the bounds of FM^2 are much larger than these of FM^1, which implies the bounds of FM^3 are mainly resulted from the gap between sMGPS and pMGPS and the gap exists due to discrete EV charging rates and the safety margin.

To study the impact of the sub-grid capacity, we evaluate the performance of EDR versus different transformer capacities. The transformer capacity varies from 60 to 100 kVA, where 60 kVA is the least possible transformer capacity since it is equal to the peak rate of all non-EV appliances. When the transformer capacity is 60 kVA, the sub-grid capacity is constrained by the transformer capacity; when the transformer capacity is 100 kVA, the sub-grid capacity is constrained by voltage. The simulation results are shown in Fig. 11. Under different penetration levels, EDR decreases as transformer capacity increases. This result is reasonable: when the transformer capacity is larger, an EV can be charged with more energy during the same period, so the

TABLE 1
Distribution of the Experimental Bounds of FM^3 in 30 Days

Penetration level	Experimental bound (=x) of FM^3 (packets)			EDR
	$0 \le x \le 2$	$2 < x \le 4$	$5 < x \le 9$	
100%	94.5%	5.2%	0.3%	3.46%
70%	98.4%	1.5%	0.1%	1.50%
50%	99.9%	0.1%	0	0.43%

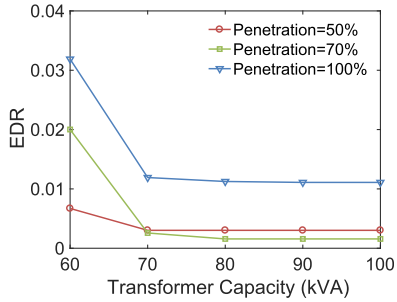
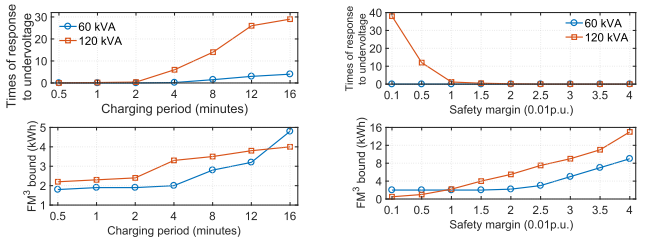


Fig. 11. EDR versus different capacities of the distribution transformer.



(a) Case 3: fixed safety margin (b) Case 4: fixed charging period

Fig. 12. Performance gap under different charging periods and safety margins.

performance gap between pMGPS and pSTFQ amounts to a smaller percent of total charged energy. Thus, increasing transformer capacity improves fairness performance. Moreover, as shown in Fig. 11, the largest EDR in all scenarios is below 4 percent, which illustrates satisfactory fairness performance achieved by our proposed scheme.

6.3 Performance Gap Under Different System Parameters

The performance gaps among the three systems are studied under different charging periods and safety margins. Simulations are conducted under two cases as shown in Fig. 12. In Case 3, the safety margin is fixed to be 0.01 p.u. and the charging periods change from 0.5 to 16 minutes. In Case 4, the charging period is fixed to 3 minutes and the safety margin changes from 0.001 to 0.04 p.u. In both cases, simulations are conducted under 100 percent penetration level of EVs for 30 days and EVs have fixed weights. Two performance metrics, the times of response to under-voltage and the bounds of FM^3, are collected. Fig. 12a shows that the bounds of FM^3 are not sensitive to the charging periods under both transformer capacities, but the times of response to under-voltage are significantly reduced with decreasing charging periods when the transformer capacity is 120 kVA. As opposed to the charging period, the safety margin closely impacts the bounds of FM^3. As shown in Fig. 12b, when safety margin drops, the bounds of FM^3 decreases significantly. When the safety margin decreases to 0.001 p.u., the bound of FM^3 is about the length of two packets. These results clearly demonstrate that the gap between pSTFQ and pMGPS can be reduced by using a smaller safety margin. A similar conclusion can be made to the gap between sMGPS and pMGPS with respect to the discrete charging rates of EVs. However, we cannot use a very small safety margin; otherwise, the response time to under-voltage will not be held below a satisfactory level.

7 Conclusion

In a distribution grid, charging EVs at home results in the under-voltage and transformer overloading problems and also unfairness among EVs. These problems were solved by a physical fair queueing framework developed in this paper. This framework consists of two fair queueing schemes: physical multi-server GPS and physical start-time fair queueing. pMGPS captures the characteristics of EV charging in a sub-grid, but cannot be directly applied, so pSTFQ was then designed as a practical packetized scheme to schedule EV charging. The performance gap between these fair queueing schemes was studied through theoretical analysis and was also evaluated via simulations. Both analysis and simulations showed that the pSTFQ scheme gracefully avoided the physical problems such as under-voltage and transformer overloading. Moreover, the energy difference ratio between pSTFQ and the idealized pMGPS is less than 4 percent in different scenarios, which illustrated that fair sharing of energy among EVs was achieved. How to design a communication network to support the scheduling scheme developed in this paper will be studied in the future work. It should be noted that the physical fair queueing framework can be applied to other scheduling scenarios in cyber-physical systems.

