

Adaptive Pricing for Optimal Coordination in Networked Energy Systems with Nonsmooth Cost Functions

Jiayi Li¹, Jiale Wei², Matthew Motoki¹, Yan Jiang², and Baosen Zhang¹

Abstract— Incentive-based coordination mechanisms for distributed energy consumption have shown promise in aligning individual user objectives with social welfare, especially under privacy constraints. Our prior work proposed a two-timescale adaptive pricing framework, where users respond to prices by minimizing their local costs, and the system operator iteratively updates the prices based on aggregate user responses. A key assumption was that the system cost need to smoothly depend on the aggregate of the user demands. In this paper, we relax this assumption by considering the more realistic model in which the system cost is determined by solving a Direct Current Optimal Power Flow (DCOPF) problem with constraints. We present a generalization of the pricing update rule that leverages the generalized gradients of the system cost function, which may be nonsmooth due to the structure of DCOPF. We prove that the resulting dynamic system converges to a unique equilibrium, which solves the social welfare optimization problem. Our theoretical results provide guarantees on convergence and stability using tools from nonsmooth analysis and Lyapunov theory. Numerical simulations on networked energy systems illustrate the effectiveness of the proposed scheme.

I. INTRODUCTION

Modern energy systems are undergoing a significant transformation, marked by the increasing prevalence of distributed energy resources (DERs), responsive loads, and the emergence of more autonomous devices. These developments have created opportunities for customers to actively participate in system operations. However, unlike dispatchable resources, customers often cannot be directly controlled by an operator¹ and must be coordinated through some form of incentives [3]. But the system and its customers often have competing objectives: system operators strive to achieve global objectives like efficiency, reliability, fairness, and stability of the network, while individual users optimize their private costs and preferences that are often unknown or unobservable to the system. In this paper, we study how

¹J. Li, M. Motoki, and B. Zhang are with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195, USA. Emails: {1jy9712, mmotoki, zhanbao}@uw.edu. They are partially supported by NSF grant ECCS-2153937 and the Washington Clean Energy Institute.

²J. Wei and Y. Jiang are with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, GD 518172, CHN. Emails: jialewei@link.cuhk.edu.cn; yjiang@cuhk.edu.cn. Y. Jiang is partially supported by CUHK SZ University Development Fund and part of this research was performed while the author was visiting the Institute for Mathematical and Statistical Innovation at University of Chicago, which is supported by the NSF grant DMS-1929348.

(Corresponding author: Yan Jiang.)

¹Direct load control exists and have been implemented, but are often constrained by the number of times and duration that they can be called [1], [2]. We do not explore this class of resources in this paper.

to achieve alignment between the system objective and user objective while keeping most of the information about the users private.

Incentive-based coordination mechanisms have received extensive attention and are one of the main features of power systems with communication capabilities [4]. In the context of demand response in electricity markets, incentives can take many different forms, ranging from alert/text-based signals [5] to pricing [6]. In this paper, we focus on a price-based incentive mechanism, in which a system operator broadcasts pricing signals, users respond by adjusting their consumption to minimize individual costs, and then the operator iteratively updates the prices based on user responses. Ideally, this iterative interaction should converge to an optimal solution that balances user cost and system performance. However, the major obstacle is that the operator typically lacks access to the cost functions of individual users, either due to privacy concerns or reliance on complex or black-box control strategies (e.g., reinforcement learning) of users themselves [7]–[9]. This limits the effectiveness of many pricing schemes and makes theoretical analysis difficult.

A previous work [10] proposed a two-timescale adaptive pricing framework adopted from a dynamic incentive [11] that evolves with the actions of users. In this framework, users act as price takers, optimizing their local behavior in response to a broadcast price signal, while the operator iteratively updates prices based on observed aggregate consumption. This iterative update circumvents the need for user-specific knowledge. Moreover, under mild conditions, such as monotonicity of user response with respect to price, this adaptive scheme has been shown to converge to the solution of the global social welfare optimization problem.

However, [10] made a key simplifying assumption that prices and the operator's objective are functions of aggregate demand alone (hence the prices are uniform across the users). Of course, in real-world power systems, electricity must be delivered over a physical network, where supply and demand must balance at each node, and transmission line capacities impose additional constraints. On top of these constraints, the operator solves an optimization problem which is often modeled as a Direct Current Optimal Power Flow (DCOPF) problem to determine the best way to satisfy the demands. This introduces new layers of complexity, since the cost depends nonlinearly and nonsmoothly on the demand, and the prices can exhibit discontinuities.

The nonsmoothness of the price arises quite naturally. In DCOPF problems, the feasible regions are polytopic and,

when the generator costs are linear, the optimal solutions can occur at the vertices of the feasible region [12]. Therefore, a small change in load can change the set of binding constraints and in turn causes discontinuous jumps in prices [13]. The algorithms in [10] and [11] use prices to infer gradient information about the system cost, but when the cost in DCOPF is nonsmooth and the prices are discontinuous, gradients are no longer well-defined.

This paper extends the adaptive pricing framework by embedding DCOPF constraints into the system operator's objective and carefully designing pricing updates based on generalized gradients of the (possibly nonsmooth) cost. Our formulation integrates network constraints directly into the operator's cost, and we propose a pricing update rule based on generalized gradients. This rule accounts for the potential non-differentiability of the cost function due to network constraints. Our main contributions are:

- 1) We design a vectorized price update rule based on the generalized gradient of the nonsmooth system cost induced by DCOPF, enabling implementation in realistic grid models.
- 2) We prove that the proposed iterative mechanism converges to a unique equilibrium that aligns user behavior with the social welfare solution, using tools from Lyapunov stability theory for nonsmooth systems.

This work offers a scalable and theoretically grounded approach to align local and global objectives in networked energy systems, opening the door to practical decentralized control under realistic grid constraints. The proposed mechanism is robust to privacy constraints, as the operator requires only demand observations and users do not need to disclose their cost functions or internal constraints, thus preserving privacy. We also demonstrate through simulations on networked scenarios that the mechanism effectively induces socially optimal behavior while maintaining system feasibility under DCOPF.

II. PROBLEM FORMULATION

In this section, we motivate the need to design an incentive pricing mechanism to achieve the global social welfare in networked energy systems while preserving user privacy.

