A Mean Field Game Theoretic Approach to Electric Vehicles Charging

ZIMING ZHU[1], (Member, IEEE), SANGARAPILLAI LAMBOTHARAN[2], (Senior Member, IEEE), WOON HAU CHIN[1], (Senior Member, IEEE), AND ZHONG FAN[1]

[1]Telecommunications Research Laboratory, Toshiba Research Europe Ltd., Bristol BS1 4ND, U.K.
[2]Advanced Signal Processing Group, Loughborough University, Leicestershire LE11 3TU, U.K.

Corresponding author: Z. Zhu (ziming.zhu@toshiba-trel.com)

ABSTRACT Electric vehicles (EVs) provide environmentally friendly transport and they are considered to be an important component of distributed and mobile electric energy storage and supply system. It is possible that EVs can be used to store and transport energy from one geographical area to another as a supportive energy supply. Electricity consumption management should consider carefully the inclusion of EVs. One critical challenge in the consumption management for EVs is the optimization of battery charging. This paper provides a dynamic game theoretic optimization framework to formulate the optimal charging problem. The optimization considers a charging scenario where a large number of EVs charge simultaneously during a flexible period of time. Based on stochastic mean field game theory, the optimization will provide an optimal charging strategy for the EVs to proactively control their charging speed in order to minimize the cost of charging. Numerical results are presented to demonstrate the performance of the proposed framework.

INDEX TERMS EV consumption management, optimal charging, stochastic optimisation, mean field game.

I. INTRODUCTION

The future electric power generation and supply system which is recognised as smart grids, is expected to bring significant benefits to energy generation and dispatch. The direction of power flow will no longer be just downhill from the bulk power plants to consumers, but it can start from any energy generation sources and end up anywhere on the grid. Electric vehicle (EV) is expected to be one of the main components of distributed energy consumption, storage and supply system in smart grids. EVs can serve as a distributed and mobile energy source in the electricity market [1], [2]. Facilitated by the advanced information and communication technologies (ICT), EVs can be optimally scheduled and dispatched to meet the dynamic demand of energy and to respond swiftly to emergency situations [3]. The storage and transportation of energy from one geographical area to another as supportive supply enhances the overall flexibility of the grid [4]. As EVs will eventually be employed at household level, as alternative to traditional petrol cars, it is necessary to include them into home electricity demand management and consumption optimisation [5].

EV is a major electricity consumer and draws a significant amount of power in order to retain sufficient battery capacity.

For the grid operators, such high loads attached to the grid will have to be managed carefully [6], [7]. A scheduling method proposed in [8] minimises the system operation costs as well as the difference between the minimum and maximum system demand. To cope with the potential new load with minimal additional infrastructure, The authors of [9] proposed a load shaping tool to improve the usage of distribution transformers. The authors of [10] proposed an EV classification scheme for a renewable charging station to reduce the effect of intermittency of electricity supply and the cost of energy trading. The EV owners should consider the best charging times and charging rates (speed) to reduce the cost of energy consumption. Both centralised and decentralised approaches are proposed for optimal charging. In the centralised approach, a scheduling agent is responsible for handling all EVs connected to the power grid and to optimise the charging schedules globally. These techniques are able to provide globally optimal solutions. However, undesired computational complexity and delay are usually seen. Optimal charging can also be achieved via decentralised algorithms where each EV manages its own charging according to the operational condition of the grid as well as economic incentives [11], [12]. Optimal charging

becomes one of the critical challenges in the utilisation of EVs, as evidenced by the emerging work in the literature. For example, a charging strategy with a genetic algorithm (GA) which obtains the stochastic features in order to reduce the power fluctuation level caused by EV charging was presented in [13]. Paper [14] presented a model which facilitates a cooperative participation of EVs in residential buildings and a parking lot. Heuristic methods for optimal charging considering the acceptable charging power at different state-of-charges and in response to variable pricing policies in a regulated market were studied in [15]. Vehicle-to-grid (V2G), the provision of energy from the EV to the grid as an ancillary service, has the potential to offer financial benefits to both EV owners and the power system. In [16], a V2G algorithm was developed to provide additional system flexibility and peak load shaving to the utility and low costs of EV charging to the customer. An algorithm was developed in [17] for use by a V2G aggregator to bid into energy markets. Considering the energy flow between the EVs and the grid is bidirectional, a load shifting technique by optimally scheduling the charging and discharging of EVs in a decentralised fashion was proposed in [18]. The authors of [19] developed a global EV power management scheme to minimise the cost of charging and discharging over the day. A decentralised scheme was further developed in order to minimise the cost of the EVs in the local groups. The local scheduling was claimed scalable to a large EV population and also resilient to the dynamic EV arrivals.

Game theory is a powerful tool for understanding and modelling mathematically the interaction of various rational decision makers. Applying game theoretic formulation to various optimisation problems has attracted a lot of interests in recent research on communication networks and signal processing [20]. The work in [21] proposed a strategic game where each greedy base station in a multiple input and multiple output (MIMO) system determines its optimal downlink beamformer but without any coordination with other base stations. Compared to a fully coordinated design where the optimal beamformers are jointly designed, the scheme provided benefits in terms of lower system complexity and overheads. Game theory is very suitable for analysing the interaction of consumers and utility operators in order to achieve efficient distributed demand management [22], [23]. More detailed discussions on demand management and game theory can be found in [24]. There have been recent interests in the application of game theory in smart grids as well as EV charging optimisation [25], [26]. Paper [27] proposed a decentralised charging control method based on the Nash Certainty Equivalence Principle that considers network impacts. A Stackelberg game was proposed in [28] to model the energy exchange between the grid and the EVs. In this leader-follower game, the smart grid decides its price to optimise its revenue while ensuring the participation of EVs. On the other hand, the EVs optimise their charging strategies. The Stackelberg equilibrium is obtained via the proposed decentralised algorithm. In particular, a mean field game theoretic framework was

developed in [29] to minimise the consumption cost for the EVs within a predefined period of time. The energy consumption behaviour of a large number (tends to infinity) of EVs, including charging power from the grid and releasing power to the grid, was modelled as a stochastic optimisation where individual player (EV) chooses its optimal strategy (the amount of energy charging/discharging at any particular time) according to the statistical behaviour of the total group of players. The mean field game is a novel differential game theoretic modelling mechanism which was first proposed in [30]. It provides a powerful mathematical modelling based on the formulation of two coupled backward-forward partial differential equations (PDE) for problems with a large number of indistinguishable players. The optimal game solution, which is claimed as the Nash-Mean Field Equilibrium (Nash-MFE) is derived by solving the coupled equations. The applications of mean field game theory are also proposed for security enhancements and power control in mobile networks [31], [32].