Appendix A
Proof of Lemma 1

Let us consider the first busy period where the pSTFQ and sMGPS systems both start at $t = 0$ and become idle at time t_1 and t_2 respectively. Assume EV i finishes charging at time t'. If $W_i^{\mathrm{sMGPS}}(0,t') > W_i^{\mathrm{pSTFQ}}(0,t')$, queue i is first deleted from $\mathcal{Q}^{\mathrm{sMGPS}}(t')$ and then deleted from $\mathcal{Q}^{\mathrm{pSTFQ}}$ at some time t^* where $W_i^{\mathrm{pSTFQ}}(0,t^*) = W_i^{\mathrm{sMGPS}}(0,t')$, i.e., all backlogged packets in the pSTFQ system finish service. If $W_i^{\mathrm{sMGPS}}(0,t') \leq W_i^{\mathrm{pSTFQ}}(0,t')$, queue i is first deleted from $\mathcal{Q}^{\mathrm{pSTFQ}}$ and then deleted from $\mathcal{Q}^{\mathrm{sMGPS}}$ until $W_i^{\mathrm{sMGPS}}(0,t'') = W_i^{\mathrm{pSTFQ}}(0,t')$ at some time t''. Thus, when a queue is empty in both the pSTFQ and sMGPS systems, the amount of services the queue has been served in the two systems should be equal. We have $W_i^{\mathrm{pSTFQ}}(0,t_1) = W_i^{\mathrm{sMGPS}}(0,t_2)$ for all $i \in \mathcal{N}$ and thus $\sum_{i\in\mathcal{N}} W_i^{\mathrm{pSTFQ}}(0,t_1) = \sum_{i\in\mathcal{N}} W_i^{\mathrm{sMGPS}}(0,t_2)$. In addition, considering that the pSTFQ and sMGPS systems have the same output rates, we can conclude $t_1 = t_2$. When the pSTFQ and sMGPS systems are both idle at time t, $W_i^{\mathrm{sMGPS}}(0,t) = W_i^{\mathrm{pSTFQ}}(0,t), \forall i \in \mathcal{N}$. Therefore, the fairness measure of EVs can be bounded for each busy period. Because the pSTFQ and sMGPS systems have the same output rate and the same busy periods, Eq. (5) follows.

Appendix B
Proof of Lemma 2

In the pSTFQ system, when queue i is in service at time t, the slope of W_i^{pSTFQ} is $\hat{P}_i(t)$; otherwise, the slope is 0. In the sMGPS system, the slope of W_i^{sMGPS} varies between 0 and $\hat{P}_i(t)$. Thus, during a charging period if queue i is in service, $FM_i^1(0,t)$ is a non-increasing function with time; otherwise, $FM_i^1(0,t)$ is a non-decreasing function with time. $FM_i^1(0,t)$ reaches the minimum and maximum values at the end of a charging period when a

packet in queue i in the pSTFQ system finishes and begins service respectively. Because packets start and finish service only at the beginning of charging periods, $FM_i^1(0,t)$ reaches the maximum and minimum only at the beginning of charging periods.

APPENDIX C
PROOF OF LEMMA 3

In the sMGPS system, for queues i and j that are continuously backlogged during $[t_1,t_2]$,

$$\frac{W_i^{\text{sMGPS}}(t_1,t_2)}{W_j^{\text{sMGPS}}(t_1,t_2)}=\frac{\int_{t_1}^{t_2}P_i(t)dt}{\int_{t_1}^{t_2}P_j(t)dt}\leq \max_{t\in\mathbb{R}^+}\left\{\frac{P_i(t)}{P_j(t)}\right\},$$

where $\mathbb{R}^+$ is the set of positive real numbers. Based on Eq. (3), these exists an x such that

$$\frac{P_i(t)}{P_j(t)}=\frac{\min(w_i x,\hat{P}_i(t))}{\min(w_j x,\hat{P}_j(t))}.$$

We bound $\frac{P_i(t)}{P_j(t)}$ in three cases. *Case 1:* $P_i(t)<\hat{P}_i(t)$ *and* $P_j(t)<\hat{P}_j(t)$. $\frac{P_i(t)}{P_j(t)}=\frac{w_i}{w_j}$. *Case 2:* $P_i(t)=\hat{P}_i(t)$ *and* $P_j(t)\leq\hat{P}_j(t)$. We have $\frac{P_i(t)}{P_j(t)}\leq\frac{w_i}{w_j}$. *Case 3:* $P_i(t)\leq\hat{P}_i(t)$ *and* $P_j(t)=\hat{P}_j(t)$. We have $\frac{P_i(t)}{P_j(t)}\leq\frac{\hat{P}_i(t)}{P_j(t)}=\frac{\hat{P}_i}{\hat{P}_j}$. Thus, the lemma follows.

