Real-Time Welfare-Maximizing Regulation Allocation in Dynamic Aggregator-EVs System

Sun Sun, *Student Member, IEEE*, Min Dong, *Senior Member, IEEE*, and Ben Liang, *Senior Member, IEEE*

***Abstract*—The concept of vehicle-to-grid (V2G) has gained recent interest as more and more electric vehicles (EVs) are put to use. In this paper, we consider a dynamic aggregator-EVs system, where an aggregator centrally coordinates a large number of dynamic EVs to provide regulation service. We propose a Welfare-Maximizing Regulation Allocation (WMRA) algorithm for the aggregator to fairly allocate the regulation amount among the EVs. Compared with previous works, WMRA accommodates a wide spectrum of vital system characteristics, including dynamics of EV, limited EV battery size, EV battery degradation cost, and the cost of using external energy sources for the aggregator. The algorithm operates in real time and does not require any prior knowledge of the statistical information of the system. Theoretically, we demonstrate that WMRA is away from the optimum by $O(1/V)$, where V is a controlling parameter depending on EVs' battery size. In addition, our simulation results indicate that WMRA can substantially outperform a suboptimal greedy algorithm.**

***Index Terms*—Aggregator-EVs system; electric vehicles; real-time algorithm; V2G; welfare-maximizing regulation allocation.**

I. Introduction

Electrification of personal transportation is expected to become prevalent in the near future. For example, from one report of the U.S. department of energy [1], the government sets an ambitious goal to put one million EVs on the road by 2015. Besides serving the purpose of transportation, EVs can also be used as distributed electricity generation/storage devices when plugged-in [2]. Hence, the concept of vehicle-to-grid (V2G), referring to the integration of EVs to the power grid, has received increasing attention [2], [3].

Frequency regulation service is to balance power generation and load demand in a short time scale, so as to maintain the frequency of a power grid at its nominal value. Traditionally, regulation service is provided by fast responsive generators, which vary their output to alleviate power deficits or surpluses, and is the most expensive ancillary service [4]. Experiments show that EV's power electronics and battery can well respond to the frequent regulation signals. Thus it is possible to exploit a plugged-in EV as a promising alternative to provide regulation service through charging/discharging, which potentially could reduce the cost of regulation service significantly [5]. However, since the regulation service is generally requested on the order of megawatts while the power capacity of an EV is typically 5-20 kW, it is often necessary for an aggregator to coordinate a large number of EVs to provide regulation service [6]. In addition, frequent charging/discharging has a detrimental effect on EV's battery life. Thus, it is important to design proper algorithm for regulation allocation in the aggregator-EVs system, especially in a real-time fashion.

There is a growing body of recent works on V2G regulation service. Specific to the aggregator-EVs system, which focuses on the interaction between the aggregator and the EVs, centralized regulation allocation is studied in [7]–[11], where the objective is to maximize the profit of the aggregator or the EVs. In [7], a set of schemes based on different criteria of fairness among the EVs are provided. In [8], the regulation allocation problem is formulated as quadratic programming. In [9], considering both regulation service and spinning reserves, the underlying problem is formulated as linear programming. In [10], the charging behavior of EVs is also considered, and the underlying problem is then reduced to the control of charging sequence and charging rate of each EV, which is solved by dynamic programming. In [11], a real-time regulation control algorithm is proposed by formulating the problem as a Markov decision process, with the action space consisting of charging, discharging, and regulation. Finally, a distributed regulation allocation system is proposed in [12] using game theory, and a smart pricing policy is developed to incentivize EVs.

In addressing the regulation allocation problem, however, these earlier works have omitted to consider some essential characteristics of the aggregator-EVs system. For example, deterministic model is used in [7] and [10], which ignore the uncertainty of the system, e.g., the uncertainty of the electricity prices. The dynamics of the regulation signals is not incorporated in [12], nor the energy restriction of EV battery is considered. The self-charging/discharging activities in support of EV's own need are omitted in [7] and [12]. The potential cost of using external energy sources for the aggregator to accomplish regulation service is ignored in [7]–[11], and the cost of EV battery degradation due to frequent charing/discharging in regulation service is not considered in [8], [10]–[12].

In this work, we consider all of the above factors in a more complete aggregator-EVs system model, and develop a real-time algorithm for the aggregator to fairly allocate the regulation amount among the EVs. Specifically, considering an aggregator-EVs system providing long-term regulation service to a power grid, we aim to maximize the long-term social

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

Sun Sun and Ben Liang are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada (email: {ssun, liang}@comm.utoronto.ca).

Min Dong is with the Department of Electrical Computer and Software Engineering, University of Ontario Institute of Technology, Toronto, Canada (email: min.dong@uoit.ca).

welfare of the aggregator-EVs system, with the constraints on each EV's regulation amount and degradation cost. To solve such a stochastic optimization problem, we adopt Lyapunov optimization technique, which is also used in [13]–[15] for demand side management in smart grid. We demonstrate how a solution to this maximization can be formulated under a general Lyapunov optimization framework [16], and propose a real-time allocation strategy specific to the aggregator-EVs system. The proposed Welfare-Maximizing Regulation Allocation (WMRA) algorithm does not rely on any statistical information of the system, and is shown to be asymptotically close to the optimum as EV's battery capacity increases. Finally, WMRA is compared to a greedy algorithm through simulation and is shown to offer substantial performance gains.

In our preliminary version of this work [17], the EVs are ideally assumed to be static, *i.e.*, they are in the aggregator-EVs system throughout the operational time. In this paper, to more realistically capture the dynamics of the aggregator-EVs system, we generalize the system model in [17] to accommodate dynamic EVs, which is considered in none of the previous works [7]–[12]. This generalization is challenging for the centralized control of regulation allocation, since the returning EV may have a different energy state compared with the last leaving energy state, and this energy difference will impose much more difficulties on the aggregator for handling EV's battery size constraint. To tackle this difficulty, we design a novel virtual queue to track the energy state of each EV. Through a careful design of the dynamics of the virtual queue, we can ensure that the battery size constraint of the EV is always satisfied.

The remainder of this paper is organized as follows. We describe the system model and formulate the regulation allocation problem in Section II. In Section III, we propose WMRA, and in Section IV we analyze its performance. Simulations are exhibited in Section V, and we conclude in Section VI.

Notation: Denote $[a]^+$ as $\max\{a,0\}$, $[a,b]^+$ as $\max\{a,b\}$, and $[a,b]^-$ as $\min\{a,b\}$. The main symbols used in this paper are summarized in Table I.

TABLE I
LIST OF MAIN SYMBOLS

G_t	regulation signal at time slot t
$\mathbf{1}_{d,t}$	indicator of regulation down at time slot t
$\mathbf{1}_{u,t}$	indicator of regulation up at time slot t
Δt	interval of regulation signals
N	number of registered EVs
$t_{ir,k}$	k-th returning time slot of the i-th EV
$t_{il,k}$	k-th leaving time slot of the i-th EV
$\mathcal{T}_{i,r}$	set of returning time slots for the i-th EV
$\mathcal{T}_{i,l}$	set of leaving time slots for the i-th EV
$\mathcal{T}_{i,p}$	set of all participating time slots for the i-th EV
$\mathbf{1}_{i,t}$	indicator of the i-th EV's dynamics at time slot t
$x_{id,t}$	regulation down amount of the i-th EV at time slot t
$x_{iu,t}$	regulation up amount of the i-th EV at time slot t
$x_{i,\max}$	upper bound on $x_{id,t}$ and $x_{iu,t}$
$x_{i,t}$	regulation amount of the i-th EV at time slot t
$h_{id,t}$	effective upper bound on $x_{id,t}$
$h_{iu,t}$	effective upper bound on $x_{iu,t}$
$s_{i,t}$	energy state of the i-th EV at the beginning of time slot t
$s_{i,\text{cap}}$	battery capacity of the i-th EV
$s_{i,\min}$	lower bound on $s_{i,t}$
$s_{i,\max}$	upper bound on $s_{i,t}$
$\Delta_{i,k}$	difference between the i-th EV's $(k+1)$-th returning energy state and the k-th leaving energy state
$C_i(\cdot)$	degradation cost function of the i-th EV
$c_{i,\max}$	upper bound on $C_i(\cdot)$
$c_{i,\text{up}}$	upper bound on long-term degradation cost of the i-th EV
$e_{s,t}$	unit cost of clearing energy surplus
$e_{d,t}$	unit cost of clearing energy deficit
$e_{\min}$	lower bound on $e_{s,t}$ and $e_{d,t}$
$e_{\max}$	upper bound on $e_{s,t}$ and $e_{d,t}$
ω_i	normalized weight of the i-th EV

II. System Model and Problem Formulation

In this section, we propose a centralized dynamic aggregator-EVs system and formulate the regulation allocation problem mathematically.

A. Aggregator-EVs System and Regulation Service

Consider a long-term time-slotted system, in which the regulation service is provided over equal time intervals of length Δt. At the beginning of each time slot $t \in \mathcal{T} \triangleq \{0,1,\cdots\}$, the aggregator receives a random regulation signal G_t from a power grid. If $G_t > 0$, the aggregator is required to provide *regulation down* service by absorbing G_t units of energy from the power grid during time slot t; if $G_t < 0$, the aggregator is required to provide *regulation up* service by contributing $|G_t|$ units of energy to the power grid during time slot t. To represent the type of the regulation service at time slot t, we define the indicator random variables $\mathbf{1}_{d,t} \triangleq \begin{cases} 1, & \text{if } G_t > 0 \\ 0, & \text{otherwise} \end{cases}$ and $\mathbf{1}_{u,t} \triangleq \begin{cases} 1, & \text{if } G_t < 0 \\ 0, & \text{otherwise} \end{cases}$. Note that the product $\mathbf{1}_{d,t} \cdot \mathbf{1}_{u,t} = 0$, since regulation down and up services cannot happen simultaneously.

To provide regulation service, the aggregator coordinates N registered EVs and can communicate with each EV bi-directionally when the EV is plugged-in. Each EV can leave the system for personal reason or for self-charging/discharging purpose and re-join the system later. Assume that each EV provides regulation service only if it is in the system.