A. Planner's Optimization Problem

We consider a supply-demand balancing electricity market with n users indexed by $i \in \mathcal{N} := \{1, \dots, n\}$. The power demand of user i is denoted by $\mathbf{x}_i \in \mathbb{R}$. For a given demand profile of users $\vec{\mathbf{x}} := (\mathbf{x}_i, i \in \mathcal{N}) \in \mathbb{R}^n$ which is the column vector obtained from the concatenation of all user demands, the disutility of the power consumption of user i is given by $f_i(\mathbf{x}_i)$, while the system cost in serving the demand profile of users is given by $J(\vec{\mathbf{x}})$.

We now discuss them separately:

Assumption 1 (Cost assumption). *The following assumptions on cost functions are made throughout this paper:*

- *Each user disutility function $f_i(\mathbf{x}_i)$ is strictly convex and twice continuously differentiable;*

- *The system cost function $J(\vec{\mathbf{x}})$ is a parametric programming determined from the DCOPF problem with linear generation costs:*

$$J(\vec{\mathbf{x}}) := \min_{\xi} \mathbf{c}^T \xi \quad (1)$$

s.t. Linear constraints on ξ depending on $\vec{\mathbf{x}}$

where $\mathbf{c} := (c_i, i \in \mathcal{G}) \in \mathbb{R}^{|\mathcal{G}|}$ denotes the vector of generation cost coefficients and $\xi := (\xi_i, i \in \mathcal{G}) \in \mathbb{R}^{|\mathcal{G}|}$ denotes the power generation from a set of generators \mathcal{G} . A nice feature of the optimal cost $J(\vec{\mathbf{x}})$ is that it is a convex function of the user demand profile $\vec{\mathbf{x}}$ [14]. However, although $J(\vec{\mathbf{x}})$ is continuous, it is not differentiable everywhere.

Then the system operator is interested in solving the following *global social welfare problem*:

$$\min_{\vec{\mathbf{x}} \in \mathbb{R}^n} C(\vec{\mathbf{x}}) := \sum_{i \in \mathcal{N}} f_i(\mathbf{x}_i) + J(\vec{\mathbf{x}}), \quad (2)$$

which minimizes the sum of the total disutility of all users and the system cost to serve users.

Remark 1 (Linear generation cost assumption). *Note that the linear cost in (1) is in some sense the most difficult cost function to deal with, at least in our setting. If the cost is strongly convex function, for example, a quadratic cost, J becomes differentiable everywhere. All of the results in the paper still hold since in that case the generalized gradient reduces to the (standard) gradient and all sets are singletons. Therefore, we focus on linear cost functions in this paper.*

As discussed before, we adopt the standard assumption that each disutility function $f_i(\mathbf{x}_i)$ is strictly convex and twice continuously differentiable [15], [16] while the system cost function $J(\vec{\mathbf{x}})$ is convex and locally Lipschitz². Hence, it is easy to see that the entire objective function is strictly convex and locally Lipschitz, which implies the existence of a unique global minimizer $\vec{\mathbf{x}}^*$ to problem (2) [18, Proposition 3.1.1]. By [19, Theorem 8.2], $\vec{\mathbf{x}}^*$ is such a minimizer if and only if

$$\begin{aligned} \mathbf{0} &\in \partial \left(\sum_{i \in \mathcal{N}} f_i(\mathbf{x}_i^*) + J(\vec{\mathbf{x}}^*) \right) \\ &= (\nabla f_i(\mathbf{x}_i^*), i \in \mathcal{N}) + \partial J(\vec{\mathbf{x}}^*), \end{aligned} \quad (3)$$

where the equality is due to the sum rule of the generalized gradient for convex functions [20, Chapter 2.4]. The so-called generalized gradient is a counterpart to gradient for nonsmooth functions, which is often known to be subdifferential by optimization community. As mentioned in Assumption 1, $J(\vec{\mathbf{x}})$ is continuous but it is not differentiable everywhere, which forces us to borrow the generalized gradient concept.

²Every convex function is locally Lipschitz [17]. We list locally Lipschitz property explicitly for the purpose of emphasis.

Definition 1 (Generalized gradient [20]). If $g : \mathbb{R}^d \mapsto \mathbb{R}$ is a locally Lipschitz continuous function, then its generalized gradient $\partial g : \mathbb{R}^d \mapsto \mathcal{B}(\mathbb{R}^d)$ at $z \in \mathbb{R}^d$ is defined by³

$$\partial g(z) := \text{co} \left\{ \lim_{k \rightarrow \infty} \nabla g(z_k) : z_k \rightarrow z, z_k \notin \Omega_g \cap \mathcal{S} \right\},$$

where co denotes convex hull, $\Omega_g \subset \mathbb{R}^d$ denotes the set of points where g fails to be differentiable, and $\mathcal{S} \subset \mathbb{R}^d$ is a set of measure zero that can be arbitrarily chosen to simply the computation.

Remark 2 (Relation to gradient). Unlike a gradient which gives a single vector, a generalized gradient is a set-valued map. The generalized gradient is the generalization of the gradient in the sense that, if g is differentiable at z , then $\partial g(z) = \{\nabla g(z)\}$.

However, in practice, the planner's optimization problem in (2) is not implementable due to the lack of knowledge of the exact disutility functions of users for privacy concerns. This poses challenges for the system operator to realize economic dispatch by solving (2) directly. An important way to address this issue by the system operator is to update its power price $p_i \in \mathbb{R}$ for individual users iteratively based on how users adjust their desired power. By doing so, the system operator hopes to encourage users to align their individual goals of cost minimization with the goal of problem (2). The design of such an adaptive price update will be discussed later, which is the core of this paper.