APPENDIX D
PROOF OF LEMMA 4

Because EVs may depart at any time t, $\mathcal{Q}^{\text{pSTFQ}}(t)$ is not always equal to $\mathcal{Q}^{\text{sMGPS}}(t)$. Let queue $i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)$ for $t\in[t_1,t_2)$ and depart at time t_2. According to pSTFQ, queue i is in $\mathcal{Q}_+^{\text{sMGPS}}(t)$ during $[t_1,t_2)$. After time t_2, queue i is deleted from $\mathcal{Q}_+^{\text{sMGPS}}(t)$, but queue i is still in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$ until all packets in queue i in the pSTFQ system finish service. Hence, if queue $i\in\mathcal{Q}_+^{\text{sMGPS}}(t)$, queue i must be in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$, i.e., $\mathcal{Q}_+^{\text{sMGPS}}(t)\subseteq\mathcal{Q}_+^{\text{pSTFQ}}(t)$. Similarly, we can prove $\mathcal{Q}_-^{\text{pSTFQ}}(t)\subseteq\mathcal{Q}_-^{\text{sMGPS}}(t)$. The set of queues that are nonempty in either the pSTFQ or sMGPS system is equal to the union of $\mathcal{Q}^{\text{pSTFQ}}(t)$ and $\mathcal{Q}^{\text{sMGPS}}(t)$, denoted as $\mathcal{Q}^{\text{pSTFQ}}(t)\cup\mathcal{Q}^{\text{sMGPS}}(t)$. By the definition of $\mathcal{Q}^{\text{pSTFQ}}(t)$ and $\mathcal{Q}^{\text{sMGPS}}(t)$, we have

$$\begin{aligned}&\mathcal{Q}^{\text{pSTFQ}}(t)\cup\mathcal{Q}^{\text{sMGPS}}(t)\\&\quad=\mathcal{Q}_+^{\text{pSTFQ}}(t)\cup\mathcal{Q}_-^{\text{pSTFQ}}(t)\cup\mathcal{Q}_-^{\text{sMGPS}}(t)\cup\mathcal{Q}_+^{\text{sMGPS}}(t)\\&\quad=\mathcal{Q}_+^{\text{pSTFQ}}(t)\cup\mathcal{Q}_-^{\text{sMGPS}}(t).\end{aligned}\tag{8}$$

For queue $i\notin\mathcal{Q}^{\text{pSTFQ}}(t)\cup\mathcal{Q}^{\text{sMGPS}}(t)$, from Lemma 1, we have $FM_i^1(0,t)=0$. Thus, Eq. (5) becomes,

$$\sum_{i\in\mathcal{Q}(t)}W_i^{\text{pSTFQ}}(0,t)=\sum_{i\in\mathcal{Q}(t)}W_i^{\text{sMGPS}}(0,t),$$

where $\mathcal{Q}(t)=\mathcal{Q}_+^{\text{pSTFQ}}(t)\cup\mathcal{Q}_-^{\text{sMGPS}}(t)$, which implies

$$\sum_{i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)}FM_i^1(0,t)+\sum_{i\in\mathcal{Q}_-^{\text{sMGPS}}(t)}FM_i^1(0,t)=0.\tag{9}$$

APPENDIX E
PROOF OF THEOREM 1

Denote $\sum_{i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)}\hat{P}_i(t)$ as $A(t)$ and the total power that can be used for EV charging as $C(t)$. We first prove the lower bound in three cases.

Case 1: $C(t)<A(t)$ *for all time* t. Let t_1 be the time a packet in queue i in the pSTFQ system finishes service. Then, we have $W_i^{\text{pSTFQ}}(0,t_1)=\sum_{k=1}^{n_i(t_1)}L_i^k$. At time t_1, the packet must have known virtual start time, so it has started service in the sMGPS system, which implies $W_i^{\text{sMGPS}}(0,t_1)>\sum_{k=1}^{n_i(t_1)-1}L_i^k$. Hence, $FM_i^1(0,t_1)\geq -L_i^{n_i(t_1)}\geq -L_i^{\max}$.

Let t_2 be the time a packet in queue k begins service in the pSTFQ system. At time t_2, the $n_i(t_2)$th packet for all queue $i\in\mathcal{N}$ have finished service in the pSTFQ system. We want to prove that there must exist a queue $j\in\mathcal{Q}_-^{\text{pSTFQ}}(t_2)$ such that $S_j^{n_j(t_2)}\leq S_k^{n_k(t_2)+1}$. Let us assume that for all queue $j\in\mathcal{Q}_-^{\text{pSTFQ}}(t_2)$, $S_j^{n_j(t_2)}>S_k^{n_k(t_2)+1}$. Let p' be the packet that has the largest virtual start time among all the $n_i(t_2)$th packets for queue $i\in\mathcal{Q}_-^{\text{pSTFQ}}(t_2)$ and let t_3 be the time p' is served in the pSTFQ system. At time t_3, since p' has the largest virtual start time, if p' is served, all the HOL packets in other queues must be served, i.e., all queues are served. However, as $C(t)<A(t)$, it is impossible to serve the HOL packets from all queues. Thus, there must exist a packet p^* having virtual start time $S^*<S_k^{n_k(t_2)+1}$, where packet p^* is in queue i^*.