In this paper, a mean field game theoretic framework is proposed for a scenario where individual EVs manage their charging at an aggregated charging station. The charging station has the capability of charging a large number of EVs during a period of time. We consider that, in the sense of demand management, the charging station has scheduled operation times in order to prevent unexpected peak load on the grid. Acknowledging this, EVs arrive at the station on time and charge their battery within the defined charging period. We assume that the charging station allows a minor delay in operation time. The EV is able to continue charging in excess of the scheduled operation time, subject to a penalty cost. The novelty of the proposed framework lies in the consideration of a degree of flexibility in the length of charging process. Compared with determined scheduling as in [29], the proposed setting provides more customer comfort. The optimisation enables the EVs to dynamically control the charging process and to be able to finish at an appropriate time so that the total cost of charging is minimised.

This paper is organised as follows. The game theoretic optimisation framework is formulated in section II. Section III provides detailed discussion on the issue of Nash-Mean Field Equilibrium and the methodology of obtaining the optimal solution of the game. Numerical simulations are presented in Section IV to demonstrate the performance of the proposed framework. Section V draws the conclusion.

II. THE GAME THEORETIC OPTIMISATION FRAMEWORK

The system consists of an EV charging station where a large fleet can be charged simultaneously. The EVs are aggregated at the station and recharge their batteries in a predefined period of time. The charging station can tolerate certain delay in the scheduled finishing time. The actual time of ending the charging service will depend on the dynamics of the charging EVs, i.e., when a certain quorum of fully charged EVs is reached.

The station charges an EV according to its power consumption, which can be represented by the EV's charging rates (speed) over the charging time. Detailed description of cost of charging is provided in the subsequent section. Given the pricing information, individual EV is encouraged to optimise and control its charging in order to minimise the cost of charging. Due to the non-deterministic charging time, it is necessary to formulate the charging optimisation framework as a dynamic control process, based on the knowledge of the current charging status of all EVs in the station. The proposed optimisation framework aims to provide an optimal control strategy that minimises the cost of charging for every individual EV, i.e., a profile that defines the dynamics of the charging rates for the whole charging duration. The system is illustrated in Figure 1.
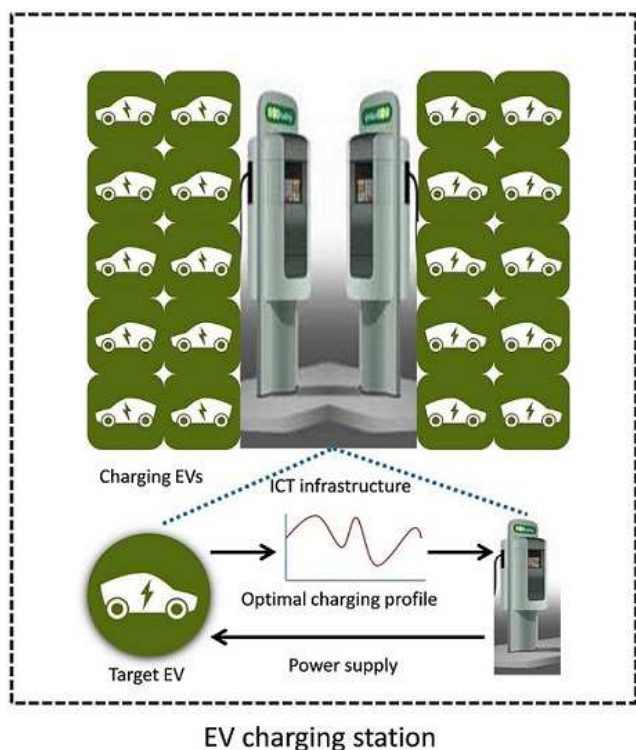

FIGURE 1. System model: aggregated EV charging at a station.

A. OPTIMISATION COSTS

The minimisation of cost of charging is the objective of the optimisation for the EVs. The cost consists of the energy consumption cost during charging and the endpoint costs that are related to the finishing time of charging. Consider a finite set of aggregated charging EVs $\mathbb{K} = \{1, \ldots, K\}$. The charging station's pricing policy for any EV k is defined as a continuous function $u_t^{(k)}$ of the charging speed $a_t^{(k)}$, i.e., its energy consumption at time t,

$$u_t^{(k)} = \frac{1}{2}(a_t^{(k)})^2. \tag{1}$$

The pricing plan can be viewed as a continuous function approximating the stepwise (multi-level) pricing models

adopted in various research on electricity markets [33]–[35]. The simplified quadratic representation is also widely used to formulate both the production cost function and the revenue function in economics. This pricing policy provides incentive for the EVs to maintain charging at low power in order to achieve low cost for the whole charging duration. The charging station can also benefit in terms of lower accumulated load generated on the grid, especially when massive EVs are charging at the same time.

The issue of charging time is now described. Denote the station's scheduled operation period as $[0, \hat{t}], \hat{t} > 0$. All connected EVs are expected to start their charging at time 0. They should be willing to maintain lowest possible charging rates until their batteries are fully charged at time $\hat{t}$, when the station is scheduled to terminate its charging service. However in reality the actual charging duration of any particular EV can vary from the expected time. This variation is mainly due to charging efficiency/loss, degree of degradation of individual battery. Denote an EV's actual finishing time as $\tau^{(k)}$. As mentioned above, the charging station will tolerate a modest delay in terms of the finishing time, up to a maximum of t_{max}. The actual charging finishing time, denoted by T, between $[\hat{t}, t_{max}]$ will depend on the dynamics $\{\tau^{(k)}, \forall k\}$ of all EVs.

In the following, several endpoint costs are defined as functions of the charging finishing times $\hat{t}, \tau$ and T. Firstly, a punctuality cost is set. It can be viewed as a price paid for lateness, payable to the charging station, for the EV k:

$$c_1^{(k)}(\hat{t}, \tau^{(k)}) = f_1([\tau^{(k)} - \hat{t}]_+). \tag{2}$$

The charging station can issue such lateness penalty to regulate the punctuality behaviour of charging EVs. The selection of this cost function $f_1(\cdot)$ will have influence on the result of charging optimisation directly, as discussed in Section IV.

Secondly, a cost of the lateness is defined in terms of the actual finishing time of charging. This reflects the loss of incomplete battery recharge because the charging station will have to stop power supply after this time:

$$c_2^{(k)}(T, \tau^{(k)}) = f_2([\tau^{(k)} - T]_+). \tag{3}$$

This represents the cost of inefficiency for the charging EV. Although not actual financial cost, it is included in the optimisation. Besides, some EVs may opt for very fast charging as their priority is the charging time. For those EVs, the utility function should be modelled differently. They are not considered to participate in the same game. Therefore this type of charging is not covered in this paper.