B. User's Optimization Problem

All users are assumed to be rational price takers. More precisely, given the power price p_i , each user i adjusts its power consumption by solving the following optimization problem:

$$\min_{\mathbf{x}_i} f_i(\mathbf{x}_i) + p_i \mathbf{x}_i, \quad (4)$$

which minimizes the total cost of user i induced by disutility and payment for power consumption. Since (4) is an unconstrained convex optimization problem, the necessary and sufficient condition for \mathbf{x}_i^* to be a minimizer is [21, Chapter 4.2.3]

$$\nabla f_i(\mathbf{x}_i^*) + p_i = 0, \quad (5)$$

which yields a unique global solution \mathbf{x}_i^* by the strict convexity of $f_i(\mathbf{x}_i)$ [18, Proposition 3.1.1]. Basically, as the system operator updates its price signal p_i , user i adjusts its power demand \mathbf{x}_i^* accordingly to satisfy (5) in a unique way. To put it another way, for any given price p_i , the demand \mathbf{x}_i^* is unique. Hence, \mathbf{x}_i^* is clearly a function [22, Definition 2.1] of the current price p_i and can be expressed as

$$\mathbf{x}_i^*(p_i) := \arg \min_{\mathbf{x}_i} f_i(\mathbf{x}_i) + p_i \mathbf{x}_i.$$

³Throughout this paper, we use $\mathcal{B}(\mathbb{R}^d)$ to denote the collection of all subsets of \mathbb{R}^d and $B(z; \epsilon)$ to denote the open ball centered at $z \in \mathbb{R}^d$ with radius $\epsilon > 0$.

An important feature of this function $\mathbf{x}_i^*(p_i)$ is that it is a continuously differentiable and strictly decreasing function, which is highlighted by the following lemma.

Lemma 1 (Bijective demand update). Under Assumption 1, the demand update $\mathbf{x}_i^*(p_i)$ is a bijection given by a continuously differentiable and strictly decreasing function

$$\mathbf{x}_i^*(p_i) = \nabla^{-1} f_i(-p_i), \quad (6)$$

which naturally has the properties that $\nabla \mathbf{x}_i^*(p_i) < 0$ and, $\forall \tilde{p}_i, \hat{p}_i \in \mathbb{R}$, if $\tilde{p}_i \neq \hat{p}_i$, then $\mathbf{x}_i^*(\tilde{p}_i) \neq \mathbf{x}_i^*(\hat{p}_i)$.

Proof. First, by [23, Theorem 2.14], the strict convexity of $f_i(\mathbf{x}_i)$ ensures that its gradient $\nabla f_i(\mathbf{x}_i)$ is a strictly increasing function. Then, this strict monotonicity implies that $\nabla f_i(\mathbf{x}_i)$ is a bijection, which further implies that $\nabla f_i(\mathbf{x}_i)$ has a unique inverse function written as $\nabla^{-1} f_i$ that is also a bijection. Now, we note that, as an optimal solution, $\mathbf{x}_i^*(p_i)$ must satisfy (5), i.e.,

$$\nabla f_i(\mathbf{x}_i^*(p_i)) + p_i = 0. \quad (7)$$

Since $\nabla^{-1} f_i$ is well-defined, we are allowed to represent $\mathbf{x}_i^*(p_i)$ in (7) as (6), from which it is easy to see that $\mathbf{x}_i^*(p_i)$ is a bijection since $\nabla^{-1} f_i$ is a bijection.

Moreover, an important property of any bijective function is that it is one-to-one, which means that every element in the codomain is mapped to by at most one element in the domain. Thus, $\forall \tilde{p}_i, \hat{p}_i \in \mathbb{R}$, if $\mathbf{x}_i^*(\tilde{p}_i) = \mathbf{x}_i^*(\hat{p}_i)$, then $\tilde{p}_i = \hat{p}_i$, which is logically equivalent to the contrapositive, i.e., $\forall \tilde{p}_i, \hat{p}_i \in \mathbb{R}$, if $\tilde{p}_i \neq \hat{p}_i$, then $\mathbf{x}_i^*(\tilde{p}_i) \neq \mathbf{x}_i^*(\hat{p}_i)$.

Finally, we would like to show that $\mathbf{x}_i^*(p_i)$ is a continuously differentiable function. By Assumption 1, $f_i(\mathbf{x}_i)$ is twice continuously differentiable, which implies that $\nabla f_i(\mathbf{x}_i)$ is continuously differentiable. That is, $\nabla^2 f_i(\mathbf{x}_i)$ exists everywhere and is continuous. As mentioned in the beginning of the proof, $\nabla f_i(\mathbf{x}_i)$ is strictly increasing, which implies that $\nabla^2 f_i(\mathbf{x}_i) > 0$ everywhere. By inverse function theorem, $\nabla^{-1} f_i$ is continuously differentiable and its derivative at $(-\mathbf{x}_i)$ is given by $1/\nabla^2 f_i(\nabla^{-1} f_i(-\mathbf{x}_i)) = 1/\nabla^2 f_i(\mathbf{x}_i^*(p_i)) > 0$ since $\nabla^2 f_i(\mathbf{x}_i) > 0$ everywhere. Thus, $\nabla \mathbf{x}_i^*(p_i) = -1/\nabla^2 f_i(\nabla^{-1} f_i(-\mathbf{x}_i)) = -1/\nabla^2 f_i(\mathbf{x}_i^*(p_i)) < 0$, which is clearly continuous since $\nabla^2 f_i(\mathbf{x}_i)$ is continuous. This concludes the proof that $\mathbf{x}_i^*(p_i)$ is continuously differentiable and strictly decreasing. \square

Lemma 1 shows that the demand update $\mathbf{x}_i^*(p_i)$ is a bijective function which naturally enjoys a nice property. That is, $\forall \tilde{p}_i, \hat{p}_i \in \mathbb{R}$, if $\tilde{p}_i \neq \hat{p}_i$, then $\mathbf{x}_i^*(\tilde{p}_i) \neq \mathbf{x}_i^*(\hat{p}_i)$, which means that it is impossible for distinct price signals to motivate the same power demand. As the analysis will unfold later, this “uniqueness” plays a role in the convergence of the pricing mechanism which we will propose.

Therefore, our goal is to design a suitable update for price profile $\tilde{\mathbf{p}} := (\mathbf{p}_i, i \in \mathcal{N}) \in \mathbb{R}^n$ which can leverage the nice demand update $\tilde{\mathbf{x}}^*(\tilde{\mathbf{p}}) := (\mathbf{x}_i^*(p_i), i \in \mathcal{N}) \in \mathbb{R}^n$ induced by individual user's optimization problem (4) in each iteration to gradually gear the demand profile $\tilde{\mathbf{x}}$ of users toward the minimizer of (2) after enough iterations. This incentive

pricing mechanism allows the system operator to achieve the desired solution to the planner's optimization problem (2) without solving it directly.