Let t^* and t' be the time packet p^* starts and finishes service in the sMGPS system. At time t^*, $p_k^{n_k(t_2)+1}$ has not started service in the sMGPS system, so $W_k^{\text{sMGPS}}(0,t^*)\leq\sum_{m=1}^{n_k(t_2)}L_k^m=W_k^{\text{pSTFQ}}(0,t_2)$. At time t_2, p^* has not started service in the sMGPS system, so $t'>t_2$. According to the charging profile of EVs [15], since $W_k^{\text{sMGPS}}(t^*,t')$ and $W_{i^*}^{\text{sMGPS}}(t^*,t')$ are very small compared to the battery sizes of EVs k and i^*, $\hat{P}_k(t)$ and $\hat{P}_{i^*}(t)$ are nearly constant during $[t',t^*]$. Let $\hat{P}_i$ be the value of $\hat{P}_i(t)$ during $[t',t^*]$. Then,

$$\begin{aligned}FM_k^1(0,t_2)&\leq W_k^{\text{sMGPS}}(0,t')-W_k^{\text{sMGPS}}(0,t^*)\\&\leq\hat{P}_{i^*}T\frac{W_k^{\text{sMGPS}}(t^*,t')}{W_{i^*}^{\text{sMGPS}}(t^*,t')}.\end{aligned}\tag{10}$$

Based on Lemma 3, if $\frac{\hat{P}_k}{\hat{P}_{i^*}}>\frac{w_k}{w_{i^*}}$, $FM_k^1(0,t_2)\leq\hat{P}_kT\leq L_k^{\max}$; otherwise, $FM_k^1(0,t_2)\leq w_k\frac{\hat{P}_{i^*}T}{w_{i^*}}\leq w_k\max_{i\in\mathcal{N}}\{\frac{L_i^{\max}}{w_i}\}$. Therefore, $FM_k^1(0,t)\leq w_k\max_{i\in\mathcal{N}}\{\frac{L_i^{\max}}{w_i}\}$.

Case 2: $C(t)\geq A(t)$ *for all time* t. Queues in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$ are always being served. Let queue $i\in\mathcal{Q}_+^{\text{pSTFQ}}(t)$ for $t\in[t_1,t_2)$. From Lemma 2, we have $FM_i^1(0,t_1)\geq FM_i^1(0,t_2)$. Thus, if $FM_i^1(0,t_1)\leq a_i$, we must have $FM_i^1(0,t)\leq a_i$ for any $t\in[t_1,t_2]$. Let t_1 be any time when $FM_i^1(0,t)$ becomes positive from negative and b_m be the beginning time of the mth charging period. There exists a time b_m such that $t_1\in[b_m,b_{m+1}]$ and $FM_i^1(0,b_m)\leq 0$. From Lemma 2, we can conclude FM_i^1 reaches the maximum at time b_{m+1}. Since $FM_i^1(0,b_m)\leq 0$, $FM_i^1(0,t_1)\leq FM_i^1(0,b_{m+1})\leq L_i^{\max}$. The theorem follows.

Case 3: $C(t)$ *is larger and less than* $A(t)$ *alternatively at different times.* Let $[b_{m_1},b_{m_2})$ be the first time interval where $C(t)>A(t)$ and $[b_{m_2},b_{m_3})$ be the time interval where $C(t)\leq A(t)$. During $t\in[0,b_{m_1})$, we have $C(t)\leq A(t)$, the upper bound of $FM_i^1(0,t)$ follows Case 1 for all $i\in\mathcal{N}$. From Lemma

2, we know that for any queue $i \in \mathcal{Q}_+^{\text{pSTFQ}}(t)$ during $[b_{m_1}, b_{m_2})$, as long as $FM_i^1(0, b_{m_1}) < a_i$, $FM_i^1(0, b_{m_2}) < a_i$ for sure. Then, the upper bound in the theorem follows at time b_{m_2}. We need to obtain the upper bound of $FM_i^1(0,t)$ for $t \in [b_{m_2}, b_{m_3})$.

It is noticeable that during $[b_{m_2}, b_{m_3})$, $C(t)$ becomes less than $A(t)$ again. If the packets scheduling during $[b_{m_1}, b_{m_2})$ does not impact the scheduling during $[b_{m_2}, b_{m_3})$, the fairness upper bound during $[b_{m_2}, b_{m_3})$ should remain the same as that during $[b_{m_1}, b_{m_2})$. From Case 1, we can see that for any EV i, its upper bound is less than $w_i \max_{i \in \mathcal{N}}\{\frac{L_i^{\max}}{w_i}\}$ regardless of how other EVs are charged as long as the HOL packets for queues in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$ are served based on virtual start times. The packet scheduling during $[b_{m_1}, b_{m_2})$ is a special case where all the HOL packets for queues in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$ are served. In other words, the fairness upper bound during $[b_{m_2}, b_{m_3})$ is not affected by the packet scheduling during $[b_{m_1}, b_{m_2})$ and thus is the same as that during $[0, b_{m_1})$. It is easy to see that the upper bound follows in other time intervals.