For the i-th EV, denote $t_{ir,k} \in \mathcal{T}$ as its k-th returning time slot and $t_{il,k} \in \mathcal{T}$ as its k-th leaving time slot with $t_{ir,k} < t_{il,k}$, $\forall k \in \{1,2,\cdots\}$. For simplicity of analysis, assume that all EVs are in the system at the initial time and thus $t_{ir,1} = 0, \forall i$. Define the set of the returning time slots of the i-th EV as $\mathcal{T}_{i,r} \triangleq \{t_{ir,1}, t_{ir,2}, \cdots\}$ and the set of its leaving time slots as $\mathcal{T}_{i,l} \triangleq \{t_{il,1}, t_{il,2}, \cdots\}$, with $t_{ir,k} < t_{ir,k+1}$ and $t_{il,k} < t_{il,k+1}$. Define

$$\mathcal{T}_{i,p} \triangleq \cup_{k=1}^{\infty} \{t_{ir,k}, t_{ir,k}+1, \cdots, t_{il,k}-1\}$$

as the set containing all participating time slots of the i-

th EV for regulation service. Hence, the i-th EV is in the system for any $t \in \mathcal{T}_{i,p}$. Define the indicator random variable $\mathbf{1}_{i,t} \triangleq \begin{cases} 1, & \text{if } t \in \mathcal{T}_{i,p} \\ 0, & \text{otherwise} \end{cases}$ to represent the dynamics of the i-th EV at time slot t (*i.e.*, whether the i-th EV is in the system at time slot t). Define the vector $\mathbf{1}_t \triangleq [\mathbf{1}_{1,t}, \cdots, \mathbf{1}_{N,t}]$ to represent the dynamics of all EVs at time slot t.

At the beginning of each time slot, the aggregator allocates regulation amount among all *participating* EVs. Denote $x_{id,t} \geq 0$ as the amount of regulation down energy allocated to the i-th EV through charging, and $x_{iu,t} \geq 0$ as the amount of regulation up energy contributed by the i-th EV through discharging. Due to the limitation of charging/discharging circuit in battery, assume that $x_{id,t}$ and $x_{iu,t}$ are upper bounded by $x_{i,\max} > 0$. Note that if the i-th EV is out of the system at time slot t, *i.e.*, $\mathbf{1}_{i,t} = 0$, then it cannot provide regulation service and we have $x_{id,t} = x_{iu,t} = 0$. Define the vectors $\mathbf{x}_{d,t} \triangleq [x_{1d,t}, \cdots, x_{Nd,t}]$ and $\mathbf{x}_{u,t} \triangleq [x_{1u,t}, \cdots, x_{Nu,t}]$ to represent the regulation amounts of all EVs at time slot t.

For the i-th EV, assume that it is in the system at time slot t (*i.e.*, $\mathbf{1}_{i,t} = 1$), and thus can provide regulation service. Denote $s_{i,t} \in [0, s_{i,\text{cap}}]$ as its energy state at the beginning of time slot t, with $s_{i,\text{cap}}$ being its battery capacity. Due to the regulation service, the energy state of the i-th EV at the beginning of time slot $t+1$ is given by

$$s_{i,t+1} = s_{i,t} + \mathbf{1}_{d,t} x_{id,t} - \mathbf{1}_{u,t} x_{iu,t} = s_{i,t} + b_{i,t}, \quad (1)$$

where

$$b_{i,t} \triangleq \mathbf{1}_{d,t} x_{id,t} - \mathbf{1}_{u,t} x_{iu,t} \quad (2)$$

is defined to be the effective charging/discharging amount of the i-th EV at time slot t. Charging a battery near its capacity or discharging it close to the zero energy state can significantly reduce battery's lifetime [18]. Therefore, lower and upper bounds on the battery energy state are usually imposed by its manufacturer or user. Denote the interval $[s_{i,\min}, s_{i,\max}]$ as the preferred energy range of the i-th EV with $0 \leq s_{i,\min} < s_{i,\max} \leq s_{i,\text{cap}}$. Then, the resultant energy state $s_{i,t+1}$ in (1) should lie in $[s_{i,\min}, s_{i,\max}]$, which indicates that the regulation amounts $x_{id,t}$ and $x_{iu,t}$ must satisfy $0 \leq x_{id,t} \leq \mathbf{1}_{i,t} h_{id,t}$ and $0 \leq x_{iu,t} \leq \mathbf{1}_{i,t} h_{iu,t}$, respectively, where $h_{id,t}$ and $h_{iu,t}$ are effective upper bounds on the regulation amounts and are defined as

$$h_{id,t} \triangleq [x_{i,\max}, s_{i,\max} - s_{i,t}]^-,$$

and

$$h_{iu,t} \triangleq [x_{i,\max}, s_{i,t} - s_{i,\min}]^-,$$

respectively.

From time to time, the i-th EV may need to stop its regulation service and leave the system. When the EV is out of the system (*i.e.*, $\mathbf{1}_{i,t} = 0$), it cannot offer regulation service and the aggregator has no information of the EV's energy state. Moreover, the dynamics of the energy state may not follow (1) when $\mathbf{1}_{i,t} = 0$. When returning, the EV may have a different energy state compared with its last leaving energy state. Assume that all returning energy states of the i-th EV are confined in the preferred energy range by the EV's self-control, *i.e.*, $s_{i,t} \in [s_{i,\min}, s_{i,\max}], \forall t \in \mathcal{T}_{i,r}$. Define

$$\Delta_{i,k} \triangleq s_{i,t_{ir,k+1}} - s_{i,t_{il,k}}, \forall k \in \{1, 2, \cdots\} \quad (3)$$

as the difference between the i-th EV's $(k+1)$-th returning energy state and its last leaving energy state. We assume that

A1) $\Delta_{i,k}$ is bounded, *i.e.*, $|\Delta_{i,k}| \leq \Delta_{i,\max}$, where the constant $\Delta_{i,\max} \geq 0$.

A2) $\Delta_{i,k}$ has mean zero, *i.e.*, $\mathbb{E}[\Delta_{i,k}] = 0$, $\forall k$.

Note that A2 is a mild assumption, based on the random behavior of each EV when it is out of the system.

For each EV, providing regulation service incurs battery degradation due to frequent charging/discharging activities. Denote $C_i(x)$ as the degradation cost function of the regulation amount of the i-th EV, with $0 \leq C_i(x) \leq c_{i,\max}$ and $C_i(0) = 0$. Since faster charging or discharging, *i.e.*, larger value of $x_{id,t}$ or $x_{iu,t}$, has a more detrimental effect on the battery's lifetime, we assume $C_i(x)$ to be convex, continuous, and non-decreasing on the interval $[0, x_{i,\max}]$. We further assume that each EV imposes an upper bound $c_{i,\text{up}} \in [0, c_{i,\max}]$ on the time-averaged battery degradation, expressed by

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\mathbf{1}_{d,t} C_i(x_{id,t}) + \mathbf{1}_{u,t} C_i(x_{iu,t})\right] \leq c_{i,\text{up}}.$$

The total regulation amount provided by the EVs may be insufficient to meet the requested regulation amount due to, for example, a lack of participating EVs, or high battery degradation cost. For brevity, define

$$x_{i,t} \triangleq \mathbf{1}_{d,t} x_{id,t} + \mathbf{1}_{u,t} x_{iu,t}, \quad 0 \leq x_{i,t} \leq x_{i,\max}$$

as the regulation amount allocated to the i-th EV at time slot t, which equals either $x_{id,t}$ or $x_{iu,t}$. Then, the insufficiency of the regulation amount is indicated by $\sum_{i=1}^{N} x_{i,t} < |G_t|$, with the gap $|G_t| - \sum_{i=1}^{N} x_{i,t}$ representing an energy surplus in the case of regulation down or an energy deficit in the case of regulation up. Assume that energy surplus or energy deficit must be cleared, or the regulation service fails. Therefore, from time to time, the aggregator has to exploit more expensive external energy sources, such as from the traditional regulation market, so as to fill the energy gap. Denote the unit costs for clearing energy surplus and energy deficit at time slot t as $e_{s,t}$ and $e_{d,t}$, respectively, which are both random but are restricted in the interval $[e_{\min}, e_{\max}]$. Then, the cost of the aggregator for using the external energy sources at time slot t is given by

$$e_t \triangleq \mathbf{1}_{d,t} e_{s,t} \Big(G_t - \sum_{i=1}^{N} x_{id,t} \Big) + \mathbf{1}_{u,t} e_{d,t} \Big(|G_t| - \sum_{i=1}^{N} x_{iu,t} \Big),$$

where we have implicitly assumed that the total regulation amount provided by all EVs cannot exceed the requested amount.

B. Fair Regulation Allocation through Welfare Maximization

The objective of the aggregator is to maximize the long-term social welfare of the aggregator-EVs system. Specifically, the aggregator aims to fairly allocate the regulation amount among EVs and to reduce the cost for the expensive external energy sources, with the constraints on each EV's regulation amount and degradation cost. To this end, we formulate the regulation allocation problem as the following stochastic optimization problem[1]:

P1:

$$\max_{\mathbf{x}_{d,t},\mathbf{x}_{u,t}} \sum_{i=1}^{N} \omega_i U\big(\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[x_{i,t}]\big) - \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[e_t]$$

$$\text{s.t.} \quad 0 \le x_{id,t} \le \mathbf{1}_{i,t} h_{id,t}, \quad \forall i,t \tag{4}$$

$$0 \le x_{iu,t} \le \mathbf{1}_{i,t} h_{iu,t}, \quad \forall i,t \tag{5}$$

$$\sum_{i=1}^{N} x_{id,t} \le \mathbf{1}_{d,t} G_t, \quad \forall t \tag{6}$$

$$\sum_{i=1}^{N} x_{iu,t} \le \mathbf{1}_{u,t} |G_t|, \quad \forall t \tag{7}$$

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\mathbf{1}_{d,t}C_i(x_{id,t}) + \mathbf{1}_{u,t}C_i(x_{iu,t})\right] \le c_{i,\text{up}}, \forall i, \tag{8}$$

where $\omega_i > 0$ is the normalized weight associated with the i-th EV, and $U(\cdot)$ is a utility function assumed to be concave, continuous, and non-decreasing, with $U(0) = 0$. Furthermore, to facilitate later analysis, we make a mild assumption that the utility function $U(\cdot)$ satisfies

$$U(x) \le U(0) + \mu x, \forall x \in \left[0, \max_{1\le i\le N}\{x_{i,\max}\}\right], \tag{9}$$

where the constant $\mu > 0$. One sufficient condition for (9) to hold is that $U(\cdot)$ has finite positive derivate at zero, such as $U(x) = \log(1+x)$. The expectations in the above optimization problem are taken over the randomness of the system and the possible randomness of the regulation allocation.

In the objective function of **P1**, the first term includes each EV's welfare under the utility function $U(\cdot)$ and the weight ω_i, and the second term reflects the aggregator's cost for exploiting external energy sources. Note that the fairness of the regulation allocation among EVs is ensured by the utility function $U(\cdot)$, and various types of fairness can be achieved by using different utility functions [19]. For each EV, in (4) and (5), hard constraints on the regulation amounts are set at each time slot, while in (8), a long-term time-averaged constraint on the regulation amount is set due to the battery degradation. The constraints (6) and (7) ensure that $x_{id,t} = 0$ for regulation up and $x_{iu,t} = 0$ for regulation down.

III. Welfare-Maximizing Regulation Allocation

In this section, we first apply a sequence of two reformulations to **P1**, then propose a real-time welfare-maximizing regulation allocation (WMRA) algorithm to solve the resultant optimization problem. The performance analysis of the proposed WMRA will be shown in Section IV.