III. ADAPTIVE PRICE UPDATE UNDER NONSMOOTHNESS

In terms of incentive pricing mechanism, our recent work [10] proposes for a similar but simpler setting to adopt price dynamics utilizing the gradient information of the system cost function $J(\cdot)$ to incentivize users to adjust their power consumption towards a point where the planner's problem and user's problem are simultaneously solved. However, the underlying assumption that $J(\cdot)$ is smooth does not hold in our case due to the particular choice of $J(\cdot)$ as (1) which makes $J(\cdot)$ convex and locally Lipschitz but not differentiable everywhere. Thus, we propose an adaptive price update by leveraging the generalized gradient in Definition 1 as follows:

$$\dot{\vec{p}} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p}. \quad (8)$$

This is well-defined since the fact that $J(\vec{x})$ is a locally Lipschitz continuous function ensures that $J(\vec{x})$ has a nonempty compact set as its generalized gradient at any $\vec{x} \in \mathbb{R}^n$ [24, Proposition 6].

Based on (8), we now illustrate the incentive pricing mechanism in more detail. As shown in Fig. 1, we consider a two-timescale design of incentive pricing mechanism, where individual users solve (4) for $\vec{x}_i^*(\vec{p}_i)$ much faster than the system operator updates the price \vec{p} via (8). This timescale separation allows users to consider the price signal \vec{p} as static when solving for $\vec{x}^*(\vec{p})$. Thus, following any given price \vec{p} provided by the system operator, users adjust their power consumption towards $\vec{x}^*(\vec{p})$ almost immediately by solving (5) and then the system operator updates the price \vec{p} according to (8) in response to the current demand profile $\vec{x}^*(\vec{p})$. It should be intuitively clear that (8) provides users with incentives to align their own interests with social welfare, given that adjustments to \vec{p} intend to reduce the difference between the marginal cost of the individual user quantified by \vec{p} and the marginal cost of the system operator characterized by $\partial J(\vec{x}^*(\vec{p}))$.

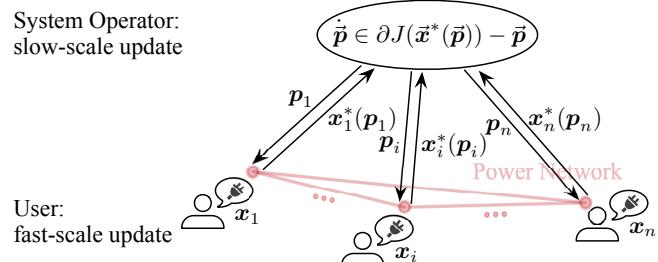


Fig. 1: Two-timescale design of incentive pricing mechanism.

With this in mind, as the system operator iteratively updates the price \vec{p} , the nonsmooth dynamical system composed of (5) and (8) ideally should settle down at a point that achieves the optimal solution to the planner's optimization problem (2). That is, at the equilibrium price $\vec{p}^* :=$

$(\vec{p}_i^*, i \in \mathcal{N}) \in \mathbb{R}^n$, we would like to have $\vec{x}^*(\vec{p}^*)$ satisfy (3), which is captured by the following theorem.

Theorem 1 (Unique equilibrium with incentive aligned).

Under Assumption 1, the demand profile $\vec{x}^*(\vec{p}^*)$ occurring at the unique equilibrium price \vec{p}^* of the dynamical system composed of (5) and (8) is the unique global minimizer to the planner's optimization problem (2), i.e.,

$$\mathbf{0} \in \partial J(\vec{x}^*(\vec{p}^*)) + (\nabla f_i(\vec{x}_i^*(\vec{p}_i^*)), i \in \mathcal{N}). \quad (9)$$

Proof. The point \vec{p}^* is an equilibrium of the price update (8) if and only if [20, Chapter 4.4]

$$\mathbf{0} \in \partial J(\vec{x}^*(\vec{p}^*)) - \vec{p}^*. \quad (10)$$

Note that $\vec{x}^*(\vec{p}^*)$ generated from the demand update satisfies (5), i.e.,

$$\nabla f_i(\vec{x}_i^*(\vec{p}_i^*)) + \vec{p}_i^* = \mathbf{0}, \quad \forall i \in \mathcal{N},$$

from which we know

$$\vec{p}^* = -(\nabla f_i(\vec{x}_i^*(\vec{p}_i^*)), i \in \mathcal{N}). \quad (11)$$

Substituting (11) into (10) yields (9), which is exactly in the form of the optimality condition (3) for the planner's optimization problem (2). Thus, $\vec{x}^*(\vec{p}^*)$ corresponding to the equilibrium price \vec{p}^* is the unique global minimizer to (2). Now, it remains to show that the equilibrium price \vec{p}^* is unique. By way of contradiction, suppose that both \vec{p}^* and \vec{p}° satisfy (10), where $\vec{p}^* \neq \vec{p}^{\circ}$. Then, by a similar argument as above, we know that both $\vec{x}^*(\vec{p}^*)$ and $\vec{x}^*(\vec{p}^{\circ})$ must satisfy the optimality condition (3). Thus, $\vec{x}^*(\vec{p}^*)$ and $\vec{x}^*(\vec{p}^{\circ})$ are both optimizers of problem (2). Notably, by Lemma 1, our assumption $\vec{p}^* \neq \vec{p}^{\circ}$ directly implies $\vec{x}^*(\vec{p}^*) \neq \vec{x}^*(\vec{p}^{\circ})$. Therefore, we now reach a situation where there are two distinct optimizers to problem (2), which contradicts the fact that problem (2) has a unique minimizer. This concludes the proof of the uniqueness of the equilibrium price \vec{p}^* . \square

Theorem 1 verifies that the proposed incentive pricing mechanism is guaranteed to settle down at a unique equilibrium price \vec{p}^* whose corresponding demand profile $\vec{x}^*(\vec{p}^*)$ is exactly the unique global minimizer to the planner's optimization problem (2). In other words, by adopting the proposed adaptive price update, the system operator can encourage users to align their individual benefits with the social welfare. Thus, the system objective of economic dispatch is achieved without disclosure of user privacy.

IV. NONSMOOTH STABILITY ANALYSIS

Having characterized the equilibrium point and confirmed the incentive alignment at that point, we are now ready to investigate the stability of the nonsmooth dynamical system composed of the demand update (5) and the price update (8) by performing the natural extension of Lyapunov stability analysis provided in [24, Theorem 1]. More precisely, the stability under the incentive pricing mechanism can be certified by finding a well-defined Lyapunov function that is decreasing along the trajectories of the system comprising (5)

and (8). The main result of this section is presented below, whose proof is enabled by a sequence of intermediate results that we discuss next.