For the lower bound, as mentioned in [17], [18], if packets with unknown virtual start times are to be served, there does not exist a constant $c < 0$ such that $FM_i^1(0,t) > c$. When EV $i \in \mathcal{Q}_-^{\text{pSTFQ}}(t)$ departs, EV i is deleted from $\mathcal{Q}^{\text{pSTFQ}}(t)$, but is still in $\mathcal{Q}^{\text{sMGPS}}(t)$. If all EVs but one in $\mathcal{Q}_-^{\text{pSTFQ}}(t)$ departs simultaneously at time t, $\mathcal{Q}^{\text{sMGPS}}(t)$ remains unchanged, but there is only one EV k in $\mathcal{Q}_-^{\text{pSTFQ}}(t)$. Based on Lemma 4, since $\mathcal{Q}_-^{\text{pSTFQ}}(t) \subseteq \mathcal{Q}_-^{\text{sMGPS}}(t)$,

$$FM_k^1(0,t) \geq \sum_{i \in \mathcal{Q}_-^{\text{sMGPS}}(t)} FM_i^1(0,t) = -\sum_{j \in \mathcal{Q}_+^{\text{pSTFQ}}(t)} FM_j^1(0,t).$$

The lower bound of $FM_i^1(0,t)$ depends on the fairness bounds of other EVs in $\mathcal{Q}_+^{\text{pSTFQ}}(t)$. Since $FM_j^1(0,t) \leq w_j \max_{i \in \mathcal{N}}\{\frac{L_i^{\max}}{w_i}\}$, the theorem follows.

Appendix F
The Proof of Theorem 2

The lower bound in Theorem 1 can be reached when all EVs but one in $\mathcal{Q}_-^{\text{pSTFQ}}(t)$ depart simultaneously. Even though the FM^1 bound of the left EV is small, since no other EVs in $\mathcal{Q}_-^{\text{pSTFQ}}(t)$ can be selected, the left EV will still be selected and turned on, which further decreases the FM^1 bound of the EV. However, the probability of many EVs departing within a short time interval is extremely low. In most cases, EVs may depart randomly or their departure times spread over a long time interval. There are always some EVs that can be selected before the FM^1 bounds of one EV becomes much smaller than others unless the number of EVs being charged is small. Let us study the lower bound in most cases.

Denote $-FM_i^1(0,t)$ as $\Delta_i(t)$. Let t_n be the time that $\Delta_{i_n}(t)$ of queue i_n reaches the maximum and assume $\Delta_{i_n}(t_n) > KL_{i_n}^{\max}$. From Lemma 2, we know t_n is a multiple of charging periods. Because $\Delta_{i_n}(t_n)$ is the maximum, EV i must be served at $t_n - T$; otherwise, $\Delta_{i_n}(t_n - T) > \Delta_{i_n}(t_n)$, contradicting the assumption that $\Delta_{i_n}(t_n)$ is the maximum. Nonetheless, $\Delta_{i_n}(t_n)$ is not the only maxima and it is possible that $\Delta_{i_n}(t_n - T) = \Delta_{i_n}(t_n)$; for example, when all queues in $\mathcal{Q}^{\text{pSTFQ}}(t_n - T)$ are served, we have $W_i^{\text{sMGPS}}(t_n - T, t_n) = W_i^{\text{pSTFQ}}(t_n - T, t_n) = \int_{t_n - T}^{t_n} \hat{P}_i(t)dt$ and thus $\Delta_{i_n}(t_n - T) = \Delta_{i_n}(t_n)$. In this case, $\Delta_{i_n}(t_n - T)$ is also the maximum. Let t_{n-1} be the largest time, a multiple of charging periods, before time t_n such that $\Delta_{i_n}(t_{n-1}) < \Delta_{i_n}(t_n)$. Then, we have $W_{i_n}^{\text{pSTFQ}}(t_{n-1}, t_n) - W_{i_n}^{\text{sMGPS}}(t_{n-1}, t_n) \leq L_{i_n}^{\max}$. At time t_{n-1}, we have

$$\begin{aligned}\Delta_{i_n}(t_{n-1}) &= \Delta_{i_n}(t_n) - (\Delta_{i_n}(t_n) - \Delta(t_{n-1})) \\ &> KL_{i_n}^{\max} - \left(W_{i_n}^{\text{pSTFQ}}(t_{n-1}, t_n) - W_{i_n}^{\text{sMGPS}}(t_{n-1}, t_n)\right) \\ &\geq (K-1)L_{i_n}^{\max}.\end{aligned}$$