[1]For EVs that only visit the system finite times, since they only affect the system's transient behavior, but not the long-term behavior, we can ignore them and only consider the rest EVs that leave and re-join the system infinite times.

A. Problem Transformation

The objective of **P1** contains a function of a long-term time average, which complicates the problem. Fortunately, in general, such a problem can be transformed to a problem of maximizing the long-term time average of the function [16]. Specifically, we transform **P1** as follows.

We first introduce an auxiliary vector $\mathbf{z}_t \triangleq [z_{1,t}, \cdots, z_{N,t}]$ with the constraints

$$0 \le z_{i,t} \le x_{i,\max}, \forall i, t, \text{ and} \tag{10}$$

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[z_{i,t}] = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[x_{i,t}], \forall i. \tag{11}$$

From the above constraints, the auxiliary variable $z_{i,t}$ and the regulation allocation amount $x_{i,t}$ lie in the same range and have the same long-term time average behavior. We next consider the following problem.

P2:

$$\max_{\mathbf{x}_{d,t},\mathbf{x}_{u,t},\mathbf{z}_t} \quad \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\left(\sum_{i=1}^{N}\omega_i U(z_{i,t})\right) - e_t\right]$$
$$\text{s.t.} \quad (4), (5), (6), (7), (8), (10), \text{ and } (11).$$

Compared with **P1**, **P2** is over $\mathbf{x}_{d,t}$, $\mathbf{x}_{u,t}$ and $\mathbf{z}_t$ with two more constraints (10) and (11). Nevertheless, **P2** contains no function of time average; instead, it maximizes the long-term time average of the expected social welfare.

Denote $(\mathbf{x}_{d,t}^{\text{opt}}, \mathbf{x}_{u,t}^{\text{opt}})$ as an optimal solution to **P1**, and $(\mathbf{x}_{d,t}^{*}, \mathbf{x}_{u,t}^{*}, \mathbf{z}_t^{*})$ as an optimal solution to **P2**. Define $\bar{\mathbf{z}}_t^{\text{opt}} \triangleq [\bar{z}_{1,t}^{\text{opt}}, \cdots, \bar{z}_{N,t}^{\text{opt}}]$ with the i-th element

$$\bar{z}_{i,t}^{\text{opt}} \triangleq \lim_{T\to\infty}\frac{1}{T}\sum_{\tau=0}^{T-1}\mathbb{E}[x_{i,\tau}^{\text{opt}}], \quad \forall i, t,$$

where $x_{i,\tau}^{\text{opt}} \triangleq \mathbf{1}_{d,\tau} x_{id,\tau}^{\text{opt}} + \mathbf{1}_{u,\tau} x_{iu,\tau}^{\text{opt}}$. Denote the objective functions of **P1** and **P2** as $f_1(\cdot)$ and $f_2(\cdot)$, respectively. The equivalence of **P1** and **P2** is stated below.

Lemma 1: **P1** and **P2** have the same optimal objective, *i.e.*, $f_1(\mathbf{x}_{d,t}^{\text{opt}}, \mathbf{x}_{u,t}^{\text{opt}}) = f_2(\mathbf{x}_{d,t}^{*}, \mathbf{x}_{u,t}^{*}, \mathbf{z}_t^{*})$. Furthermore, $(\mathbf{x}_{d,t}^{\text{opt}}, \mathbf{x}_{u,t}^{\text{opt}}, \bar{\mathbf{z}}_t^{\text{opt}})$ is an optimal solution to **P2**, and $(\mathbf{x}_{d,t}^{*}, \mathbf{x}_{u,t}^{*})$ is an optimal solution to **P1**.

Proof: The proof follows the general framework given in [16]. Details specific to our system are given in Appendix A. ■

Lemma 1 indicates that the transformation from **P1** to **P2** results in no loss of optimality. Thus, in the following, we will focus on solving **P2** instead.

B. Problem Relaxation

P2 is still a challenging problem since in the constraints (4) and (5), the regulation allocation amount of each EV depends on its current energy state $s_{i,t}$, hence coupling with all previous regulation allocation amounts. To avoid such coupling, we relax the constraints of $x_{id,t}$ and $x_{iu,t}$, and introduce the optimization problem **P3** below.

P3:

$$\max_{\mathbf{x}_{d,t},\mathbf{x}_{u,t},\mathbf{z}_t} \quad \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\left(\sum_{i=1}^{N}\omega_i U(z_{i,t})\right) - e_t\right]$$

$$\text{s.t.}\quad 0 \le x_{id,t} \le \mathbf{1}_{i,t} x_{i,\max}, \forall i, t, \tag{12}$$

$$0 \le x_{iu,t} \le \mathbf{1}_{i,t} x_{i,\max}, \forall i, t, \tag{13}$$

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[b_{i,t}] = 0, \forall i, \tag{14}$$

$$(6), (7), (8), (10), \text{ and } (11),$$

where in (14) $b_{i,t}$ is the effective charging/discharging amount defined in (2). In **P3**, we have replaced the constraints (4) and (5) in **P2** with (12)–(14), thus have removed the dependence of the regulation amount on $s_{i,t}$. We next demonstrate that, any $(\mathbf{x}_{d,t}, \mathbf{x}_{u,t})$ that satisfies (4) and (5) also satisfies (12)–(14). Therefore, **P3** is a relaxed problem of **P2**.

Consider the i-th EV. The constraints (4) and (5) in **P2** are equivalent to the following two sub-constraints:
if $\mathbf{1}_{i,t} = 1$, then

$$0 \le x_{id,t} \le x_{i,\max} \tag{15}$$

$$0 \le x_{iu,t} \le x_{i,\max} \tag{16}$$

$$s_{i,\min} \le s_{i,t+1} \le s_{i,\max}; \tag{17}$$

if $\mathbf{1}_{i,t} = 0$, then

$$x_{id,t} = x_{iu,t} = 0. \tag{18}$$

Since (15), (16), and (18) are equivalent to (12) and (13), we are left to justify that (17) (*i.e.*, the boundedness of $s_{i,t}$) implies (14). Recall that $s_{i,t}$ is bounded for any returning time slot $t \in \mathcal{T}_{i,r}$ by the EV's self-control. Together, we need to justify that if $s_{i,t} \in [s_{i,\min}, s_{i,\max}], \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$, then the constraint (14) holds. This result is shown in the following lemma.

Lemma 2: For the i-th EV, under the assumption A2, if $s_{i,t} \in [s_{i,\min}, s_{i,\max}], \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$, then the constraint (14) holds, *i.e.*, $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[b_{i,t}] = 0$.

Proof: See Appendix B. ■

From Lemma 2, we know that, the boundedness of $s_{i,t}$ indeed implies (14), which completes our demonstration that **P3** is a relaxed version of **P2** with a larger feasible solution set. Later, we will show in Section IV-A that our proposed algorithm for **P3** in fact ensures the boundedness of $s_{i,t}$, and thus provides a feasible solution to **P2** and to the original problem **P1**.

The relaxed problem **P3** allows us to apply Lyapunov optimization to design a real-time algorithm for solving welfare maximization. To our best knowledge, this relaxation technique to accommodate the type of time-coupled action constraints such as (4) and (5) is first introduced in [20] for a power-cost minimization problem in data centers equipped with an energy storage device. Unlike in [20], the structure of our problem is more complicated, where the dynamics of the distributed storage devices (EVs) are considered, as well as a nonlinear objective which allows both positive and negative values for the energy requirement G_t. Thus, the algorithm design is more involved to ensure that the original constraints in **P2** are satisfied.

C. WMRA Algorithm

In this subsection, we propose a WMRA algorithm to solve **P3** by employing Lyapunov optimization technique.

We first define three virtual queues for each EV with the associated queue backlogs $J_{i,t}$, $H_{i,t}$, and $K_{i,t}$. The evolutionary behaviors of $J_{i,t}$, $H_{i,t}$, and $K_{i,t}$ are as follows:

$$J_{i,t+1} = [J_{i,t} + \mathbf{1}_{d,t} C_i(x_{id,t}) + \mathbf{1}_{u,t} C_i(x_{iu,t}) - c_{i,\text{up}}]^+; \tag{19}$$

$$H_{i,t+1} = H_{i,t} + z_{i,t} - x_{i,t}; \tag{20}$$

$$K_{i,t} = \begin{cases} s_{i,t} - c_i, & \text{if } t \in \mathcal{T}_{i,r} \quad \text{(21a)} \\ K_{i,t-1} + b_{i,t-1}, & \text{otherwise,} \quad \text{(21b)} \end{cases}$$

where in (21a) we have designed the constant $c_i = s_{i,\min} + 2x_{i,\max} + V(\omega_i \mu + e_{\max})$ with $V \in (0, V_{\max}]$ and

$$V_{\max} = \min_{1\le i\le N} \left\{ \frac{s_{i,\max} - s_{i,\min} - 4x_{i,\max}}{2(\omega_i \mu + e_{\max})} \right\}. \tag{22}$$

The role of V will be explained later. It will also be clear in Section IV-A that the specific expressions of c_i and $V_{\max}$ are designed to ensure the boundedness of $s_{i,t}$. Note that $x_{i,\max}$ is generally much smaller than the energy capacity. For example, for the Tesla Model S base model [21], the energy capacity is 40 kWh, and $x_{i,\max} = 0.166$ kWh if the maximum charging rate 10 kW is applied and the regulation duration is 1 minute. Therefore, $V_{\max} > 0$ holds in general.

From (21a), $K_{i,t}$ is re-initialized as a shifted version of $s_{i,t}$ every time the i-th EV returning to the aggregator-EVs system; also, from (21b), $K_{i,t}$ evolves the same as $s_{i,t}$ for $t \in \mathcal{T}_{i,p}$ (recall that the dynamics of $s_{i,t}$ may not follow (1) when $\mathbf{1}_{i,t} = 0$). Therefore, $K_{i,t}$ is essentially a shifted version of $s_{i,t}, \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$, *i.e.*,

$$K_{i,t} = s_{i,t} - c_i, \quad \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}. \tag{23}$$

Additionally, since the effective charging/discharging amount $b_{i,t} = 0$ when $\mathbf{1}_{i,t} = 0$, once the i-th EV leaves the system, the value of $K_{i,t}$ will be locked until the next returning time slot of the EV, *i.e.*,

$$K_{i,t} = K_{i,t_{il,k}}, \quad \forall t \in \{t_{il,k}, \cdots, t_{ir,k+1} - 1\}, \text{ and } \forall k \in \{1, 2, \cdots\}.$$

By introducing the virtual queues, the constraints (8) and (11) hold if the queues $J_{i,t}$ and $H_{i,t}$ are mean rate stable, respectively [16]. Below we give the definition of mean rate stability of a queue.

Definition: A queue Q_t is mean rate stable if $\lim_{t\to\infty} \frac{\mathbb{E}[|Q_t|]}{t} = 0$.