Finally, $C_T^{(k)}(\hat{t}, \tau^{(k)}, T) = c_1^{(k)}(\hat{t}, \tau^{(k)}) + c_2^{(k)}(T, \tau^{(k)})$ is used to represent the cost at the finishing time of the charging. Individual EV would want to minimise this cost along with the charging expense during the whole charging duration. Assume that all these cost functions are continuous and twice differentiable. The total cost function $J^{(k)} : a^{(k)} \mapsto \mathbb{R}$ of the

optimal charging is therefore

$$J^{(k)} = \int_0^{\tau^{(k)}} u(a_t^{(k)})dt + C_T^{(k)}(\hat{t}, \tau^{(k)}, T). \qquad (4)$$

B. DYNAMIC EV CHARGING PROCESS

Let us suppose that an EV's charging is represented by its battery capacity $X^{(k)} \in [0, 1]$ moving from an initial state $X_0^{(k)} > 0$ (battery capacity when charging starts) towards the fully charged point of 1. This movement is described using a dynamic process, written as a stochastic differential equation

$$dX_t^{(k)} = \eta^{(k)} a_t^{(k)} dt + \sigma_t dW_t^{(k)} + dN_t^{(k)}, \qquad (5)$$

where the charging rate $a_t^{(k)}$ is a controlled drift at time t in return for a cost as defined in (1) and $\eta^{(k)}$ represents the measurable charger efficiency for the EV, which is assumed to be 1 for simplicity. $W_t^{(k)}$ is an independent Brownian motion (Wiener process) with a diffusion coefficient σ_t. It's differentiation should follow the rules of Itō calculus [36]. The choice of $W_t^{(k)}$ represents the adjustment (uncertainty in power loss) added to the charging which indicates that the charging process is independent among the EVs at different times, due to different battery characteristics and individual EV's minor operational consumption during the charging time. The term $N_t^{(k)}$ is a reflective noise which ensures that the value of $X^{(k)}$ remains in $(0, 1)$. The reader is referred to [37] for more details on Brownian motion.

At any particular point in time during the charging process, the EV will be able to obtain the information of the charging status, i.e., the current battery capacity $\{X_t^{(1)}, \ldots, X_t^{(K)}\}$ and the charging rates $\{a_t^{(1)}, \ldots, a_t^{(K)}\}$ of all charging EVs via communications through the ICT infrastructure. An estimation of actual finishing time T is obtained based on this gathered information. Mechanisms for the estimation of T can vary depending on the particular algorithms. This paper considers the mean field game theoretic method. Due to the dynamic nature of the optimisation, such information must be exchanged continuously and in real-time during the charging period. However for each EV, the amount of data overhead required for exchanging information at any time is limited. Considering that the EVs are aggregated in the station, data communications take place at a short distance through wireless sensor networks embedded in the EVs.

Having included the information of estimated T into the cost function, the EV is able to optimise its own charging process. The optimisation is described as a stochastic control:

$$\min_{a_t^{(k)}} E\left[\frac{1}{2}\int_0^\tau (a_t^{(k)})^2 dt + C_T^{(k)}(\hat{t}, \tau^{(k)}, T)\right], \qquad (6)$$

subject to the dynamic $dX^{(k)} = a^{(k)}dt + \sigma dW^{(k)} + dN^{(k)}$ with an initial state $X_0^{(k)}$ and the expectation is over $W^{(k)}$. The optimisation will aim to find an optimal law $\gamma^*(t, X^{(k)*})$ which defines the optimal charging strategy of the control trajectory $a^{(k)*}$ and hence the movement of $X^{(k)*}$ for the particular EV. Note that in practice, the charging rate of an EV is usually valued in a range of $[a_{min}, a_{max}]$. This should be included as an additional constraint in the optimisation problem.

C. K-PERSON GAME THEORETIC FORMULATION

Based on the above formulation for a single EV, the optimisation of the total K EVs at the charging station is now formulated as a game theoretic framework.

One classical formulation is the K-person dynamic differential game where every EV is treated as an individual player, hence $\mathbb{K}$ represents set of *players*. They are assumed to be rational meaning that they will play the best *strategies*, i.e., at time t, player k optimises its charging control $a_t^{(k)}$ based on the understanding of the game situation in order to maximise its own utility. The situation of the game at time t is determined by the charging status of every individual EV in the station. Here $\Omega_t^{(k)} = (X_t^{(1)}, \ldots, X_t^{(K)}, a_t^{(1)}, \ldots, a_t^{(K)})$ denotes the set of information available to the player k at time t. It is assumed that players are memoryless as they do not have this status information of previous time instants. The level of satisfaction (utility) under certain game situation is represented by a *payoff*. In this particular optimal charging game, a player's payoff can apparently be measured by its cost of charging. Therefore, the objective of the player is to determine a dynamic trajectory of charging rate $a_t^{(k)}$ that maximises its payoff by conducting the optimisation as defined in (6).

Define a mapping $\mathfrak{B}_t^{(k)} : \Omega_t^{(k)} \mapsto \mathbb{S}_t^{(k)}$, to represent the choice of strategy for player k at time t, with $\mathbb{S}_t^{(k)}$ the set of all possible controls $a_t^{(k)}$ for the player. In particular, $\mathfrak{B}_t^{(k)}$ yields a best response control that maximises the payoff. The optimal charging action of player k at time t is therefore a (own-state) feedback strategy determined by

$$a_t^{(k)} = \mathfrak{B}^{(k)}(\Omega_t^{(k)}), \quad 0 \leq t \leq T. \qquad (7)$$

Referring to the game theoretic interpretations, $\mathbf{a}_t^{(-k)} \in \mathbb{S}_t^{(-k)}$, where $\mathbb{S}_t^{(-k)} = \prod_{i \in \mathbb{K}, i \neq k} \mathbb{S}_t^{(i)}$, denotes the joint strategy choices of all players other than k at time t. For player k, the choice of strategy $a_t^{(k)}$ is a best response to the current game status and the strategies chosen by all the players $(a_t^{(k)}, \mathbf{a}_t^{(-k)})$.