At time t_{n-1}, not all queues in $\mathcal{Q}^{\text{pSTFQ}}(t_{n-1})$ are served. There exists some queue i_{n-1} with $\Delta_{i_{n-1}}(t_{n-1}) > (K-1)L_{i_n}^{\max}$. Similarly, for queue i_n, there exists some time $t_{n-2} < t_{n-1}$, where $\Delta_{i_n}(t_{n-2}) > (K-2)L_{i_n}^{\max}$ and at least one queue i_{n-2} such that $\Delta_{i_{n-2}}(t_{n-2}) > (K-2)L_{i_n}^{\max}$; for queue i_{n-1}, there exists some time $t_{n-3} < t_{n-1}$, where $\Delta_{i_n}(t_{n-3}) > (K-2)L_{i_n}^{\max}$ and at least one queue i_{n-3} such that $\Delta_{i_{n-3}}(t_{n-3}) > (K-2)L_{i_n}^{\max}$. Based on Eq. (9), when the FM^1 of an EV increases, the FM^1 of one another EV must decrease to balance the equation. Since the power ratings of EVs are independent of each other and the charging rates of EVs vary with the SOCs of their batteries, for each EV, the charging rates of other EVs can be considered as random. Thus, when the FM^1 of an EV increases, it is expected to be an EV whose FM^1 decreases. In other words, queues i_{n-1} and i_{n-2} mentioned above are expected to be different queues. Following this rule, we have

$$\Delta_{i_{n-j}}(t_{n-j}) > (K - (\lfloor \log_2 j \rfloor + 1))L_{i_n}^{\max}, \forall j = 1, \ldots, r,$$

where $r = \left|\mathcal{Q}_-^{\text{pSTFQ}}(t_{n-j})\right|$ and $\lfloor x \rfloor$ is the largest integer less than or equal to x. Since $\Delta_i < 0$ for all EVs in $\mathcal{Q}_-^{\text{pSTFQ}}(t)$, we have $K - \lfloor \log_2 r \rfloor + 1 \leq 0$. Hence, $K \leq \lfloor \log_2 |\mathcal{N}| \rfloor + 1$.

Appendix G
Proof of Theorem 3

Under pMGPS, whenever there is a load change in the sub-grid, a new round of rate allocation is conducted for EVs. Under sMGPS, rate allocation happens only in the beginning of each charging period. Nonetheless, since the charging period is short, i.e., about several minutes, the load variations during a charging period can be considered as random. It is reasonable to assume that the expectation of the load variations is zero. We prove the lemma in two cases.

Case 1: The sub-grid capacity is constrained by transformer capacity. Let $P_0(t)$ be the power flow through the transformer before any EV is charged, $\Delta P_{0i}(t)$ and $\Delta P'_{0i}(t)$ be the power flow change through the transformer after EV i is charged under pMGPS and sMGPS, $P_i(t)$ and $P'_i(t)$ be the charging rates of EV i under pMGPS and sMGPS. Based on [14], linearity is an adequate approximation to the voltage and thermal sensitivities of users. Under pMGPS, we have $\Delta P_{0i}(t) = \alpha_i P_i(t)$. Let C be the transformer capacity. Then, we have

$$C - P_0(t) = \sum_{i\in\mathcal{N}(t)} \Delta P_{0i} = \sum_{i\in\mathcal{N}(t)} \alpha_i P_i(t) = x(t) \sum_{i\in\mathcal{N}(t)} \alpha_i w_i, \tag{11}$$

where $x(t) = \frac{P_i(t)}{w_i}$ for all $i \in \mathcal{N}(t)$.

Since sMGPS is a virtual system, it impacts the power flow change by controlling the packetized scheme, pSTFQ. During any time interval $[t_1, t_2]$, the total energy charged for EV i under sMGPS is $W_i^{\text{sMGPS}}(t_1, t_2) = \int_{t_1}^{t_2} P_i(t)dt$. Then, the average possibility of turning on EV i under pSTFQ at the beginning of charging periods is

$$q_i = \frac{W_i^{\text{sMGPS}}(t_1, t_2)}{\int_{t_1}^{t_2} \hat{P}_i(t)dt} = \frac{\int_{t_1}^{t_2} x(t)dt}{\int_{t_1}^{t_2} y(t)dt}, \tag{12}$$

where $y(t) = \frac{\hat{P}_i(t)}{w_i}$. Thus, if EVs i and j are active for any time interval $[t_1, t_2]$, $q_i = q_j$. When t_2 approaches t_1, q_i becomes $q_i(t) = \frac{x(t)}{y(t)} = \frac{1}{|\mathcal{N}(t)|}$.