Unlike $J_{i,t}$ and $H_{i,t}$, since $K_{i,t}$ is re-initialized when $t \in \mathcal{T}_{i,r}$, a new virtual queue is essentially created every time the i-th EV re-joining the system. Therefore, the mean rate stability of $K_{i,t}$ is insufficient for the constraint (14) to hold, and a stronger condition is required. Fortunately, since $K_{i,t}$ is just a shifted version of $s_{i,t}$ from (23), based on Lemma 2, the following result is straightforward.

Lemma 3: For the i-th EV, under the assumption A2, if $K_{i,t} \in [s_{i,\min} - c_i, s_{i,\max} - c_i], \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$, then the constraint (14) holds, *i.e.*, $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[b_{i,t}] = 0$.

Later in Section IV-A, we will show that by our proposed algorithm the boundedness assumption of $K_{i,t}$ in Lemma 3 can be guaranteed.

Define $\mathbf{J}_t \triangleq [J_{1,t}, \cdots, J_{N,t}]$, $\mathbf{H}_t \triangleq [H_{1,t}, \cdots, H_{N,t}]$, $\mathbf{K}_t \triangleq [K_{1,t}, \cdots, K_{N,t}]$, and $\mathbf{\Theta}_t \triangleq [\mathbf{J}_t, \mathbf{H}_t, \mathbf{K}_t]$. Initialize

$J_{i,0} = H_{i,0} = 0$, and $K_{i,0} = s_{i,0} - c_i, \forall i$. Define the Lyapunov function $L(\mathbf{\Theta}_t) \triangleq \frac{1}{2}\sum_{i=1}^N (J_{i,t}^2 + H_{i,t}^2 + K_{i,t}^2)$, and the associated one-slot Lyapunov drift as $\Delta(\mathbf{\Theta}_t) \triangleq \mathbb{E}\left[L(\mathbf{\Theta}_{t+1}) - L(\mathbf{\Theta}_t)|\mathbf{\Theta}_t\right]$. The drift-minus-welfare function is given by $\Delta(\mathbf{\Theta}_t) - V\mathbb{E}\left[\sum_{i=1}^N \omega_i U(z_{i,t}) - e_t|\mathbf{\Theta}_t\right]$, where $V \in (0, V_{\max}]$ is the weight associated with the welfare objective. Hence, the larger V, the more weight is put on the welfare objective.

Furthermore, we assume that for the i-th EV, the conditional expectation of the energy state difference $\Delta_{i,k}$, given the queue backlogs before the EV returns, is zero, *i.e.*,

A3) $\mathbb{E}[\Delta_{i,k}|\mathbf{\Theta}_t] = 0$, for $t = t_{ir,k+1} - 1, \forall k \in \{1, 2, \cdots\}, \forall i$.

Note that A3 is mild, considering the random behavior of each EV due to other activities.

Now we provide an upper bound on the drift-minus-welfare function in the following proposition.

Proposition 1: Under the assumptions A1 and A3, the drift-minus-welfare function is upper-bounded as

$$\begin{aligned}
&\Delta(\mathbf{\Theta}_t) - V\mathbb{E}\left[\sum_{i=1}^N \omega_i U(z_{i,t}) - e_t|\mathbf{\Theta}_t\right] \\
&\le B + \sum_{i=1}^N K_{i,t}\mathbb{E}[b_{i,t}|\mathbf{\Theta}_t] + \sum_{i=1}^N H_{i,t}\mathbb{E}[z_{i,t} - x_{i,t}|\mathbf{\Theta}_t] \\
&+ \sum_{i=1}^N J_{i,t}\mathbb{E}\left[\mathbf{1}_{d,t}C_i(x_{id,t}) + \mathbf{1}_{u,t}C_i(x_{iu,t}) - c_{i,\text{up}}|\mathbf{\Theta}_t\right] \\
&- V\mathbb{E}\left[\sum_{i=1}^N \omega_i U(z_{i,t}) - e_t\Big|\mathbf{\Theta}_t\right], \qquad (24)
\end{aligned}$$

where

$$B \triangleq \frac{1}{2}\sum_{i=1}^N \left[2x_{i,\max}^2 + \Delta_{i,\max}^2 + [c_{i,\text{up}}^2, (c_{i,\max} - c_{i,\text{up}})^2]^+\right]. \qquad (25)$$

Proof: See Appendix C. ■

Adopting the general framework of Lyapunov optimization [16], we now propose the WMRA algorithm by minimizing the upper bound on the drift-minus-welfare function in (24) at each time slot. We will show in Section IV that the proposed algorithm can lead to a guaranteed performance.

The minimization problem is equivalent to the following decoupled sub-problems with respect to $\mathbf{z}_t$, $\mathbf{x}_{d,t}$, and $\mathbf{x}_{u,t}$, separately. Denote the solutions produced by WMRA as $\tilde{\mathbf{z}}_t \triangleq [\tilde{z}_{1,t}, \cdots, \tilde{z}_{N,t}]$, $\tilde{\mathbf{x}}_{d,t} \triangleq [\tilde{x}_{1d,t}, \cdots, \tilde{x}_{Nd,t}]$, and $\tilde{\mathbf{x}}_{u,t} \triangleq [\tilde{x}_{1u,t}, \cdots, \tilde{x}_{Nu,t}]$, respectively. Specifically, we obtain $\tilde{z}_{i,t}, \forall i$, by solving **(a)**:

$$\textbf{(a):} \min_{z_{i,t}} \quad H_{i,t}z_{i,t} - \omega_i V U(z_{i,t}) \quad \text{s.t.} \ 0 \le z_{i,t} \le x_{i,\max}.$$

For $G_t > 0$, we obtain $\tilde{\mathbf{x}}_{d,t}$ by solving **(b1)**:

$$\begin{aligned}
\textbf{(b1):} \min_{\mathbf{x}_{d,t}} \quad & V e_{s,t}\big(G_t - \sum_{i=1}^N x_{id,t}\big) - \sum_{i=1}^N H_{i,t}x_{id,t} \\
& + \sum_{i=1}^N J_{i,t}C_i(x_{id,t}) + \sum_{i=1}^N K_{i,t}x_{id,t} \\
\text{s.t.} \quad & 0 \le x_{id,t} \le \mathbf{1}_{i,t}x_{i,\max}, \quad \sum_{i=1}^N x_{id,t} \le G_t.
\end{aligned}$$

For $G_t < 0$, we obtain $\tilde{\mathbf{x}}_{u,t}$ by solving **(b2)**:

$$\begin{aligned}
\textbf{(b2):} \min_{\mathbf{x}_{u,t}} \quad & V e_{d,t}\big(|G_t| - \sum_{i=1}^N x_{iu,t}\big) - \sum_{i=1}^N H_{i,t}x_{iu,t} \\
& + \sum_{i=1}^N J_{i,t}C_i(x_{iu,t}) - \sum_{i=1}^N K_{i,t}x_{iu,t} \\
\text{s.t.} \quad & 0 \le x_{iu,t} \le \mathbf{1}_{i,t}x_{i,\max}, \quad \sum_{i=1}^N x_{iu,t} \le |G_t|.
\end{aligned}$$

Algorithm 1 Welfare-maximizing regulation allocation (WMRA) algorithm.

1: The aggregator initializes the virtual queue vector $\mathbf{\Theta}_0$, and re-initialize $K_{i,t} = s_{i,t} - c_i$ for $t \in \mathcal{T}_{i,r}, \forall i$.

2: At the beginning of each time slot t, the aggregator performs the following steps sequentially.

(2a) Observe $G_t, e_{s,t}, e_{d,t}, \mathbf{1}_t$, $\mathbf{J}_t$, $\mathbf{H}_t$, and $\mathbf{K}_t$.

(2b) Solve (**a**) and record an optimal solution $\tilde{\mathbf{z}}_t$. If $G_t > 0$, solve (**b1**) and record an optimal solution $\tilde{\mathbf{x}}_{d,t}$. If $G_t < 0$, solve (**b2**) and record an optimal solution $\tilde{\mathbf{x}}_{u,t}$. Allocate the regulation amounts among EVs based on $\tilde{\mathbf{x}}_{d,t}$ and $\tilde{\mathbf{x}}_{u,t}$. If $\sum_{i=1}^N \tilde{x}_{id,t} < G_t$ or $\sum_{i=1}^N \tilde{x}_{iu,t} < |G_t|$, clear the imbalance using external energy sources.

(2c) Update the virtual queues $J_{i,t}$, $H_{i,t}$, and $K_{i,t}, \forall i$, based on (19), (20), and (21b), respectively.

Note that **(a)**, **(b1)**, and **(b2)** are all convex problems, so they can be efficiently solved using standard methods such as the interior point method [22]. We summarize WMRA in Algorithm 1. Note from Steps (2b) and (2c) that, the solutions of **(a)** and **(b1)** (or **(b2)**) affect each other over multiple time slots through the update of $H_{i,t}, \forall i$. To perform WMRA, no statistical information of the system is needed, which makes the algorithm easy to implement.

IV. Performance Analysis

In this section, we characterize the performance of WMRA with respect to our original problem **P1**.

A. Properties of WMRA Algorithm

We now show that WMRA can ensure the boundedness of each EV's energy state. The following lemma characterizes sufficient conditions under which the solution of $\tilde{x}_{id,t}$ and $\tilde{x}_{iu,t}$ under WMRA is zero.

Lemma 4: Under the WMRA algorithm, for any $t \in \mathcal{T}_{i,p}$,

1) for $G_t > 0$, if $K_{i,t} > x_{i,\max} + V(\omega_i\mu + e_{\max})$, then $\tilde{x}_{id,t} = 0$, which means that $K_{i,t+1}$ cannot be increased at the next time slot; and
2) for $G_t < 0$, if $K_{i,t} < -x_{i,\max} - V(\omega_i\mu + e_{\max})$, then $\tilde{x}_{iu,t} = 0$, which means that $K_{i,t+1}$ cannot be decreased at the next time slot.

Proof: See Appendix D. ■

Since Lemma 4 on the other hand provides conditions under which queue backlog $K_{i,t}$ can no longer increase or decrease, using Lemma 4, we can prove the boundedness of $K_{i,t}$ below.

Lemma 5: Under the WMRA algorithm, queue backlog $K_{i,t}$ associated with the i-th EV is bounded within $[s_{i,\min} - c_i, s_{i,\max} - c_i], \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$.

Proof: See Appendix E. ∎

In the proof of Lemma 5, we remark on the specific designs of c_i and $V_{\max}$, which are to ensure the boundedness of $K_{i,t}$ within a shifted preferred energy range.

From Lemma 5, the boundedness condition of $K_{i,t}$ in Lemma 3 is now satisfied, therefore the conclusion there is true under WMRA. Since $K_{i,t} = s_{i,t} - c_i, \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$, using Lemma 5, the following lemma is straightforward.