The formulation is completed by defining the strategy set over the total charging period for player k as $\mathbb{S}^{(k)} = \{\mathbb{S}_t^{(k)}, 0 \leq t \leq T\}$, and the overall strategy space for all players as $\mathbb{S} = \prod_{k \in \mathbb{K}} \mathbb{S}^{(k)}$. Having the above formulation, the game can be viewed as a dynamic optimisation process where every individual player (EV k) chooses best strategy (charging rate $a^{(k)}$) to maximise its payoff (minimise cost $J^{(k)}$), for the whole charging duration. The solution of the game is considered as a feedback Nash Equilibrium (NE), as defined below:

Definition 1: The feedback Nash Equilibrium of the K-person charging optimisation game is a joint strategy profile $\mathbf{a}^* = \{a^{(1)*}, a^{(2)*}, \ldots, a^{(K)*}\}$, $\mathbf{a}^* \in \mathbb{S}$, where $a^{(k)*} = \{a_t^{(k)*} = \mathfrak{B}^{(k)}(\Omega_t^{(k)*}), 0 \leq t \leq T\}$, and satisfies

for all $k \in \mathbb{K}$,

$$J^{(k)}(\mathbf{a}^*) \leq J^{(k)}(a^{(k)}, \mathbf{a}^{(-k)*}), \quad \forall a^{(k)} \in \mathbb{S}^{(k)}, a^{(k)} \neq a^{(k)*}. \tag{8}$$

This definition states that given the equilibrium strategy choices of other players $\mathbf{a}^{(-k)*}$, player k has no incentive to change its own strategy from $a^{(k)*}$ unilaterally. Nash Equilibrium is critical because, if exists, it guarantees a stable game situation where every player plays the best strategy responding to the strategic choices of all other players. For the particular charging game, obtaining the NE point is equivalent to achieving an optimal charging result for every EV in the system.

Analysing the NE in terms of showing its existence and uniqueness is never obvious [38]. Even if NE does exist, it may be difficult to develop convergence algorithms to exploit. This is more complex in the case of K-person games, where K can be considerably large. Various mechanisms for analysing multi-player differential games can be found in [39]. Nevertheless, it can be claimed that under the proposed game formulation, any change of any player at any time, i.e., changes in Ω_t, has impact on all players' payoffs. They will have to be acknowledged and respond accordingly. This results in significant computation complexity and increased ICT overhead. In order to resolve these potential issues, it is necessary to modify the formulation and propose the following mean field game approach.

D. MEAN FIELD GAME REPRESENTATION

Mean field game theory is powerful in modelling and analysing games with numerous players. For the simultaneous charging scenario involving a large number of EVs, it is possible to formulate a statistical performance of the whole population to represent the mean field, and every player optimises its charging strategy accordingly.

In order to model the above discussed charging optimisation scenario as a mean field game, two additional assumptions in relation to the players are required. Firstly, the total number of players is very large so that they can be viewed as a continuum instead of individuals. In other words, we now consider the charging of infinite EVs. Having this assumption in place, we are able to analyse the charging status based on a statistical distribution of the population, without the need of detailed observation of individual EVs. This condition is justified later. The second assumption is that the players are indistinguishable. This implies that all EVs have similar type of batteries and charging control abilities (however still their initial battery states and the efficiency loss, etc., may vary). They are modeled mathematically identical.

Having made these assumptions, we are able to remove the notations k and X in the classical formulation and use a state variable $x_t \in [0, 1]$ to represent the battery capacity of the indistinguishable player at time $t \in [0, T]$. The movement of x_t is indicated using a differential equation

$$dx_t = a_t dt + \sigma_t dW_t + dN_t, \tag{9}$$

which is similar to the one in (5) however without the player index k. The choice of W_t represents the uncertainty of x_t at different times. Now the charging processes of all the EVs can be modelled using the same formula. Considering the independence of individual EVs and the uncertainty in different times, the result of their participation of the charging can be described using a statistical distribution of x and t, denoted by $m(t, x_t)$. The distribution is defined in the compact domain of $[0, T] \times [0, 1]$ and has a compact support. This means that we are able to analyse the behaviour of the overall charging process through investigating the dynamic trajectory of m. The charging of m moves from initial state $m(0, x_0) = m_0$ towards the completed charging state indicated by $m(\cdot, 1)$.

In this context, the time when the dynamic flow reaching 1, can be seen as $\tilde{t} \mapsto \partial_x m(\tilde{t}, x)|_{x=1}$. The cumulative distribution function (CDF) F of finishing times can be defined by

$$F(\tilde{t}) = \int_0^{\tilde{t}} \partial_x m(\tilde{t}, x)|_{x=1} dt. \tag{10}$$

The actual finishing time T of the charging can be defined by this information of the dynamics of EVs. For example, T is fixed by a quorum rule of θ, which means θ percent of the EVs have finished their charging,

$$T = \begin{cases} \hat{t}, & \text{if } F^{-1}(\theta) \leq \hat{t} \\ t_{max}, & \text{if } F(t_{max}) \leq \theta \\ F^{-1}(\theta), & \text{otherwise} \end{cases} \tag{11}$$

Under the above formulation, the charging optimisation game is played from the viewpoint of an 'average' player. The player minimises his cost of charging with the optimal control $a(t, x) \in \bar{\mathbb{S}}$ for the trajectory of the battery state x according to the statistical behaviour m which determines T. $\bar{\mathbb{S}}$ denotes the space of all controls for the state dynamics. The actual players (the individual EVs) will be argued into the optimal strategy a. Their repartition will lead to the optimal trajectory of distribution m, which in return feeds back to the optimisation of charging strategy.

The game can be considered as a coupled system of cost minimisation and the optimal behaviour of the statistical trajectory. Solving the optimisation system will naturally lead to the optimal solution of the game [40]. In this way, the optimisation of the cost of charging no longer requires information Ω_t of all individuals however it knows the status m_t. System complexity and communications overhead are therefore reduced.

III. ANALYSIS OF THE MEAN FIELD GAME

The solution of the mean field game theoretic optimal charging framework, known as the Nash-Mean Field Equilibrium (Nash-MFE) [30], is discussed. As the mean field game is transformed from a K-person game, the definition of feedback Nash-MFE is stated based on *Definition 1*, as follows.

Definition 2: The Nash-Mean Field Equilibrium in feedback strategies of the charging optimisation game is

a control $a^* \in \bar{\mathbb{S}}$, consistent with the distribution m^* of the charging dynamics for a given initial state of m_0, and satisfies

$$J(a^*, m^*) \leq J(a, m^*), \quad \forall a \in \bar{\mathbb{S}}, a \neq a^*. \quad (12)$$

Having followed the equilibrium strategy of a^*, individual players of the game (EVs) have no incentive to deviate from a^*. Hence the dynamics of battery states will be according to m^*. Therefore it is claimed that their optimal charging processes with an actual finishing time of T^* are determined.

A. THE COUPLED STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

Based on the work in [30], a mathematical scheme is formulated using coupled stochastic partial differential equations (SPDE) in order to obtain cost minimisation and the optimal behaviour of the statistical trajectory, and to generate the MFE of the game.