Under pSTFQ, we have

$$\begin{aligned} C - P_0(t) - \Delta P'(t) &= \sum_{i\in\mathcal{N}(t)} \mathbb{E}[\Delta P'_{0i}] \\ &= \sum_{i\in\mathcal{N}(t)} \alpha_i q_i(t) \hat{P}_i(t) = x'(t) \sum_{i\in\mathcal{N}(t)} \alpha_i w_i, \end{aligned}$$

where $x'(t) = \frac{P'_i(t)}{w_i}$ for all $i \in \mathcal{N}(t)$ and $\Delta P'(t)$ is the unutilized power due to the discrete charging rates of EVs. The unutilized power is less than the charging rate of the lucky EV. For any $i \in \mathcal{N}(t)$, since $\alpha_i \geq 1$,

$$\begin{aligned} \mathbb{E}[FM_i^2(t_1, t_2)] &= \int_{t_1}^{t_2} w_i(x(t) - x'(t))dt \\ &= \int_{t_1}^{t_2} \frac{w_i \Delta P'(t)}{\sum_{i\in\mathcal{N}(t)} \alpha_i w_i} dt \leq \int_{t_1}^{t_2} \frac{w_i \Delta P'(t)}{\sum_{i\in\mathcal{N}(t)} w_i} dt. \end{aligned} \tag{13}$$

Case 2: The sub-grid capacity is constrained by voltage. Let i_t and i'_t be the users with the minimum voltage in the sub-grid under pMGPS and sMGPS respectively at time t. Let $V_i(t)$ be the voltage of user i before any EVs are charged. Let $\Delta V_i(t)$ and $\Delta V'_i(t)$ be the voltage changes of user i after EVs are charged under pMGPS and sMGPS respectively. Similar to Eq. (11), we have

$$V_{i_t}(t) - V_{\min} = \Delta V_{i_t}(t) = \sum_{j\in\mathcal{N}(t)} \beta_{ij} P_j(t) = x(t) \sum_{j\in\mathcal{N}(t)} \beta_{ij} w_j,$$

and

$$\begin{aligned} V_{i'_t}(t) - V'_{\min} &= \mathbb{E}[\Delta V'_{i'_t}(t)] \\ &= \sum_{j\in\mathcal{N}(t)} \beta_{ij} q_i(t) \hat{P}_j(t) = x'(t) \sum_{i\in\mathcal{N}(t)} \beta_{ij} w_j, \end{aligned}$$

where $V'_{\min} = V_{\min} + \Delta V_{\min} + \Delta V(P'(t))$ and $\Delta V(P'(t))$ is the unutilized voltage due to the discrete charging rates.

Under pMGPS, all EVs are allocated with a rate, while under pSTFQ, only part of EVs are charged with full rates and the rest are not charged. The charging rates of EVs under pMGPS are distributed more evenly than those under pSTFQ. For users with the minimum voltage before EVs are charged, they have a higher probability of still having the minimum voltage under pMGPS than that under pSTFQ. In other words, the expected voltage change of EV i_t is smaller or equal than the expected voltage of EV i'_t, i.e., $\mathbb{E}[V_{i_t}(t) - V_{\min}] \leq \mathbb{E}[V_{i'_t}(t) - V_{\min}]$. Thus, we have

$$\begin{aligned} \mathbb{E}[FM_i^2(t_1, t_2)] &= \int_{t_1}^{t_2} w_i(x(t) - x'(t))dt \\ &\leq \int_{t_1}^{t_2} \frac{w_i(\Delta V_{\min} + \Delta V(P'(t)))}{\sum_{i\in\mathcal{N}(t)} \beta_{ij} w_j} dt. \end{aligned}$$

Since $\Delta V(P'(t))$ is the unutilized voltage due to the discrete charging rates, the fairness gap caused by $\Delta V(P'(t))$ is the same as Eq. (13). Hence,

$$\mathbb{E}[FM_i^2(t_1, t_2)] \leq \int_{t_1}^{t_2} \left(\frac{w_i \Delta V_{\min}}{\sum_{i\in\mathcal{N}(t)} \beta_{ij} w_i} + \frac{w_i \Delta P'(t)}{\sum_{i\in\mathcal{N}(t)} w_i} \right) dt.$$

REFERENCES

[1] P. J. Balducci, "Plug-in hybrid electric vehicle market penetration scenarios," Pacific Northwest Nat. Laboratory, Richland, WA, USA, Tech. Note JRC 58748, 2008.

[2] *IEEE Guide for Loading Mineral-Oil-Immersed Power Transformers Up to and Including 100 mVA with 55 Degrees C or 65 Degrees C Average Winding Rise*, ANSI/IEEE Standard C57.92–1981.

[3] T. A. Short, *Electric Power Distribution Handbook*. Boca Raton, FL, USA: CRC Press, 2004.

[4] K. Clement-Nyns , E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 371–380, Feb. 2010.

[5] O. Sundstrom and C. Binding, "Flexible charging optimization for electric vehicles considering distribution grid constraints," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 26–37, Mar. 2012.

[6] E. Sortomme, M. M. Hindi, S. D. J. MacPherson, and S. S. Venkata, "Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 198–205, Mar. 2011.

[7] Q. Li, T. Cui, R. Negi, F. Franchetti, and M. D. llić, "On-line decentralized charging of plug-in electric vehicles in power systems," [Online]. Available: http://arxiv.org/abs/1106.5063

[8] W. Tang, S. Bi, and Y. J. Zhang, "Online coordinated charging decision algorithm for electric vehicles without future information," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2810–2824, Nov. 2014.