Lemma 6: Under the WMRA algorithm, the energy state of the i-th EV is bounded within $[s_{i,\min}, s_{i,\max}], \forall t \in \mathcal{T}_{i,p} \cup \mathcal{T}_{i,l}$.

From Lemma 6, the constraints (4) and (5) in **P2** are met under WMRA.

B. Optimality of WMRA Algorithm

In this subsection, we investigate the optimality of WMRA by considering EVs with both predictable and random dynamics, which are described below.

1) EVs with predictable dynamics: Predictable dynamics could happen when each EV joins and leaves the aggregator-EVs system regularly (e.g. from 9am to 12pm in the morning, then from 2pm to 6pm in the afternoon). Therefore, the leaving and returning time slots of each EV can be predicted by the aggregator. In other words, the aggregator is aware of the realization of $\mathbf{1}_t, \forall t$ in advance. In this case, the random system state at time slot t is defined as $A_t \triangleq (G_t, e_{s,t}, e_{d,t})$. A specific case of EVs with predictable dynamics is static EVs, *i.e.*, $\mathbf{1}_{i,t} = 1, \forall i, t$[2].
2) EVs with random dynamics: If the EVs do not participate in the aggregator-EVs system regularly, then the aggregator cannot predict their dynamics beforehand, and therefore has to observe $\mathbf{1}_t$ every time slot. In this case, the random system state at time slot t is defined as $A_t \triangleq (G_t, e_{s,t}, e_{d,t}, \mathbf{1}_t)$.

Note that the WMRA algorithm is the same under both of the above cases. The only difference between them is that, in the optimization problem **P3**, the expectations are taken over different randomness of the system state. The performance under WMRA as compared to the optimal solution of **P1** is given in the following theorem, which applies to both predictable and random dynamics.

Theorem 1: Under the assumptions A1, A2, and A3, given the system state A_t is i.i.d. over time,

1) $(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t})$ is feasible for **P1**, *i.e.*, it satisfies (4)–(8).
2) $f_1(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t}) \geq f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}) - \frac{B}{V}$, where B is defined in (25) and $V \in (0, V_{\max}]$.

Proof: See Appendix F. ∎

Remarks: From Theorem 1, the welfare performance of WMRA is away from the optimum by $O(1/V)$. Hence, the larger V, the better the performance of WMRA. However, in practice, due to the boundedness condition of EV's battery capacity, V cannot be arbitrarily large and is upper bounded by $V_{\max}$, which is defined in (22). Note that $V_{\max}$ increases with the smallest span of the EVs' preferred battery capacity ranges, *i.e.*, $\min_{1 \leq i \leq N}\{s_{i,\max} - s_{i,\min}\}$. Therefore, roughly speaking, the performance gap between WMRA and the optimum decreases as the smallest battery capacity increases. Asymptotically, as the EVs' battery capacities go to infinity, WMRA would achieve exactly the optimum.

In Theorem 1, the i.i.d. condition of A_t can be relaxed to Markovian, and a similar performance bound can be obtained. In particular, this relaxed condition can accommodate the case where G_t is Markovian and has a ramp rate constraint ($|G_t - G_{t-1}| \leq$ ramp rate $\times \Delta t$), by properly designing the transition probability matrix of G_t.

Theorem 2: Under the assumptions A1, A2, and A3, given that the system state A_t evolves based on a finite state irreducible and aperiodic Markov chain,

1) $(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t})$ is feasible for **P1**, *i.e.*, it satisfies (4)–(8).
2) $f_1(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t}) \geq f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}) - O(1/V)$, where $V \in (0, V_{\max}]$.

Proof: The above results can be proved by expanding the proof of Theorem 1 using a multi-slot drift technique [16]. We omit the proof here for brevity. ∎

V. Simulation Results

Besides the analytical performance bound derived above, we are further interested in evaluating WMRA in example numerical settings. Towards this goal, we have developed an aggregator-EVs model in Matlab and compared WMRA with a greedy algorithm.

Suppose that the aggregator is connected with $N = 100$ EVs, evenly split into Type I (based on the 2012 Ford Focus Electric) and Type II (based on the Tesla Model S base model). The parameters of Type I and Type II EVs are summarized in Table I [21], [23]. The maximum regulation amount $x_{i,\max}$ can be derived by multiplying the maximum charging/discharging rate with the regulation interval Δt. In current practice, Δt is of the order of seconds. For example, for PJM, $\Delta t = 2$ seconds [24], and for NYISO, $\Delta t = 6$ seconds [25]. In simulations, we set $\Delta t = 5$ seconds as an example.

Consider that the system state $A_t = (G_t, e_{s,t}, e_{d,t}, \mathbf{1}_t)$ follows a finite state irreducible and aperiodic Markov chain. For the regulation signal G_t, we ignore the ramp rate constraint in our simulations. At each time slot, we draw a sample of G_t from a uniformly distributed set $\{-1.15, -1.15+\Delta_1, -1.15+2\Delta_1, \cdots, 1.15\}$ (kWh) with the cardinality 200, where 1.15 kWh is the maximum allowed regulation amount at each time slot if all N EVs are in the system. The unit costs of the external sources, $e_{s,t}$ and $e_{d,t}$, are drawn uniformly from a discrete set $\{0.1, 0.1+\Delta_2, 0.1+2\Delta_2, \cdots, 0.12\}$ (dollars/kWh) with the cardinality 200. The lower bound 0.1 dollars/kWh and the upper bound 0.12 dollars/kWh correspond to the mid-peak and the on-peak electricity prices in Ontario, respectively [26]. The dynamics of each EV is described by the indicator random variable $\mathbf{1}_{i,t}$, which represents whether the i-th EV is in the system at time slot t. In particular, we assume that

[2]In this work, we focus on the investigation of non-static EVs. For static EVs, the interested reader is referred to [17] for details.

TABLE II
PARAMETERS OF TYPE I AND TYPE II EVS

	Type I EV	Type II EV
Energy capacity $s_{i,\text{cap}}$ (kWh)	23	40
Maximum charging/discharging rate (kW)	6.6	10

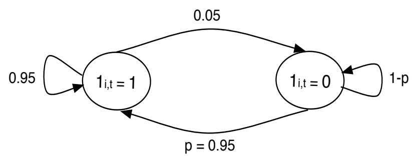


Fig. 1. Transition probabilities of $\mathbf{1}_{i,t}, \forall i$.

$\mathbf{1}_{i,t}$ follows a two-state Markov chain as shown in Fig. 1. The state transition probability $p \triangleq \mathbb{P}(0 \to 1)$ is set to be 0.95 by default.

For the i-th EV, the $(k+1)$-th returning energy state $s_{i,t_{ir,k+1}}$ is drawn uniformly from the interval $[s_{i,t_{il,k}} - \Delta_3, s_{i,t_{il,k}} + \Delta_3]$, where $s_{i,t_{il,k}}$ is the k-th leaving energy state of the i-th EV and $\Delta_3 = 5\% s_{i,\text{cap}}$[3]. We set the minimum preferred energy state $s_{i,\min} = 0.1 s_{i,\text{cap}}$, and the maximum preferred energy state $s_{i,\max} = 0.9 s_{i,\text{cap}}$ except otherwise mentioned. In the objective function of **P1**, we set $U(x) = \log(1+x)$ and $\omega_i = 1, \forall i$. Since the degradation cost function $C_i(\cdot)$ is proprietary and unavailable, in simulations, we set $C_i(x) = x^2$ as an example. The upper bound $c_{i,\text{up}}$ is set to be $x_{i,\max}^2/4$.

To allocate the requested regulation amount, we apply WMRA in Algorithm 1 at each time slot. For comparison, we consider a greedy algorithm which only optimizes the system performance at the current time slot. The regulation allocation at each time slot is determined by the following optimization problem.

$$\begin{aligned} \max_{\mathbf{x}_{d,t}, \mathbf{x}_{u,t}} \quad & \left(\sum_{i=1}^{N} \omega_i U(x_{i,t}) \right) - e_t \\ \text{s.t.} \quad & (4), (5), (6), (7), \text{ and} \\ & \mathbf{1}_{d,t} C_i(x_{id,t}) + \mathbf{1}_{u,t} C_i(x_{iu,t}) \le c_{i,\text{up}}, \forall i. \end{aligned}$$

The above problem is a convex optimization problem, and we use the standard solver in MATLAB to obtain its solution.

In Figs. 2 and 3, we compare the performance of WMRA with $V = V_{\max}$ and the performance of the greedy algorithm. From Fig. 2, with $s_{i,\max} = 0.9 s_{i,\text{cap}}$, WMRA is uniformly superior to the greedy algorithm over all time slots, with the advantage about 40%. In Fig. 3, we set the transition probability p to be 0.95 and 0.05, and vary $s_{i,\max}$ from $0.3 s_{i,\text{cap}}$ to $0.9 s_{i,\text{cap}}$. For $p = 0.95$, the observations are as follows. First, WMRA uniformly outperforms the greedy algorithm over different values of $s_{i,\max}$. Second, as $s_{i,\max}$ increases, the social welfare under WMRA slightly rises. This is because increasing $s_{i,\max}$ effectively increases $V_{\max}$, which improves the performance of WMRA. This observation is also consistent with the remarks after Theorem 1. In contrast, the greedy algorithm cannot benefit from the expanded energy range. For $p = 0.05$, the trends of the curves resemble those

[3]We ensure that all returning energy states are within the preferred range $[s_{i,\min}, s_{i,\max}]$ by ignoring unqualified samples.

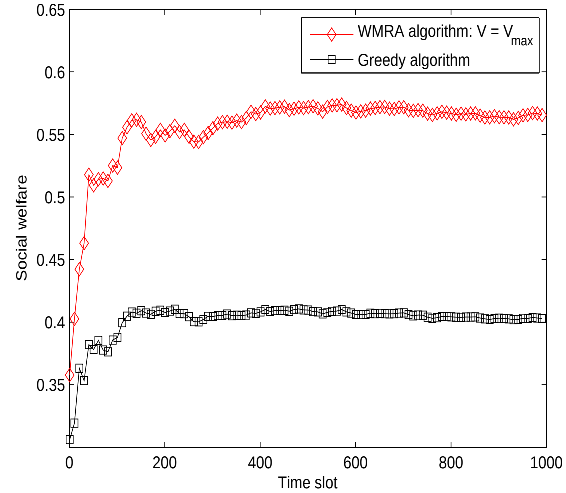


Fig. 2. Time-averaged social welfare with $V = V_{\max}$.