Firstly, consider the cost minimisation problem which has the objective as in (6), however without the index k as now only the mean field 'average' player is considered. At any particular time t, the player will obtain a T fixed by the observation of m_t, and the agent is looking for the optimal control a^* for the minimum cost-to-go. The cost-to-go value function $U(t', x) : [0, \tau] \times [0, 1] \mapsto \mathbb{R}$ has the following form:

$$U(t', x) = \min_{a_t, t \leq t' \leq \tau} E\left[\frac{1}{2}\int_{t'}^{\tau} a_t^2 dt + C_T(\hat{t}, \tau, T)\right], \quad (13)$$

subject to the dynamics of x as defined in (9).

The optimal solution of the cost minimisation is the value function U which satisfies the backward Hamilton-Jacobi-Bellman (HJB) equation:

$$\partial_t U + \min_a \left(\frac{1}{2}a^2 + a\partial_x U\right) + \frac{\sigma^2}{2}\partial_{xx}^2 U = 0. \quad (14)$$

Solving the minimisation part by using the optimal control term $a^* = -\partial_x U$, this equation is formulated as:

$$\partial_t U - \frac{1}{2}(\partial_x U)^2 + \frac{\sigma^2}{2}\partial_{xx}^2 U = 0, \quad (15)$$

with the boundary conditions $U(\tau, 1) = C_T(\hat{t}, \tau, T)$ corresponding to the endpoint cost when fully charged, and $U(t_{max}, x) = C_T(\hat{t}, t_{max}, t_{max})$ corresponding to the endpoint cost defined in terms of the maximum allowed charging delay time. Provided an optimal m_t^*, the HJB equation will determine the function U and hence indicate the optimal $a^*(t)$ of the player.

The optimal movement of m_t^*, for a given m_0, is determined by the following forward Fokker-Planck-Kolmogorov (FPK) equation,

$$\partial_t m + \partial_x(a^* m) - \frac{\sigma^2}{2}\partial_{xx}^2 m = 0, \quad (16)$$

with the compact boundary conditions of $m(\cdot, 0) = 0$ and $m(\cdot, 1) = 0$. It is observed in the FPK equation that a^* is exactly the optimal control strategy results from the HJB equation. Solving the two coupled SPDEs will determine the MFE, if exists. There is no general methods to solve such nonlinear systems. As they are inherently distributed, iterative learning algorithms have been proposed in order to obtain the solution with reasonable computational complexity [41], [42].

B. EXISTENCE AND UNIQUENESS OF THE MFE

The justification of the above mathematical scheme stems from proving the existence and uniqueness of a MFE solution. Similar to classical games, Brouwer fixed point theorem is used for establishing the equilibrium point from the best responses mapping. For the proposed optimal charging mean field game, the mapping is between the optimal control a^* and m^* consisting all players' controls. It is discovered that, one chooses best strategy a^* by solving the HJB equation corresponding to a given T. T is determined by the dynamics of flow m^* which is given by the FPK equation. Hence it is useful to investigate the time T coherent with the rational behaviours of the players. The MFE is eventually a matter of locating the fixed point of the mapping $T \mapsto T$.

Consider the following representation of the SPDE scheme:

$$T \mapsto C_T \mapsto U \mapsto -\partial_x U \mapsto m \mapsto \partial_x m(\tilde{t}, x)|_{x=1} \mapsto T, \quad (17)$$

It can be seen that the scheme is from $[\hat{t}, t_{max}]$ to $[\hat{t}, t_{max}]$ itself. In order to obtain a fixed point result for the mapping, it is needed only to show the scheme is continuous [43].

The first part of the scheme, $C_T(\hat{t}, \tau, T)$ is assumed to be a C^2 continuous function. Following the second part, it can be observed that function U is continuous in C_T. It is further stated that the HJB equation provides a solution of $U \in C^2$ with $-\partial_x U$ is Lipschitz continuous according to [40]. Also, the solution m of the FPK equation is C^1 continuous and $\partial_x m(\tilde{t}, x)|_{x=1} \in C^0$ admits a positive lower bound for any $T \in [\hat{t}, t_{max}]$. Now the final mapping of the scheme is considered, which is $\Gamma : \partial_x m(\tilde{t}, x)|_{x=1} \mapsto T$. Define γ_1 and γ_2 to represent the two different flows of dynamics reaching 1. They are both bounded by a common ϵ. Assuming $T_1 = \Gamma(\gamma_1)$ and $T_2 = \Gamma(\gamma_2)$, $\hat{t} \leq T_1 < T_2 \leq t_{max}$, it has

$$\int_0^{T_2} \gamma_2 \leq \int_0^{T_1} \gamma_1 \Rightarrow \int_{T_1}^{T_2} \gamma_2 \leq \int_0^{T_1} (\gamma_1 - \gamma_2). \quad (18)$$

The left term $\int_{T_1}^{T_2} \gamma_2$ is bounded by $\epsilon(T_2 - T_1)$ while the value of $\int_0^{T_1} (\gamma_1 - \gamma_2)$ is below $t_{max}(||\gamma_1 - \gamma_2||_\infty)$. Thus,

$$(T_2 - T_1) \leq \frac{t_{max}}{\epsilon}(||\gamma_1 - \gamma_2||_\infty), \quad (19)$$

which satisfies the Lipschitz condition. Therefore the mapping Γ is C^0 continuous. The overall scheme is a continuous mapping of $T \mapsto T$, which admits a fixed point T^* coherent with the behaviours a^* and m^*. Hence, the existence of a MFE solution for the charging game is established. However, in order to produce a unique MFE, the game theoretic formulation requires additional monotonicity conditions in relation

to the cost optimisation in the HJB equation [40]. It can be argued that these conditions are subject to individual game modelling, and they are not necessarily general premises in EV charging scenarios.

C. MEAN FIELD GAME VERSUS K-PERSON GAME

The formulation of mean field game introduces a generalisation approach by which the interaction among large populations can be analysed, based on the assumptions that players are treated indistinguishable and continuum. The settings of mean field players have advantages in the sense of increased computational efficiency [30]. By formulating players into a continuum, it enables the use of powerful differential calculus and statistics for analysing the optimal behaviours of the players. As they take actions based only on the statistical state of the total mass, information exchange in terms of their exact game play can be omitted. This reduces the system ICT overhead while enhancing privacy. Moreover, comparing to K-person game where players are sensitive to the changes of the others, changes of particular players in a mean field game has little impact on the performance of the total mass. Therefore the optimal strategy choice of every player can remain. This enhances the efficiency and the stability of the optimisation system. However, the mean field players result in less sophisticated than that of K-person games, because players are able to observe and respond to the exact moves of all others in a K-person game.