Theorem 2 (Asymptotic stability). *Under Assumption 1, the dynamical system composed of (5) and (8) is strongly asymptotically stable at the unique equilibrium characterized by (9).*

Of course, before showing the stability of the system, we need to show that the dynamical system has a solution. Here, we take the solution to be in the Caratheodory sense, which roughly says that, under some assumptions, there is a trajectory that satisfies (5) and (8) except for a set of time t that has Lebesgue measure zero [24]. The key for checking sufficient conditions for the existence of a Caratheodory solution here is to check whether the set-valued map $\vec{p} \mapsto [\partial J(\vec{x}^*(\vec{p})) - \vec{p}]$ in (8) has several properties [24, Proposition S2].

Since $\partial J(\vec{x}^*(\vec{p}))$ involves the composition of functions, we develop the following lemma to facilitate our analysis.

Lemma 2 (Property preservation in composition). *Assume that $h : \mathbb{R}^m \mapsto \mathbb{R}^d$ is continuous at $z \in \mathbb{R}^m$ and u is the composite of h and $g : \mathbb{R}^d \mapsto \mathcal{B}(\mathbb{R}^d)$ defined by*

$$u(z) := (g \circ h)(z) := g(h(z)). \quad (12)$$

- If $g : \mathbb{R}^d \mapsto \mathcal{B}(\mathbb{R}^d)$ is upper semicontinuous at $h(z) \in \mathbb{R}^d$, then u is upper semicontinuous at z .
- If $g : \mathbb{R}^d \mapsto \mathcal{B}(\mathbb{R}^d)$ is locally bounded at $h(z) \in \mathbb{R}^d$, then u is locally bounded at z .

Proof. We study the two cases separately.

For upper semicontinuity of a set-valued map [24], we need to show that, $\forall \epsilon > 0$, $\exists \delta > 0$ such that

$$u(\tilde{z}) \subset u(z) + B(\mathbf{0}; \epsilon), \quad \forall \tilde{z} \in B(z; \delta). \quad (13)$$

To this end, we first note that, if g is upper semicontinuous at $h(z) \in \mathbb{R}^d$, for any given $\epsilon > 0$, $\exists \eta > 0$ such that, whenever $y \in B(h(z); \eta)$, it holds that [24]

$$g(y) \subset g(h(z)) + B(\mathbf{0}; \epsilon). \quad (14)$$

Next, since h is continuous at $z \in \mathbb{R}^m$, for any given $\eta > 0$, $\exists \delta > 0$ such that, whenever $\tilde{z} \in B(z; \delta)$, it holds that [22, Definition 4.5]

$$h(\tilde{z}) \in B(h(z); \eta).$$

Now, we combine the above two arguments by setting $y = h(\tilde{z})$ in (14), which yields

$$g(h(\tilde{z})) \subset g(h(z)) + B(\mathbf{0}; \epsilon), \quad \forall \tilde{z} \in B(z; \delta). \quad (15)$$

Finally, from (12), we know $g(h(\tilde{z})) = u(\tilde{z})$ and $g(h(z)) = u(z)$, which combined with (15) gives exactly the claim (13) that we would like to prove.

For local boundedness of a set-valued map [24], we need to show that, $\exists \delta > 0$ and some constant $M > 0$ such that

$$\|\mu\|_2 \leq M, \quad \forall \tilde{z} \in B(z; \delta), \mu \in u(\tilde{z}). \quad (16)$$

With this aim, we first note that, if g is locally bounded at $h(z) \in \mathbb{R}^d$, $\exists \eta > 0$ and some constant $M > 0$ such that [24]

$$\|\mu\|_2 \leq M, \quad \forall y \in B(h(z); \eta), \mu \in g(y). \quad (17)$$

Again, since h is continuous at $z \in \mathbb{R}^m$, for any given $\eta > 0$, $\exists \delta > 0$ such that [22, Definition 4.5]

$$h(\tilde{z}) \in B(h(z); \eta), \quad \forall \tilde{z} \in B(z; \delta). \quad (18)$$

Now, we combine the above two arguments by setting $y = h(\tilde{z})$ in (17), which yields

$$\|\mu\|_2 \leq M, \quad \forall \tilde{z} \in B(z; \delta), h(\tilde{z}) \in B(h(z); \eta), \mu \in g(h(\tilde{z})).$$

Here, the second condition can be removed since it is directly implied by the first condition due to (18), which yields

$$\|\mu\|_2 \leq M, \quad \forall \tilde{z} \in B(z; \delta), \mu \in g(h(\tilde{z})). \quad (19)$$

Finally, from (12), we know $g(h(\tilde{z})) = u(\tilde{z})$, which substituted into (19) gives exactly the claim (16) that we would like to prove. \square

Lemma 2 paves us a way to show the existence of a Caratheodory solution of our dynamical system by checking conditions in [24, Proposition S2], which is the core of the next lemma.

Lemma 3 (Existence of a Caratheodory solution). *Under Assumption 1, there exists a Caratheodory solution of the dynamical system composed of (5) and (8) for any initial condition $\vec{p}(0)$.*

Proof. Basically, by [24, Proposition S2], it suffices to show that the set-valued map $\vec{p} \mapsto [\partial J(\vec{x}^*(\vec{p})) - \vec{p}]$ takes nonempty compact convex values and is also upper semicontinuous as well as locally bounded⁴.

Clearly, it is the term $\partial J(\vec{x}^*(\vec{p}))$ associated with the generalized gradient in the above mapping that makes our dynamics different from differential equations. Thus, we focus our analysis on properties of $\partial J(\vec{x}^*(\vec{p}))$, which is a composition of $\partial J(\vec{x})$ and $\vec{x}^*(\vec{p})$.

We start by investigating $\partial J(\vec{x})$. Based on [24, Proposition 6], it follows directly from the local Lipschitz continuity of $J(\vec{x})$ that $\partial J(\vec{x})$ is a nonempty compact convex set at any \vec{x} and the set-valued map $\vec{x} \mapsto \partial J(\vec{x})$ is upper semicontinuous and locally bounded at any \vec{x} .