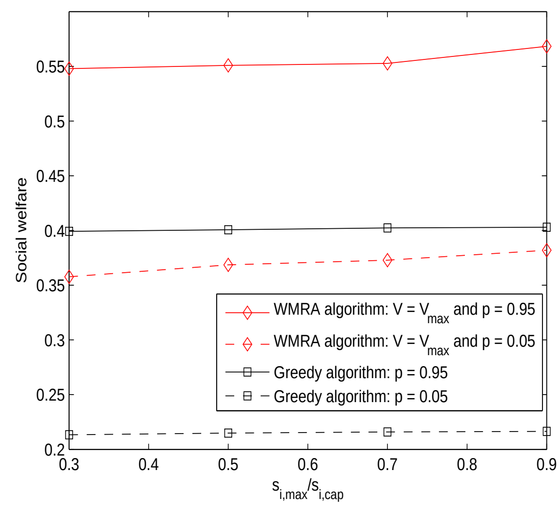


Fig. 3. Time-averaged social welfare with various $s_{i,\max}$ and $V = V_{\max}$.

for $p = 0.95$, but the social welfare of both algorithms drops. This is because when p is decreased, roughly speaking, there are fewer EVs in the system for the regulation service. Hence, to provide the requested regulation amount, the aggregator more relies on the expensive external energy sources, which leads to a decreased social welfare.

In Fig. 4, we show the performance of WMRA with the value of V ranging from $0.2V_{\max}$ to $5V_{\max}$, and compare it with the performance of the greedy algorithm. For WMRA, as expected, the social welfare grows with the value of V; also, the growing rate slows down when V gets larger. Moreover, we observe that WMRA outperforms the greedy algorithm even with $V = 0.2V_{\max}$.

In Lemma 6, the energy state of each EV is shown to be restricted within $[s_{i,\min}, s_{i,\max}]$ when $V \in (0, V_{\max}]$. In Fig. 5, for V being $V_{\max}, 2V_{\max}$, and $5V_{\max}$, we show the evolution of a Type I EV's energy state under WMRA. We see that, when $V = V_{\max}$, the energy state is always within the preferred range. In contrast, when $V = 2V_{\max}$ or $5V_{\max}$, the associated energy state can exceed the preferred range from

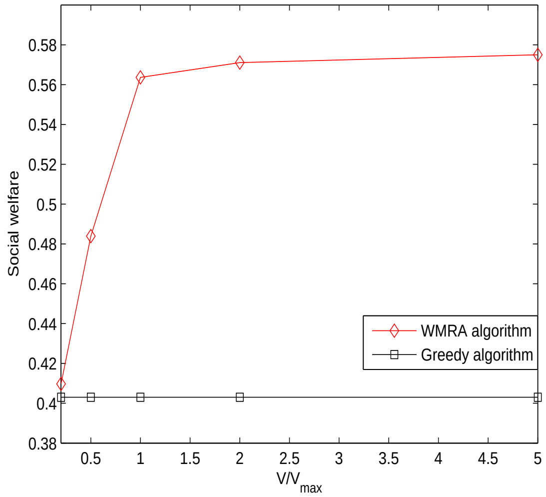


Fig. 4. Time-averaged social welfare with various values of V.

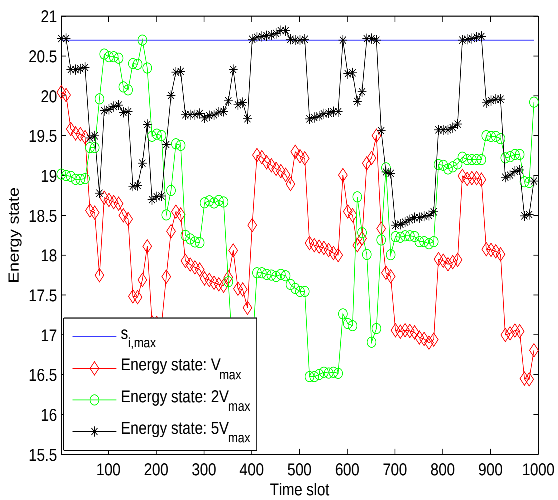


Fig. 5. Sample path of a Type I EV's energy state with $V = [1, 2, 5]V_{\max}$.

time to time. Furthermore, the larger V the more frequently such violation happens. Therefore, the observations in Figs. 4 and 5 demonstrate the significance of $V_{\max}$ in achieving the maximum social welfare under WMRA considering the constraint of EV's preferred energy range.

VI. Conclusion

We studied a practical model of a dynamic aggregator-EVs system providing regulation service to a power grid. We formulated the regulation allocation optimization as a long-term time-averaged social welfare maximization problem. Our formulation accounts for random system dynamics, battery constraints, the costs for battery degradation and external energy sources, and especially, the dynamics of EVs. Adopting a general Lyapunov optimization framework, we developed a real-time WMRA algorithm for the aggregator to fairly allocate the regulation amount among EVs. The algorithm does not require any knowledge of the statistics of the system state. We were able to bound the performance of WMRA to that under the optimal solution, and showed that the performance of WMRA is asymptotically optimal as EVs' battery capacities go to infinity. Simulation demonstrated that WMRA offers substantial performance gains over a greedy algorithm that maximizes per-slot social welfare objective.

Appendix A
Proof of Lemma 1

It is easy to see that $(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t})$ is feasible for **P1**. To show that $(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}, \bar{\mathbf{z}}^{\text{opt}}_t)$ is feasible for **P2**, it suffices to show that $\bar{\mathbf{z}}^{\text{opt}}_t$ satisfies (10) and (11). Using the definition of $\bar{z}^{\text{opt}}_{i,t}$, (11) naturally holds. Also, since $x^{\text{opt}}_{i,t}$ lies in $[0, x_{i,\max}]$, which is a closed interval, (10) holds.

We claim that

$$\begin{aligned} f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}) \quad &= f_2(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}, \bar{\mathbf{z}}^{\text{opt}}_t) \\ &\le f_2(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t}, \mathbf{z}^*_t) \\ &\le f_1(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t}) \\ &\le f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}). \end{aligned} \tag{26}$$

Using the definition of $\bar{z}^{\text{opt}}_{i,t}$ in $f_2(\cdot)$, the first equality holds. The first and the third inequalities hold since $(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t}, \mathbf{z}^*_t)$ and $(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t})$ are optimal for $f_2(\cdot)$ and $f_1(\cdot)$, respectively. The second inequality is derived using Jensen's inequality for concave functions. Since (26) is satisfied with equality, all inequalities in (26) turn into equalities, which indicates the equivalence of **P1** and **P2**.

Appendix B
Proof of Lemma 2

Let T be large enough. For the i-th EV, decompose the total effective charging/discharging amount within T time slots as

$$\sum_{t=0}^{T-1} b_{i,t} = \sum_{t=0}^{t_{il,k^*}-1} b_{i,t} + \sum_{t=t_{il,k^*}}^{T-1} b_{i,t}, \tag{27}$$

where $k^* \triangleq \max\{k : t_{il,k} \le (T-1), k \in \{1, 2, \cdots\}\}$ is defined to be the total number of the leaving times of the i-th EV up to time slot $T-1$. On the right hand side of (27), the first term corresponds to the total effective charging/discharging amount before the last leaving time, and the second term corresponds to the rest of the total effective charging/discharging amount. Using the decomposition in (27), to show (14), it suffices to show that the two limits $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}[\sum_{t=0}^{t_{il,k^*}-1} b_{i,t}]$ and $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}[\sum_{t=t_{il,k^*}}^{T-1} b_{i,t}]$ are both equal to zero.

First consider the second limit. For the i-th EV, if there is no return between t_{il,k^*} and $T-1$, then $\sum_{t=t_{il,k^*}}^{T-1} b_{i,t} = 0$ and thus $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}[\sum_{t=t_{il,k^*}}^{T-1} b_{i,t}] = 0$. Or, if there is one return, then $\sum_{t=t_{il,k^*}}^{T-1} b_{i,t} = s_{i,T} - s_{i,t_{ir,k^*+1}}$. Using the boundedness condition of $s_{i,t}$, we have $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}[\sum_{t=t_{il,k^*}}^{T-1} b_{i,t}] = 0$. Together, the second limit is zero.

Next we show that the first limit is also zero. Based on the energy state evolution in (1), there is

$$\sum_{t=0}^{t_{il,k^*}-1} b_{i,t} = \sum_{k=1}^{k^*} s_{i,t_{il,k}} - \sum_{k=1}^{k^*} s_{i,t_{ir,k}}$$

$$= s_{i,t_{il,k^*}} - s_{i,t_{ir,1}} - \sum_{k=1}^{k^*-1} \Delta_{i,k}. \quad (28)$$

Taking expectations of both sides of (28), dividing them by T, then taking limits gives

$$\lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{t_{il,k^*}-1} b_{i,t}\right] = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[s_{i,t_{il,k^*}} - s_{i,t_{ir,1}}\right] - \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=1}^{k^*-1} \Delta_{i,k}\right] = 0,$$

where the last equality is derived by the boundedness of $s_{i,t}$ and the assumption A2. This completes the proof.

Appendix C
Proof of Proposition 1

Based on the definition of $L(\mathbf{\Theta}_t)$, the difference

$$L(\mathbf{\Theta}_{t+1}) - L(\mathbf{\Theta}_t) = \frac{1}{2}\sum_{i=1}^{N} H_{i,t+1}^2 + J_{i,t+1}^2 + K_{i,t+1}^2 - H_{i,t}^2 - J_{i,t}^2 - K_{i,t}^2. \quad (29)$$

In (29), $H_{i,t+1}^2 - H_{i,t}^2$ and $J_{i,t+1}^2 - J_{i,t}^2$ can be upper bounded as follows.

$$H_{i,t+1}^2 - H_{i,t}^2 \le 2H_{i,t}(z_{i,t} - x_{i,t}) + x_{i,\max}^2 \quad (30)$$

$$J_{i,t+1}^2 - J_{i,t}^2 \le 2J_{i,t}[\mathbf{1}_{d,t}C_i(x_{id,t}) + \mathbf{1}_{u,t}C_i(x_{iu,t}) - c_{i,\text{up}}] + [c_{i,\text{up}}^2, (c_{i,\max} - c_{i,\text{up}})^2]^+. \quad (31)$$

Taking conditional expectations for both sides in (30) and (31), we have

$$\mathbb{E}[H_{i,t+1}^2 - H_{i,t}^2|\mathbf{\Theta}_t] \le 2H_{i,t}\mathbb{E}[z_{i,t} - x_{i,t}|\mathbf{\Theta}_t] + x_{i,\max}^2 \quad (32)$$

$$\mathbb{E}[J_{i,t+1}^2 - J_{i,t}^2|\mathbf{\Theta}_t] \le 2J_{i,t}\mathbb{E}[\mathbf{1}_{d,t}C_i(x_{id,t}) + \mathbf{1}_{u,t}C_i(x_{iu,t}) - c_{i,\text{up}}|\mathbf{\Theta}_t] + [c_{i,\text{up}}^2, (c_{i,\max} - c_{i,\text{up}})^2]^+. \quad (33)$$