In addition, the mean field solutions can be considered as the limit approximation of K-person games as $K \to \infty$. It is claimed that a corrective term in the order of $1/K$ is sufficient to describe the precision of the approximation [40]. Thus, the efficient mean field game approach can be applied to a wider range of practical applications including those with limited dimension (small K), for example in oil production, and in the case of EV charging.

IV. NUMERICAL RESULTS AND PERFORMANCE EVALUATION

A. SYSTEM SET UP

Consider a charging station with the total ability to charge a fleet of 500 EVs. The scheduled time length of charging is $\hat{t} = 120$ minutes, with the allowed maximum extension to $t_{max} = 150$ minutes. A quorum rule of $\theta = 90\%$ is used to determine T. The station's pricing policy has been defined as in (1). Two terminal costs for all EVs will be determined by the following linear functions: $c_1(\hat{t}, \tau) = 3([\tau - \hat{t}]_+)$, and $c_2(T, \tau) = 4([\tau - T]_+)$. The battery capacity of the EVs, as well as the charging status parameter x, are represented by percentage values in between 0 and 100. Assume that each EV has a full battery capacity of $40KWh$. However they have been assigned with different initial charging states. For the simulation the initial battery capacity value is assumed randomly between 20 and 30 percent. The charging speed of the EVs is bounded between $a_{min} = 0.25$ percent per minute and $a_{max} = 1.5$ percent per minute. Considering the
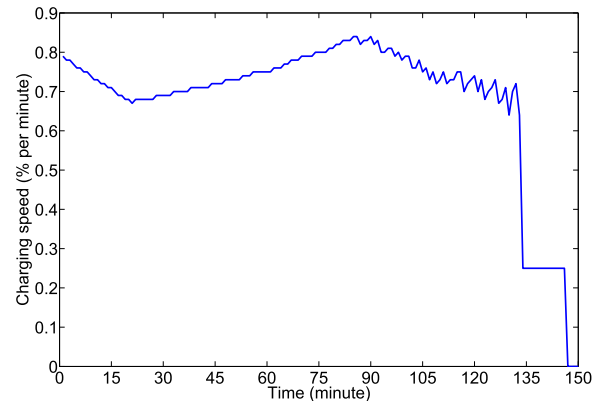
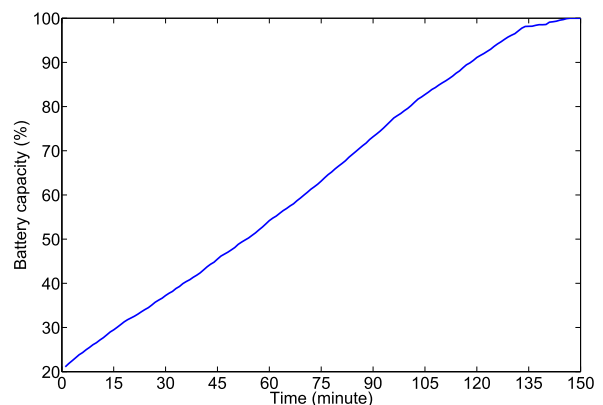**FIGURE 2.** Optimal charging speed for one EV in a mean field game scenario ($T^* = 147$).

FIGURE 3. Optimal trajectory of the charging battery capacity for one EV in a mean field game scenario ($T^* = 147$).

restraints of battery charging at different capacity levels, the maximum charging speed is halved when the capacity reaches 85 percent.

B. PERFORMANCE EVALUATION

The simulation demonstrates that all charging EVs are participating in the proposed mean field game to optimise their cost of charging. Figure 2 and Figure 3 depict the optimal charging results in terms of dynamic charging speed and the battery capacity trajectory for one randomly chosen EV from the 500 participants, respectively. As seen, the battery is charged in a dynamic however modicum fluctuated charging speed around 0.7 percent per minute. The battery capacity increases smoothly. When the capacity reaches 90 percent, the charging speed slows down significantly. The charging finishes with full capacity at the elapse of 147 minutes, which is exactly T^*, the finishing time determined by the MFE of the 500 EVs. In this way, the lateness cost at the endpoint has been successfully minimised, with a final cost of 78 which is only due to the penalty as defined in c_1. Therefore it can be claimed that the optimisation has efficiently made the full utilisation of the permitted time and obtained the optimal strategy for the EV. Since all EVs optimise the charging
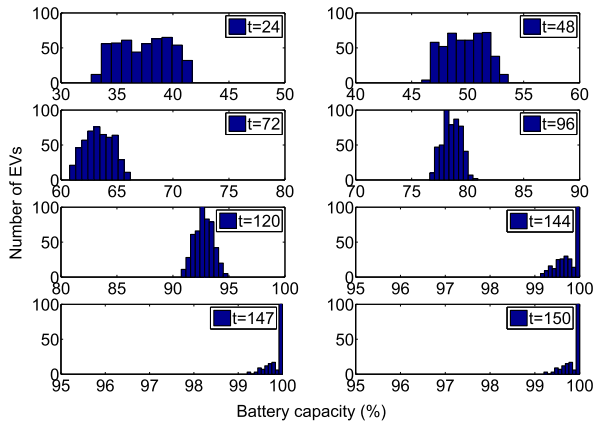
FIGURE 4. Distributions of battery capacity over charging duration ($T^* = 147$).
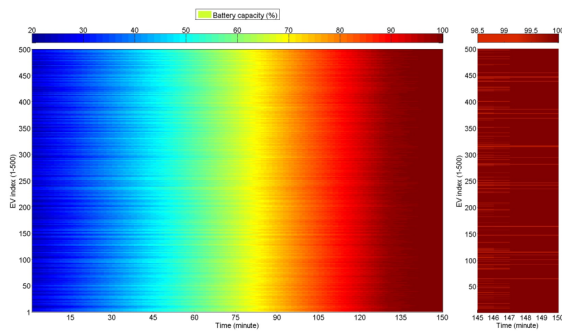

FIGURE 5. Optimal battery capacity dynamics of 500 EVs ($T^* = 147$).

FIGURE 6. Total energy consumption of 500 EVs ($T^* = 147$).
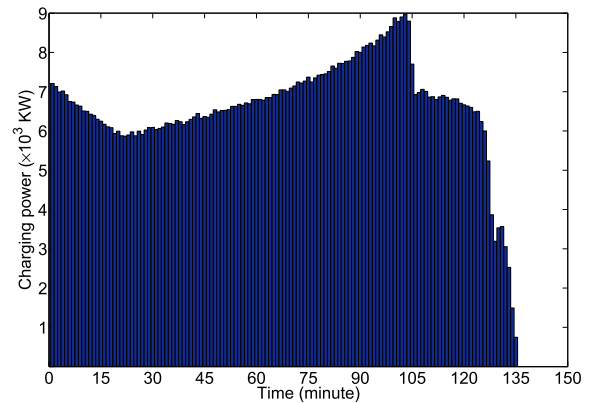
FIGURE 7. Optimal battery capacity dynamics of 500 EVs ($T^* = 135$).

according to the mean field, their behaviour will be similar and the overall result will be less dynamics in time.