As for $\vec{x}^*(\vec{p}) := (x_i^*(p_i), i \in \mathcal{N})$, it is a continuous vector-valued function since each component $x_i^*(p_i)$ is a continuous function by Lemma 1 [25, Theorem 2.4].

With above information about $\partial J(\vec{x})$ and $\vec{x}^*(\vec{p})$, we are now ready to examine properties of $\partial J(\vec{x}^*(\vec{p}))$. First, given that $\partial J(\vec{x})$ is a nonempty compact convex set at any \vec{x} , it must be true that $\partial J(\vec{x}^*(\vec{p}))$ is a nonempty compact convex set at any \vec{p} as well since $\vec{x}^*(\vec{p})$ is a bijective function by Lemma 1. This can be understood by noting that, no matter what particular value the price signal \vec{p} is taking, $\partial J(\vec{x})$ will take a corresponding value $\vec{x} = \vec{x}^*(\vec{p})$ at that \vec{p} , which

⁴There is not need to check measurability here since (8) does not explicitly depend on time t .

must produce a nonempty compact convex set $\partial J(\vec{x}^*(\vec{p}))$. Second, the upper semicontinuity and local boundedness of $\partial J(\vec{x}^*(\vec{p}))$ at any \vec{p} follow from Lemma 2 by setting $h = \vec{x}^*(\vec{p})$ which is continuous and $g = \partial J(\vec{x})$ which is upper semi-continuous and locally bounded.

Finally, the term $(-\vec{p})$ has no influence to the above properties. First, it only translates the nonempty compact convex set $\partial J(\vec{x}^*(\vec{p}))$ by $(-\vec{p})$, which is still a nonempty compact convex set. Thus, $\vec{p} \mapsto [\partial J(\vec{x}^*(\vec{p})) - \vec{p}]$ takes nonempty compact convex values. Second, it can be considered as a continuous function which is inherently upper semicontinuous and locally bounded at \vec{p} . Since the summation of two upper semicontinuous functions is still upper semicontinuous and the summation of two locally bounded functions is still locally bounded. Thus, $\vec{p} \mapsto [\partial J(\vec{x}^*(\vec{p})) - \vec{p}]$ is also upper semicontinuous and locally bounded. The result follows from [24, Propostion S2]. \square

After determining the existence of a Caratheodory solution of the system from any initial point through Lemma 3, we now examine the nonsmooth system stability by constructing a candidate Lyapunov function. According to [24, Theorem 1], we need to first seek a locally Lipschitz and regular function $V(\vec{p})$ that satisfies $V(\vec{p}^*) = 0$ and $V(\vec{p}) > 0$, $\forall \vec{p} \neq \vec{p}^*$, and then verify its monotonic evolution along the system trajectories, which is more complicated compared to a standard analysis since we need to study the Lie derivative in a nonsmooth setting.

We consider the following Lyapunov function candidate:

$$V(\vec{p}) := C(\vec{x}^*(\vec{p})) - C(\vec{x}^*(\vec{p}^*)), \quad (20)$$

where $C(\cdot)$ denotes the objective function of the planner's optimization problem (2) and \vec{p}^* corresponds to the unique equilibrium point of the system satisfying (9). The next result shows that this is a well-defined Lyapunov function candidate.

Lemma 4 (Well-defined Lyapunov function). *Under Assumption 1, $V(\vec{p})$ defined in (20) is a locally Lipschitz and regular function that satisfies $V(\vec{p}^*) = 0$ and $V(\vec{p}) > 0$, $\forall \vec{p} \neq \vec{p}^*$.*

Proof. As discussed in Section II-A, the entire objective function $C(\vec{x})$ of the planner's optimization problem (2) is locally Lipschitz and strictly convex, which together with the continuous differentiability of each $x_i^*(\mathbf{p}_i)$ by Lemma 1 allows us to show that $V(\vec{p})$ is locally Lipschitz and regular. We now illustrate this in detail.

We begin with the local Lipschitz continuity of $C(\vec{x}^*(\vec{p}))$. Clearly, the continuous differentiability of each $x_i^*(\mathbf{p}_i)$ implies that each $x_i^*(\mathbf{p}_i)$ is locally Lipschitz [26, Chapter 17.2]. Now, since each component of $\vec{x}^*(\vec{p}) := (x_i^*(\mathbf{p}_i), i \in \mathcal{N})$ is locally Lipschitz and $C(\vec{x})$ is locally Lipschitz as well, their composition $C(\vec{x}^*(\vec{p}))$ is locally Lipschitz by the chain rule [24].

We next investigate the regularity of $C(\vec{x}^*(\vec{p}))$. First of all, $\vec{x}^*(\vec{p}) := (x_i^*(\mathbf{p}_i), i \in \mathcal{N})$ is a continuously differentiable vector-valued function since each component $x_i^*(\mathbf{p}_i)$

is a continuously differentiable function [27, Theorem 2.8]. Moreover, $C(\vec{x})$ is locally Lipschitz and strictly convex, which further ensures that $C(\vec{x})$ is regular [20, Proposition 2.4.3]. By [28, Theorem 8.18], as a composite of $\vec{x}^*(\vec{p})$ and $C(\vec{x})$, $C(\vec{x}^*(\vec{p}))$ is regular.

The conclusion that $V(\vec{p})$ is locally Lipschitz and regular can be trivially drawn if one notices that the other term of $V(\vec{p})$ in (20) is just a constant.

Clearly, $V(\vec{p}^*) = 0$ by construction. To see why $V(\vec{p}) > 0$, $\forall \vec{p} \neq \vec{p}^*$, we first note that $\vec{x}^*(\vec{p}^*)$ is the unique global minimizer to problem (2) by Theorem 1, which means that

$$\forall \vec{x} \neq \vec{x}^*(\vec{p}^*), \quad C(\vec{x}) > C(\vec{x}^*(\vec{p}^*)). \quad (21)$$

Moreover, we know from Lemma 1 that, $\forall \vec{p} \neq \vec{p}^*$, it holds that $\vec{x}^*(\vec{p}) \neq \vec{x}^*(\vec{p}^*)$. Therefore, setting $\vec{x} = \vec{x}^*(\vec{p})$ in (21), we get $C(\vec{x}^*(\vec{p})) > C(\vec{x}^*(\vec{p}^*))$, i.e., $C(\vec{x}^*(\vec{p})) - C(\vec{x}^*(\vec{p}^*)) > 0$, which is equivalent to $V(\vec{p}) > 0$ by our construction of $V(\vec{p})$ in (20). This confirms that $V(\vec{p}) > 0$, $\forall \vec{p} \neq \vec{p}^*$, as desired. \square