[9] L. Gan, U. Topcu, and S. Low, "Optimal decentralized protocols for electric vehicle charging," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 940–951, May 2013.

[10] N. Chen, C. W. Tan, and T. Q. S. Quek, "Electric vehicle charging in smart grid: Optimality and valley-filling algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1073–1083, Dec. 2014.

[11] S. Chen, L. Tong, and T. He, "Optimal deadline scheduling with commitment," in *Proc. 49th Annu. Allerton Conf. Commun. Control Comput.*, 2011, pp. 111–118.

[12] S. Stdli, E. Crisostomi, R. Middleton, and R. Shorten, ," A flexible distributed framework for realising electric and plug-in hybrid vehicle charging policies," *Int. J. Control*, vol. 85, no. 5, pp. 1130–1145, 2012.

[13] O. Ardakanian, C. Rosenberg, and S. Keshav, "Distributed control of electric vehicle charging," in *Proc. 4th Int. Conf. Future Energy Syst.*, 2013, pp. 101–112.

[14] P. Richardson, D. Flynn, and A. Keane, "Optimal charging of electric vehicles in low-voltage distribution systems," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 268–279, Feb. 2012.

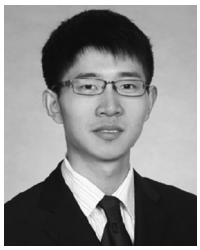
[15] [Online]. Available: http://www.teslamotors.com/supercharger

[16] A. Demers, S. Keshav, and S. Shenker, "Analysis and simultaton of a fair queueing algorithm," in *Proc. Symp. Commun. Archit. Protocols*, 1989, pp. 1–12.

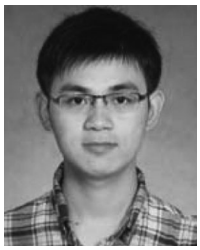
[17] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: The single-node case," *IEEE Trans. Netw.*, vol. 1, no. 3, pp. 344–357, Jun. 1993.

[18] J. M. Blanquer and B. Özden, "Fair queueing for aggregated multiple links," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, 2001, pp. 189–197.
[19] T. Sasaki, Y. Ukyo, and P. Novák, "Memory effect in a lithium-ion battery," *Nature Mater.*, vol. 12, no. 6, pp. 569–575, 2013.
[20] K. Yap, N. McKeown, and S. Katti, "Multi-server generalized processor sharing," in *Proc. 24th Int. Teletraffic Congr.*, 2012, Art. no. 29.
[21] K. Yap, T. Huang, Y. Yiakoumis, S. Chinchali, N. Mckeown, and S. Katti, "Scheduling packets over multiple interfaces while respecting user preferences," in *Proc. 9th ACM Conf. Emerging Netw. Experiments Technol.*, 2013, pp. 109–120.
[22] A. Ghodsi, V. Sekar, M. Zaharia, I. Stoica, "Multi-resource fair queueing for packet processing," in *Proc. ACM SIGCOMM Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, 2012, pp. 1–12.
[23] P. Richardson, D. Flynn, and A. Keane, "Impact assessment of varying penetrations of electric vehicles on low voltage distribution systems," presented at the *IEEE Power Energy Soc. General Meeting*, Minneapolis, MN, USA, Jul. 2010.
[24] [Online]. Available: http://www.pserc.cornell.edu/matpower/

Xudong Wang received the PhD degree in electrical and computer engineering from the Georgia Institute of Technology, in August 2003, he has been working as a senior research engineer, senior network architect, and R&D manager in several companies. He is currently with the UM-SJTU Joint Institute, Shanghai Jiao Tong University. He is a distinguished professor (Shanghai Oriental Scholar) and is the director of the Wireless and NetworkinG (WANG) Lab. He is also an affiliate faculty member in the Electrical Engineering Department, University of Washington. He has been actively involved in R&D, technology transfer, and commercialization of various wireless networking technologies. His research interests include wireless communication networks, smart grid, and cyber physical systems. He holds a number of patents on wireless networking technologies and most of his inventions have been successfully transferred to products. He is an editor of the *IEEE Transactions on Mobile Computing*, the *IEEE Transactions on Vehicular Technology*, and the Elsevier *Ad Hoc Networks*. He was also a guest editor for several journals. He was the demo co-chair of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM MOBIHOC 2006), a technical program co-chair of Wireless Internet Conference (WICON) 2007, and a general co-chair of WICON 2008. He is a senior member of the IEEE and was a voting member of IEEE 802.11 and 802.15 Standard Committees.



Yibo Pi received the BS and MS degrees in electrical and computer engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2012 and 2015, respectively. He is currently working toward the PhD degree at the University of Michigan. His current research interests include smart grid communications and wireless networks.



Aimin Tang received the BS degree in electrical and computer engineering from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2013. He is currently working toward the PhD degree in the Wireless and Networking Lab, Shanghai Jiao Tong University. His current research interests include full duplex communications, D2D communications, smart grid communications, and software defined wireless networks.