Now consider $K_{i,t+1}^2 - K_{i,t}^2$. When $\mathbf{1}_{i,t} = 1$, we have $K_{i,t+1} = K_{i,t} + b_{i,t}$ and thus

$$K_{i,t+1}^2 - K_{i,t}^2 \le 2K_{i,t}b_{i,t} + x_{i,\max}^2. \quad (34)$$

When $\mathbf{1}_{i,t} = 0$, we have $b_{i,t} = 0$ and there are two cases. First, for $t \in \{t_{il,k}, t_{il,k}+1, \cdots, t_{ir,k+1}-2\}, \forall k \in \{1, 2, \cdots\}$, there is $K_{i,t+1} = K_{i,t}$. So, we can express

$$K_{i,t+1}^2 - K_{i,t}^2 = 2K_{i,t}b_{i,t}. \quad (35)$$

Second, for $t = t_{ir,k+1} - 1, \forall k \in \{1, 2, \cdots\}$, we have $K_{i,t} = s_{i,t_{il,k}} - c_i$ and $K_{i,t+1} = K_{i,t} + \Delta_{i,k}$. Hence, by the assumption A1,

$$K_{i,t+1}^2 - K_{i,t}^2 \le 2K_{i,t}\Delta_{i,k} + \Delta_{i,\max}^2. \quad (36)$$

Using the assumption A3, from (34), (35), and (36), we have

$$\mathbb{E}[K_{i,t+1}^2 - K_{i,t}^2|\mathbf{\Theta}_t] \le 2K_{i,t}\mathbb{E}[b_{i,t}|\mathbf{\Theta}_t] + x_{i,\max}^2 + \Delta_{i,\max}^2. \quad (37)$$

Using the definition of $\Delta(\mathbf{\Theta}_t)$ and the upper bounds in (32), (33), and (37), we can derive the upper bound on the drift-minus-welfare function in Proposition 1.

Appendix D
Proof of Lemma 4

We need the following lemma.

Lemma 7: Under the WMRA algorithm, queue backlog $H_{i,t}$ associated with the i-th EV is upper bounded as follows:

$$H_{i,t} \le V\omega_i\mu + x_{i,\max}.$$

Proof: This can be shown using a similar method as in [16], and the technical condition (9) is needed. ■

1) Consider $G_t > 0$. Suppose that when $K_{i,t} > x_{i,\max} + V(\omega_i\mu + e_{\max})$, one optimal solution under WMRA is $\tilde{\mathbf{x}}_{d,t}$ with $\tilde{x}_{id,t} > 0$. Then we show that we can find another solution with $\tilde{x}_{jd,t}, \forall j \ne i$ and $\tilde{x}_{id,t} = 0$ resulting in a strictly smaller objective value, which is a contradiction.

Using the objective function of (**b1**), this is equivalent to showing that

$$Ve_{s,t}\left(G_t - \sum_{j=1}^{N}\tilde{x}_{jd,t}\right) - \sum_{j=1}^{N} H_{j,t}\tilde{x}_{jd,t} + \sum_{j=1}^{N} J_{j,t}C_j(\tilde{x}_{jd,t}) + \sum_{j=1}^{N} K_{j,t}\tilde{x}_{jd,t}$$
$$> Ve_{s,t}\left(G_t - \sum_{j=1}^{N}\tilde{x}_{jd,t} + \tilde{x}_{id,t}\right) - \sum_{j\ne i} H_{j,t}\tilde{x}_{jd,t} + \sum_{j\ne i} J_{j,t}C_j(\tilde{x}_{jd,t}) + \sum_{j\ne i} K_{j,t}\tilde{x}_{jd,t},$$

which is equivalent to

$$-H_{i,t}\tilde{x}_{id,t} + J_{i,t}C_i(\tilde{x}_{id,t}) + K_{i,t}\tilde{x}_{id,t} > Ve_{s,t}\tilde{x}_{id,t}. \quad (38)$$

Since $J_iC_i(\tilde{x}_{id,t}) \ge 0$, from (38), it suffices to show that

$$(K_{i,t} - H_{i,t} - Ve_{s,t})\tilde{x}_{id,t} > 0. \quad (39)$$

Since $\tilde{x}_{id,t} > 0$, (39) is true by using the assumption that $K_{i,t} > x_{i,\max} + V(\omega_i\mu + e_{\max})$ and Lemma 7 in which $H_{i,t}$ is upper bounded.

2) Consider $G_t < 0$. Suppose that when $K_{i,t} < -x_{i,\max} - V(\omega_i\mu + e_{\max})$, one optimal solution under WMRA is $\tilde{\mathbf{x}}_{u,t}$ with $\tilde{x}_{iu,t} > 0$. Then there is a contradiction since we can construct another solution with $\tilde{x}_{ju,t}, \forall j \ne i$ and $\tilde{x}_{iu,t} = 0$ which results in a strictly smaller objective value. The proof is similar as that in 1) and is omitted here.

Appendix E
Proof of Lemma 5

Consider the set $\{t_{ir,k}, t_{ir,k}+1, \cdots, t_{il,k}\}$ for any $k \in \{1, 2, \cdots\}$. We show below that $K_{i,t}$ is bounded for any t in such set by induction.

First consider the upper bound. For the time slot $t_{ir,k}$, based on (21) and $s_{i,t_{ir,k}} \le s_{i,\max}$, there is $K_{i,t_{ir,k}} \le s_{i,\max} - c_i$. Assume that the upper bound holds for time slot t and consider the following two cases of $K_{i,t}$.

Case 1: $x_{i,\max} + V(\omega_i\mu + e_{\max}) < K_{i,t} \le s_{i,\max} - c_i$ (We can check that $x_{i,\max} + V(\omega_i\mu + e_{\max}) < s_{i,\max} - c_i$ since $V \le V_{\max}$). For $G_t > 0$, from Lemma 4 1), there is $\tilde{x}_{id,t} = 0$.

Therefore, $K_{i,t+1} = K_{i,t} \le s_{i,\max} - c_i$. For $G_t < 0$, we have $K_{i,t+1} = K_{i,t} - x_{iu,t} \le K_{i,t} \le s_{i,\max} - c_i$.

Case 2: $K_{i,t} \le x_{i,\max} + V(\omega_i\mu + e_{\max})$. From (21), $K_{i,t+1} \le 2x_{i,\max} + V(\omega_i\mu + e_{\max}) \le s_{i,\max} - c_i$, where the last inequality holds since $V \le V_{\max}$.

Now look at the lower bound. For the time slot $t_{ir,k}$, based on (21) and $s_{i,t_{ir,k}} \ge s_{i,\min}$, there is $K_{i,t_{ir,k}} \ge s_{i,\min} - c_i$. Assume that the lower bound holds for time slot t and consider the following two cases of $K_{i,t}$.

Case 1′: $s_{i,\min} - c_i \le K_{i,t} < -x_{i,\max} - V(\omega_i\mu + e_{\max})$ (We can check that $s_{i,\min} - c_i < -x_{i,\max} - V(\omega_i\mu + e_{\max})$ since $x_{i,\max} > 0$). For $G_t < 0$, from Lemma 4 2), there is $\tilde{x}_{iu,t} = 0$. Therefore, $K_{i,t+1} = K_{i,t} \ge s_{i,\min} - c_i$, For $G_t > 0$, we have $K_{i,t+1} = K_{i,t} + x_{id,t} \ge K_{i,t} \ge s_{i,\min} - c_i$.

Case 2′: $K_{i,t} \ge -x_{i,\max} - V(\omega_i\mu + e_{\max})$. From (21), $K_{i,t+1} \ge -2x_{i,\max} - V(\omega_i\mu + e_{\max})$, which is exactly $s_{i,\min} - c_i$.

Remarks: To track the energy state $s_{i,t}$, in principle, the shift c_i can be any number. However, to make the proof in Case 2′ work, c_i is lower bounded, *i.e.*, should satisfy $c_i = s_{i,\min} + 2x_{i,\max} + V(\omega_i\mu + e_{\max}) + \epsilon_1$ where $\epsilon_1 \ge 0$. For the design of $V_{\max}$, to make the proof in Case 1 work, it is sufficient to let $V_{\max} = \min_{1\le i\le N}\left\{\frac{s_{i,\max} - s_{i,\min} - 3x_{i,\max} - \epsilon_1 - \epsilon_2}{2(\omega_i\mu + e_{\max})}\right\}$ where $\epsilon_2 > 0$. Based on the proof in Case 2, ϵ_1 and ϵ_2 are further determined as 0 and $x_{i,\max}$, respectively, to make $V_{\max}$ as large as possible.

Appendix F
Proof of Theorem 1

We first give the following fact, which is a direct consequence of the results in [16].

Lemma 8: There exists a stationary randomized regulation allocation solution $(\mathbf{x}^s_{d,t}, \mathbf{x}^s_{u,t})$ that only depends on the system state A_t, and there are

$$\mathbb{E}[x^s_{i,t}] = z^s_i, \forall i, \text{ for some } z^s_i \in [0, x_{i,\max}], \tag{40}$$

$$\mathbb{E}[e^s_t] - \sum_{i=1}^N \omega_i U(z^s_i) \le -f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t), \tag{41}$$

$$\mathbb{E}[\mathbf{1}_{d,t} C_i(x^s_{id,t}) + \mathbf{1}_{u,t} C_i(x^s_{iu,t})] \le c_{i,\text{up}}, \forall i, \text{ and} \tag{42}$$

$$\mathbb{E}[b^s_{i,t}] = 0, \forall i, \tag{43}$$

where the expectations are taken over the randomness of the system and the randomness of $(\mathbf{x}^s_{d,t}, \mathbf{x}^s_{u,t})$, and $(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t)$ is an optimal solution for **P3**.

1) For brevity, define $W_t \triangleq \left(\sum_{i=1}^N \omega_i U(z_{i,t})\right) - e_t$. Since WMRA minimizes the upper bound in (24), plug $(\mathbf{x}^s_{d,t}, \mathbf{x}^s_{u,t})$ on the right hand side of (24) together with $z_{i,t} = z^s_i, \forall t$, we have

$$\Delta(\boldsymbol{\Theta}_t) - V\mathbb{E}\left[\tilde{W}_t | \boldsymbol{\Theta}_t\right] \le B - V f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t), \tag{44}$$

where (40), (41), (42), and (43) are used. Since $\tilde{W}_t \le \sum_{i=1}^N \omega_i U(x_{i,\max})$, from (44),

$$\Delta(\boldsymbol{\Theta}_t) \le D \triangleq B + V\left(\sum_{i=1}^N \omega_i U(x_{i,\max}) - f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t)\right).$$

Using Theorem 4.1 in [16], $\mathbb{E}[|H_{i,t}|]$ and $\mathbb{E}[|J_{i,t}|]$ are upper bounded by $\sqrt{2tD + 2\mathbb{E}[L(\boldsymbol{\Theta}_0)]}, \forall t$. Hence, the virtual queues $H_{i,t}$ and $J_{i,t}$ are mean rate stable and the following limit constraints hold.