Next, we turn to verify the monotonic evolution of $V(\vec{p})$ along the system trajectories by the notion of Lie derivative in the nonsmooth setting, which requires $\max \tilde{\mathcal{L}}V(\vec{p}) < 0$, $\forall \vec{p} \neq \vec{p}^*$, with $\tilde{\mathcal{L}}V(\vec{p})$ being the set-valued Lie derivative of V regarding $[\partial J(\vec{x}^*(\vec{p})) - \vec{p}]$ in (8) at \vec{p} defined by [24], [29]

$$\begin{aligned} \tilde{\mathcal{L}}V(\vec{p}) := & \{a \in \mathbb{R} : \exists \mathbf{v} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p} \text{ such that} \\ & \zeta^T \mathbf{v} = a, \forall \zeta \in \partial V(\vec{p})\}. \end{aligned} \quad (22)$$

Lemma 5 (Negativity of Lie derivative). *Under Assumption 1, the set-valued Lie derivative of $V(\vec{p})$ defined in (22) satisfies $\max \tilde{\mathcal{L}}V(\vec{p}) < 0$, $\forall \vec{p} \neq \vec{p}^*$.*

Proof. Before delving into $\tilde{\mathcal{L}}V(\vec{p})$, we need to characterize $\partial V(\vec{p})$ which can be computed as

$$\begin{aligned} \partial V(\vec{p}) &= \partial(C \circ \vec{x}^*)(\vec{p}) \\ &\stackrel{\textcircled{1}}{=} \nabla \vec{x}^*(\vec{p}) \partial C(\vec{x}^*(\vec{p})) \\ &:= \{\nabla \vec{x}^*(\vec{p}) \boldsymbol{\eta} : \boldsymbol{\eta} \in \partial C(\vec{x}^*(\vec{p}))\} \\ &= \{\text{diag}(\nabla x_i^*(\mathbf{p}_i), i \in \mathcal{N}) \boldsymbol{\eta} : \boldsymbol{\eta} \in \partial C(\vec{x}^*(\vec{p}))\}. \end{aligned} \quad (23)$$

In $\textcircled{1}$, the chain rule of the generalized gradient [28, Theorem 8.18] can be used with equality since, as discussed in the proof of Lemma 4, the conditions that $C(\vec{x})$ is locally Lipschitz and regular and that $\vec{x}^*(\vec{p})$ is continuously differentiable both hold.

To get a more explicit expression for (23), we now derive $\partial C(\vec{x}^*(\vec{p}))$ as

$$\begin{aligned} \partial C(\vec{x}^*(\vec{p})) &= \partial \left(\sum_{i \in \mathcal{N}} f_i(x_i^*(\mathbf{p}_i)) + J(\vec{x}^*(\vec{p})) \right) \\ &= (\nabla f_i(x_i^*(\mathbf{p}_i)), i \in \mathcal{N}) + \partial J(\vec{x}^*(\vec{p})) \\ &= \partial J(\vec{x}^*(\vec{p})) - \vec{p}, \end{aligned} \quad (24)$$

where the first equality is due to the definition of $C(\vec{x})$ in (2), the second equality is due to the sum rule of the generalized gradient for convex functions [20, Chapter 2.4] as mentioned

in Section II-A, and the last equality uses the relation that $(\nabla f_i(\mathbf{x}_i^*(\mathbf{p}_i)), i \in \mathcal{N}) = -\vec{p}$ resulting from the optimality condition (7) of user's problem as discussed in the proof of Lemma 1.

Substituting (24) to (23) yields

$$\begin{aligned}\partial V(\vec{p}) &= \{\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{\eta} : \\ &\quad \boldsymbol{\eta} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p}\},\end{aligned}\quad (25)$$

which will be applied to (22) for investigating the sign of $\max \tilde{L}V(\vec{p})$. The challenging part is that $J(\cdot)$ is continuous but not differentiable everywhere, which means that there exists a set of points of \vec{p} for which $J(\cdot)$ fails to be differentiable at the corresponding $\vec{x}^*(\vec{p})$. For the ease of notation, we denote such a set of \vec{p} as $\Omega_{J(\vec{x}^*(\vec{p}))} \subset \mathbb{R}^n$. Note that we only care about points \vec{p} different from the equilibrium \vec{p}^* in this particular analysis, which are \vec{p} satisfying

$$\mathbf{0} \notin \partial J(\vec{x}^*(\vec{p})) - \vec{p} \quad (26)$$

by (10) in the proof of Theorem 1. This allows us to consider two cases based on whether $\vec{p} \neq \vec{p}^*$ is in $\Omega_{J(\vec{x}^*(\vec{p}))}$ or not.

- 1) If $\vec{p} \neq \vec{p}^*$ and $\vec{p} \notin \Omega_{J(\vec{x}^*(\vec{p}))}$: The generalized gradient $\partial J(\vec{x}^*(\vec{p}))$ reduces to a singleton, i.e.,

$$\partial J(\vec{x}^*(\vec{p})) = \{\nabla J(\vec{x}^*(\vec{p}))\}. \quad (27)$$

Thus, $\partial V(\vec{p})$ in (25) reduces to a singleton as well, i.e.,

$$\partial V(\vec{p}) = \{\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p})\} \quad (28)$$

with

$$F(\vec{p}) := \nabla J(\vec{x}^*(\vec{p})) - \vec{p}$$

in this case, which together with (27) simplifies (22) to

$$\begin{aligned}\tilde{L}V(\vec{p}) &= \{a \in \mathbb{R} : \exists \boldsymbol{v} \in \{F(\vec{p})\} \text{ such that } \boldsymbol{\zeta}^T \boldsymbol{v} = a, \\ &\quad \forall \boldsymbol{\zeta} \in \{\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p})\}\} \\ &= \{(\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p}))^T F(\vec{p})\} \\ &= \{F(\vec{p})^T \text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p})\}.\end{aligned}\quad (29)$$