Detailed distributions of the battery capacity of all the EVs over the charging period is depicted in Figure 4. The distributions are shown in eight different times from the start of charging towards the finish, with 24 minutes intervals during the first 144 minutes and three minute intervals in the final period from 144 to 150 minutes. As seen, the variation of batteries capacity remains in a range of approximately 10 percent of full capacity. The distributions at different times over the charging period have a fairly low standard deviation that is between 0.8 and 2.3. As the charging moves near to the finish point, the variation becomes smaller. The majority of the EVs finish their charging as the time approaches 147 minutes. At 147 minutes, 23 EVs do not obtain full recharge and the distribution remains unchanged until 150 minutes. This reflects the setting of the quorum rule.

Figure 5 depicts the dynamics of battery capacity for all the 500 EVs (i.e. the optimal trajectory m^*). It can be observed that all EVs behave similarly as they follow the MFE strategy. From the charging station's demand management perspective, the charging is optimised in the sense of balancing the consumption over time. Figure 6 shows the accumulated charging power for the fleet of 500 EVs over the charging time. A reasonably smooth power profile is seen without
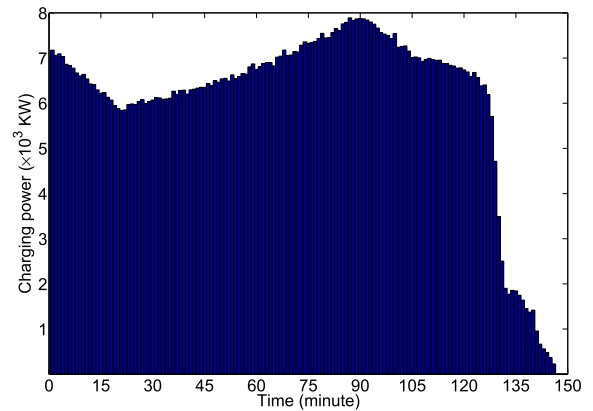
significant instant peaks. Such power profile is claimed to be beneficial to the reliability of the grid. The power profile has a maximum value of $7.9MW$, which means an approximately $18KW$ maximum charging power for every EV.

The charging finishes with a 27 minute delay to the scheduled finishing time. However, it is seen that after 130 minutes, the power drops significantly, because the majority of the EVs has already charged to a high capacity. Therefore, we claim that even delay in operation is introduced, it is not causing severe impact on the reliability of demand management. The station can issue more critical punctuality costs to regulate the charging EVs and urge them to finish sooner. Figure 7 depicts the dynamics of all EVs with an increased terminal cost of $c_1(\hat{t}, \tau) = 4([\tau - \hat{t}]_+)$, while other settings remain the same. As seen, the charging of the batteries become quicker. The charging finishes at $T^* = 135$ minutes, which is an improvement by 12 minutes compared to the previous scenario.

V. CONCLUSION

This paper proposed a dynamic game theoretic optimisation framework based on stochastic mean field game approach for charging electric vehicles in smart grids. It is designed for an optimal charging scenario where a large number of

EVs charge simultaneously in an aggregated charging station. Given the pricing policy of the charging station and the statistical charging status of all EVs, the game theoretic framework provides an optimal solution for every individual EV to proactively control its charging rate in order to minimise the cost of charging. Numerical results have been presented to demonstrate the performance of the proposed framework.

REFERENCES

[1] A. Schuller, B. Dietz, C. M. Flath, and C. Weinhardt, "Charging strategies for battery electric vehicles: Economic benchmark and V2G potential," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2014–2022, Sep. 2014.

[2] W. Su, H. Eichi, W. Zeng, and M.-Y. Chow, "A survey on the electrification of transportation in a smart grid environment," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 1–10, Feb. 2012.

[3] E. Veldman and R. A. Verzijlbergh, "Distribution grid impacts of smart electric vehicle charging from different perspectives," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 333–342, Jan. 2015.

[4] G. K. Venayagamoorthy, P. Mitra, K. Corzine, and C. Huston, "Real-time modeling of distributed plug-in vehicles for V2G transactions," in *Proc. IEEE Energy Convers. Congr. Expo.*, Sep. 2009, pp. 3937–3941.

[5] B.-G. Kim, S. Ren, M. van der Schaar, and J.-W. Lee, "Bidirectional energy trading and residential load scheduling with electric vehicles in the smart grid," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1219–1234, Jul. 2013.

[6] K. Qian, C. Zhou, M. Allan, and Y. Yuan, "Modeling of load demand due to EV battery charging in distribution systems," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 802–810, May 2011.

[7] G. Li and X.-P. Zhang, "Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 492–499, Mar. 2012.

[8] H. Morais, T. Sousa, Z. Vale, and P. Faria, "Evaluation of the electric vehicle impact in the power demand curve in a smart grid environment," *Energy Convers. Manage.*, vol. 82, pp. 268–282, Jun. 2014.

[9] S. Shao, M. Pipattanasomporn, and S. Rahman, "Demand response as a load shaping tool in an intelligent grid with electric vehicles," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 624–631, Dec. 2011.

[10] W. Tushar, C. Yuen, S. Huang, D. B. Smith, and H. V. Poor, "Cost minimization of charging stations with photovoltaics: An approach with EV classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 156–169, Jan. 2016.

[11] Z. Zhu, S. Lambotharan, W. H. Chin, and Z. Fan, "A stochastic optimization approach to aggregated electric vehicles charging in smart grids," in *Proc. IEEE Innov. Smart Grid Technol.-Asia (ISGT Asia)*, May 2014, pp. 51–56.

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

[13] J. Zheng, X. Wang, K. Men, C. Zhu, and S. Zhu, "Aggregation model-based optimization for electric vehicle charging strategy," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 1058–1066, Jun. 2013.

[14] M. H. Amirioun and A. Kazemi, "A new model based on optimal scheduling of combined energy exchange modes for aggregation of electric vehicles in a residential complex," *Energy*, vol. 69, pp. 186–198, May 2014.

[15] Y. Cao *et al.*, "An optimized EV charging model considering TOU price and SOC curve," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 388–393, Mar. 2012.

[16] E. Sortomme and M. A. El-Sharkawi, "Optimal scheduling of vehicle-to-grid energy and ancillary services," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 351–359, Mar. 2012.

[17] E. Sortomme and M. A. El-Sharkawi, "Optimal charging strategies for unidirectional vehicle-to-grid," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 131–138, Mar. 2011.