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[\tilde{z}_{i,t}] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[\tilde{x}_{i,t}], \forall i, \tag{45}$$

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left[\mathbf{1}_{d,t} C_i(\tilde{x}_{id,t}) + \mathbf{1}_{u,t} C_i(\tilde{x}_{iu,t})\right] \le c_{i,\text{up}}, \forall i.$$

Since $s_{i,t}$ is bounded under WMRA by Lemma 6, using Lemma 2, we have $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[\tilde{b}_{i,t}] = 0, \forall i$. In addition, note that $(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t})$ is derived under the constraints of the optimization problems (**a**), (**b1**), and (**b2**). Therefore, we have that $(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t})$ is feasible for **P3**, **P2**, and **P1**.

2) Taking expectations of both sides of (44) and summing over $t \in \{0, 1, \cdots, T-1\}$ for some $T > 1$, we have

$$\begin{aligned}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{W}_t] &\ge \frac{\mathbb{E}\left[L(\boldsymbol{\Theta}_T) - L(\boldsymbol{\Theta}_0)\right]}{VT} + f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t) - B/V \\ &\ge f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t) - B/V - \mathbb{E}[L(\boldsymbol{\Theta}_0)]/VT, \end{aligned} \tag{46}$$

where (46) holds since $L(\boldsymbol{\Theta}_T)$ is non-negative. Also,

$$\begin{aligned}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{W}_t] &= \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\left(\sum_{i=1}^N \omega_i U(\tilde{z}_{i,t})\right) - \tilde{e}_t\right] \\ &\le \sum_{i=1}^N \omega_i U\left(\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{z}_{i,t}]\right) - \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{e}_t], \end{aligned} \tag{47}$$

where the inequality in (47) is derived using Jensen's inequality for concave functions. Combining (46) and (47) and taking limits on both sides, there is

$$\sum_{i=1}^N \omega_i U\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{z}_{i,t}]\right) - \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{e}_t]$$

$$\ge f_2(\hat{\mathbf{x}}_{d,t}, \hat{\mathbf{x}}_{u,t}, \hat{\mathbf{z}}_t) - B/V \tag{48}$$

$$\ge f_2(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t}, \mathbf{z}^*_t) - B/V \tag{49}$$

$$= f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}) - B/V, \tag{50}$$

where $(\mathbf{x}^*_{d,t}, \mathbf{x}^*_{u,t}, \mathbf{z}^*_t)$ and $(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t})$ are defined in Section III-A, (48) holds since $\mathbb{E}[L(\boldsymbol{\Theta}_0)]$ is bounded, (49) holds since the feasible set of the optimization variables is enlarged from **P2** to **P3**, and (50) is true due to Lemma 1.

Rewrite the objective function of **P1** under WMRA, *i.e.*, $f_1(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t})$, as

$$\begin{aligned}&\sum_{i=1}^N \omega_i U\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{z}_{i,t}]\right) - \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{e}_t] \\ &+\sum_{i=1}^N \omega_i U\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{x}_{i,t}]\right) \\ &-\sum_{i=1}^N \omega_i U\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\tilde{z}_{i,t}]\right).\end{aligned}$$

Due to (45), the last two terms cancel each other. Hence, by (50), we have $f_1(\tilde{\mathbf{x}}_{d,t}, \tilde{\mathbf{x}}_{u,t}) \ge f_1(\mathbf{x}^{\text{opt}}_{d,t}, \mathbf{x}^{\text{opt}}_{u,t}) - B/V$, which completes the proof.

References

[1] U.S. Dept. Energy, "One million electric vehicles by 2015," Tech. Rep., Feb. 2011.

[2] C. Guille and G. Gross, "A conceptual framework for the vehicle-to-grid (V2G) implementation," *Energy Policy*, vol. 37, pp. 4379–4390, Nov. 2009.

[3] W. Kempton and J. Tomic, "Vehicle-to-grid power fundamentals: calculating capacity and net revenue," *J. Power Sources*, vol. 144, pp. 268–279, Jun. 2005.

[4] B. Kirby, "Frequency regulation basics and trends," U.S. Dept. Energy, Tech. Rep., 2005.

[5] W. Kempton, V. Udo, K. Huber, K. Komara, S. Letendre, S. Baker, D. Brunner, and N. Pearre, "A test of vehicle-to-grid (V2G) for energy storage and frequency regulation in the PJM system," Tech. Rep., Nov. 2008. [Online]. Available: http://www.udel.edu/V2G/resources/test-v2g-in-pjm-jan09.pdf

[6] R. Bessa and M. Matos, "Economic and technical management of an aggregation agent for electric vehicles: a literature survey," *Eur. Trans. Elect. Power*, vol. 22, pp. 334–350, Apr. 2011.

[7] J. Garzas, A. Armada, and G. Granados, "Fair design of plug-in electric vehicles aggregator for V2G regulation," *IEEE Trans. Veh. Technol.*, vol. 61, pp. 3406–3419, Oct. 2012.

[8] S. Han, S. Han, and K. Sezaki, "Optimal control of the plug-in electric vehicles for V2G frequency regulation using quadratic programming," in *Proc. IEEE ISGT*, Jan. 2011.

[9] E. Sortomme and M. Sharkawi, "Optimal scheduling of vehicle-to-grid energy and ancillary services," *IEEE Trans. Smart Grid*, vol. 3, pp. 351–359, Mar. 2012.

[10] S. Han, S. Han, and K. Sezaki, "Development of an optimal vehicle-to-grid aggregator for frequency regulation," *IEEE Trans. Smart Grid*, vol. 1, pp. 65–72, Jun. 2010.

[11] W. Shi and V. Wong, "Real-time vehicle-to-grid control algorithm under price uncertainty," in *Proc. IEEE SmartGridComm*, Oct. 2011.

[12] C. Wu, H. Rad, and J. Huang, "Vehicle-to-aggregator interaction game," *IEEE Trans. Smart Grid*, vol. 3, pp. 434–441, Mar. 2012.

[13] M. Neely, A. Tehrani, and A. Dimakis, "Efficient algorithms for renewable energy allocation to delay tolerant consumers," in *Proc. IEEE SmartGridComm*, Oct. 2010.

[14] S. Chen, P. Sinha, and N. Shroff, "Scheduling heterogeneous delay tolerant tasks in smart grid with renewable energy," in *Proc. IEEE CDC*, Dec. 2012.

[15] Y. Huang, S. Mao, and R. Nelms, "Adaptive electricity scheduling in microgrids," in *Proc. IEEE INFOCOM*, Apr. 2013.

[16] M. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[17] S. Sun, M. Dong, and B. Liang, "Real-time welfare-maximizing regulation allocation in aggregator-EVs systems," in *Proc. IEEE INFOCOM Workshop on CCSES*, Apr. 2013.

[18] S. Han, S. Han, and K. Sezaki, "Economic assessment on V2G frequency regulation regarding the battery degradation," in *Proc. IEEE ISGT*, Jan. 2012.

[19] S. Shakkottai and R. Srikant, *Network Optimization and Control*. Now Publishers Inc, 2007.

[20] R. Urgaonkar, B. Urgaonkar, M. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *Proc. ACM SIGMETRICS*, 2011.

[21] Tesla Model S. [Online]. Available: http://www.teslamotors.com/

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

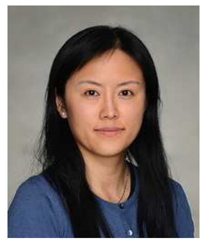
[23] Ford Focus Electric. [Online]. Available: http://www.ford.ca/cars/focus/

[24] "Fast response regulation signal." [Online]. Available: http://www.pjm.com/markets-and-operations/ancillary-services/mkt-based-regulation/fast-response-regulation-signal.aspx

[25] "Ancillary services manual." [Online]. Available: http://www.nyiso.com/public/webdocs/markets_operations/documents/Manuals_and_Guides/Manuals/Operations/ancserv.pdf

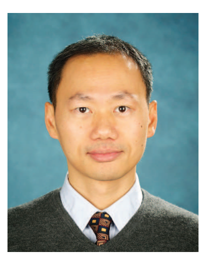
[26] "Electricity prices in ontario." [Online]. Available: http://www.ontarioenergyboard.ca/OEB/Consumers/Electricity/Electricity+Prices

Sun Sun (S'11) received the B.S. degree in Electrical Engineering and Automation from Tongji University, Shanghai, China, in 2005. From 2006 to 2008, she was a software engineer in the Department of GSM Base Transceiver Station of Huawei Technologies Co. Ltd.. She received the M.Sc. degree in Electrical and Computer Engineering from University of Alberta, Edmonton, Canada, in 2011. Now, she is pursuing her Ph.D. degree in the Department of Electrical and Computer Engineering of University of Toronto, Toronto, Canada. Her current research interest lies in the areas of stochastic optimization and distributed control, with the application of energy management in smart grid.



Min Dong (S'00-M'05-SM'09) received the B.Eng. degree from Tsinghua University, Beijing, China, in 1998, and the Ph.D. degree in electrical and computer engineering with minor in applied mathematics from Cornell University, Ithaca, NY, in 2004. From 2004 to 2008, she was with Corporate Research and Development, Qualcomm Inc., San Diego, CA. In 2008, she joined the Department of Electrical, Computer and Software Engineering at University of Ontario Institute of Technology, Ontario, Canada, where she is currently an Associate Professor. She also holds a status-only Associate Professor appointment with the Department of Electrical and Computer Engineering, University of Toronto since 2009. Her research interests are in the areas of statistical signal processing for communication networks, cooperative communications and networking techniques, and stochastic network optimization in dynamic networks and systems.

Dr. Dong received the Early Researcher Award from Ontario Ministry of Research and Innovation in 2012, the Best Paper Award at IEEE ICCC in 2012, and the 2004 IEEE Signal Processing Society Best Paper Award. She was an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS during 2009-2013, and currently serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. She has been an elected member of IEEE Signal Processing Society Signal Processing for Communications and Networking (SP-COM) Technical Committee since 2013.



Ben Liang (S'94-M'01-SM'06) received honors-simultaneous B.Sc. (valedictorian) and M.Sc. degrees in Electrical Engineering from Polytechnic University in Brooklyn, New York, in 1997 and the Ph.D. degree in Electrical Engineering with Computer Science minor from Cornell University in Ithaca, New York, in 2001. In the 2001 - 2002 academic year, he was a visiting lecturer and post-doctoral research associate at Cornell University. He joined the Department of Electrical and Computer Engineering at the University of Toronto in 2002, where he is now a Professor. His current research interests are in mobile communications and networked systems. He has served as an editor for the IEEE Transactions on Wireless Communications and an associate editor for the Wiley Security and Communication Networks journal, in addition to regularly serving on the organizational or technical committee of a number of conferences. He is a senior member of IEEE and a member of ACM and Tau Beta Pi.