[18] H. Xing, M. Fu, Z. Lin, and Y. Mou, "Decentralized optimal scheduling for charging and discharging of plug-in electric vehicles in smart grids," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1-10, 2015.

[19] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1095–1105, Sep. 2012.

[20] S. Lasaulce, M. Debbah, and E. Altman, "Methodologies for analyzing equilibria in wireless games," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 41–52, Sep. 2009.

[21] D. H. N. Nguyen and T. Le-Ngoc, "Multiuser downlink beamforming in multicell wireless systems: A game theoretical approach," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3326–3338, Jul. 2011.

[22] I. Atzeni, L. G. Ordonez, G. Scutari, D. P. Palomar, and J. R. Fonollosa, "Noncooperative day-ahead bidding strategies for demand-side expected cost minimization with real-time adjustments: A GNEP approach," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2397–2412, May 2014.

[23] Z. Zhu, S. Lambotharan, W. H. Chin, and Z. Fan, "A game theoretic optimization framework for home demand management incorporating local energy resources," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 353–362, Apr. 2015.

[24] Z. Zhu, "Mathematical optimization techniques for demand management in smart grids," Ph.D. dissertation, School Electron., Elect.Syst. Eng., Loughborough Univ., Loughborough, U.K., 2014.

[25] I. Atzeni, L. G. Ordonez, G. Scutari, D. P. Palomar, and J. R. Fonollosa, "Noncooperative and cooperative optimization of distributed energy generation and storage in the demand-side of the smart grid," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2454–2472, May 2013.

[26] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 86–105, Sep. 2012.

[27] E. L. Karfopoulos and N. D. Hatziargyriou, "A multi-agent system for controlled charging of a large population of electric vehicles," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1196–1204, May 2013.

[28] W. Tushar, W. Saad, H. V. Poor, and D. B. Smith, "Economics of electric vehicle charging: A game theoretic approach," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1767–1778, Dec. 2012.

[29] R. Couillet, S. M. Perlaza, H. Tembine, and M. Debbah, "Electrical vehicles in the smart grid: A mean field game analysis," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 6, pp. 1086–1096, Jul. 2012.

[30] J.-M. Lasry and P.-L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, no. 1, pp. 229–260, 2007.

[31] Y. Wang, F. R. Yu, H. Tang, and M. Huang, "A mean field game theoretic approach for security enhancements in mobile ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1616–1627, Mar. 2014.

[32] P. Semasinghe and E. Hossain, "Downlink power control in self-organizing dense small cells underlaying macrocells: A mean field game," *IEEE Trans. Mobile Comput.*, vol. 15, no. 2, pp. 350–363, Feb. 2016.

[33] C. Li *et al.*, "A new stepwise power tariff model and its application for residential consumers in regulated electricity markets," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 300–308, Feb. 2013.

[34] M. Greer, *Electricity Cost Modeling Calculations*. Amsterdam, The Netherlands: Elsevier, 2010.

[35] R. E. Bohn, "Spot pricing of public utility services," Energy Lab., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. MIT-EL 82-031, May 1982.

[36] K. Itô, "Multiple Wiener integral," *J. Math. Soc. Jpn.*, vol. 3, no. 1, pp. 157–169, 1951.

[37] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*. Berlin, Germany: Springer, 1972.

[38] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, Jul. 1965.

[39] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory* (Classics in Applied Mathematics), 2nd ed. Philadelphia, PA, USA: SIAM, 1999.

[40] O. Guéant, P.-L. Lions, and J.-M. Lasry, "Mean field games and applications," in *Paris-Princeton Lectures on Mathematical Finance 2010*. Berlin, Germany: Springer, 2011.

[41] A. F. Hanif, H. Tembine, M. Assaad, and D. Zeghlache, "Mean-field games for resource sharing in cloud-based networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 624–637, Feb. 2016.

[42] F. Parise, S. Grammatico, M. Colombino, and J. Lygeros, "On constrained mean field control for large populations of heterogeneous agents: Decentralized convergence to Nash equilibria," in *Proc. Eur. Control Conf. (ECC)*, Jul. 2015, pp. 3316–3321.

[43] K. C. Border, *Fixed Point Theorems With Applications to Economics and Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

ZIMING ZHU (M'14) received the B.Eng. degree in measuring and control technology and instrumentations from the Huazhong University of Science and Technology, China, the M.Sc. degree in communication networks and signal processing from Bristol University, U.K., and the Ph.D. degree in electronic and electrical engineering from Loughborough University, U.K. He is currently a Research Engineer with Toshiba Research Europe, Bristol, U.K. His research interests include game theory and optimization techniques for communication networks and smart grids.

SANGARAPILLAI LAMBOTHARAN (SM'06) received the Ph.D. degree in signal processing from Imperial College London, U.K., in 1997. He was a Post-Doctoral Research Associate until 1999. He was a Visiting Scientist with the Engineering and Theory Center, Cornell University, NY, USA, in 1996. From 1999 to 2002, he was with the Motorola Applied Research Group, U.K., as a Research Engineer, working in many various projects, including physical-link layer modeling and performance characterization of GPRS, EGPRS, and UTRAN. From 2002 to 2007, he was with King's College London, U.K., and Cardiff University, U.K., as a Lecturer and Senior Lecturer, respectively. He is currently a Professor of Digital Communications and the Head of Signal Processing and Networks Research Group, Loughborough University, U.K. He has authored over 150 conference and journal articles in these areas. His current research interests include wireless communications, cognitive radio networks, smart grids, radars, convex optimizations, and game theory. He serves as an Associate Editor of the *EURASIP Journal on Wireless Communications and Networking*.

WOON HAU CHIN (SM'10) received the B.Eng. (Hons.) degree, the M.Eng. degree in electrical engineering from the National University of Singapore, and the Ph.D. degree in electrical and electronic engineering from Imperial College London. He was with the Institute for Infocomm Research, Singapore. He is currently the Research Manager with the Physical Layer Research Group, Toshiba Research Europe, Bristol, U.K. His research interests are statistical signal processing and smart grid communications. He has contributed to multiple standards, including ETSI TC SmartBAN, the IEEE 802.11n, and 3GPP LTE.

ZHONG FAN received the B.S. and M.S. degrees in electronics engineering from Tsinghua University, China, and the Ph.D. degree in telecommunication networks from Durham University, U.K. He was a Research Fellow with Cambridge University, a Lecturer with Birmingham University, and a Researcher with Marconi Labs Cambridge. He is currently a Chief Research Fellow with Toshiba Research Europe, Bristol, U.K. His research interests are wireless networks, IP networks, M2M, and smart grid communications. He was also awarded a BT Short-Term Fellowship to work at BT Labs.

● ● ●