We claim that

$$F(\vec{p})^T \text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p}) < 0 \quad (30)$$

for two reasons. First, $\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N}) \prec 0$ since each $\nabla \mathbf{x}_i^*(\mathbf{p}_i) < 0$ by Lemma 1. Second, since $\vec{p} \neq \vec{p}^*$, we know from (26) that $F(\vec{p}) \neq \mathbf{0}$. Then, (30) follows directly. Combining (29) and (30), we know

$$\max \tilde{L}V(\vec{p}) = F(\vec{p})^T \text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})F(\vec{p}) < 0.$$

- 2) If $\vec{p} \neq \vec{p}^*$ and $\vec{p} \in \Omega_{J(\vec{x}^*(\vec{p}))}$: Substituting (25) to (22) yields

$$\begin{aligned}\tilde{L}V(\vec{p}) &= \{a \in \mathbb{R} : \exists \boldsymbol{v} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p} \text{ such that } \boldsymbol{\zeta}^T \boldsymbol{v} = a, \\ &\quad \forall \boldsymbol{\zeta} \in \{\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{\eta} : \boldsymbol{\eta} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p}\}\}.\end{aligned}\quad (31)$$

If $\tilde{L}V(\vec{p}) = \emptyset$, then $\max \tilde{L}V(\vec{p}) = -\infty$ by convention [24]. If $\tilde{L}V(\vec{p}) \neq \emptyset$, we claim that, $\forall a \in \tilde{L}V(\vec{p})$

in (31), $a < 0$. To see this, we first note that, for any such a , $\exists \boldsymbol{v} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p}$, such that $\boldsymbol{\zeta}^T \boldsymbol{v} = a$, $\forall \boldsymbol{\zeta} \in \{\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{\eta} : \boldsymbol{\eta} \in \partial J(\vec{x}^*(\vec{p})) - \vec{p}\}$. Clearly, we can pick $\boldsymbol{\eta} = \boldsymbol{v}$ and then $\boldsymbol{\zeta} = \text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{v}$ to solve for

$$\begin{aligned}a &= \boldsymbol{\zeta}^T \boldsymbol{v} = (\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{v})^T \boldsymbol{v} \\ &= \boldsymbol{v}^T \text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N})\boldsymbol{v} < 0.\end{aligned}$$

Here, the inequality is due to a similar argument for the previous case. That is, $\text{diag}(\nabla \mathbf{x}_i^*(\mathbf{p}_i), i \in \mathcal{N}) \prec 0$ and $\boldsymbol{v} \neq \mathbf{0}$ from (26). Now, we have shown that, if $\tilde{L}V(\vec{p}) \neq \emptyset$, $\forall a \in \tilde{L}V(\vec{p})$ in (31), $a < 0$, which ensures that $\max \tilde{L}V(\vec{p}) < 0$. Therefore, no matter whether $\tilde{L}V(\vec{p})$ is empty or not, it must be true that $\max \tilde{L}V(\vec{p}) < 0$.

In sum, $\max \tilde{L}V(\vec{p}) < 0$, $\forall \vec{p} \neq \vec{p}^*$. \square

According to [24, Theorem 1], with the aid of Lemmas 3, 4, and 5, we now have all the elements necessary to establish the strongly asymptotical stability of the unique equilibrium as summarized in Theorem 2.

V. NUMERICAL ILLUSTRATIONS

In this section, we present simulation results to show the convergence of our proposed incentive pricing mechanism to the desired optimal solution to the global social welfare problem. The simulations are conducted on the IEEE 14-bus system, which contains 5 generators and 20 transmission lines. The generation cost of each generator is assumed to be linear as in (1), with cost coefficient c_i generated uniformly at random from [5, 20]. We assume that there is one user on each bus and each user i has disutility $f_i(\mathbf{x}_i) = (\mathbf{x}_i - \bar{\mathbf{x}}_i)^2$, where $\bar{\mathbf{x}}_i$ is some constant that represents, for example, the targeted consumption of user i .

In order to achieve the optimal solution to the global social welfare problem (2) without solving it directly, we randomly initialize the price signal for individual users from [5, 15] and run our proposed incentive pricing mechanism, whose dynamics together with the evolution of user demand profile are provided in Fig. 2. Obviously, both price signals \vec{p} and user demands \vec{x} converge very fast. Particularly, \vec{x} successfully converges to the optimal solution of problem (2). Note that the cost on each bus might converge to one of a few values, which is typical when a small number of lines are congested [30].

VI. CONCLUSIONS AND OUTLOOK

This paper extends adaptive pricing mechanisms for social welfare optimization to network-constrained energy systems with nonsmooth cost structures. By embedding DCOPF constraints into the operator's objective and introducing a generalized gradient-based price update rule, we establish a provably convergent and privacy-preserving incentive design framework. Our theoretical analysis demonstrates the existence, uniqueness, and strong asymptotic stability of the equilibrium at which the global social welfare is achieved. Simulation results validate the effectiveness of the proposed

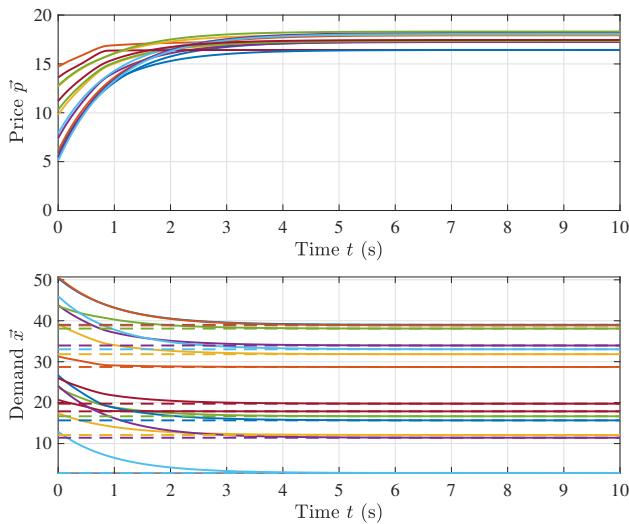


Fig. 2: Convergence of price signals and user demands of our proposed mechanism in IEEE 14-bus system with 14 users, where the dashed lines represent the optimal demand profile to the global social welfare problem.

mechanism in guiding user behavior toward globally optimal outcomes under realistic power network constraints.

Looking ahead, several important extensions remain open. First, practical systems are subject to uncertainty from renewable generation and stochastic user demand. Extending the current framework to handle uncertainty explicitly, either through robust or stochastic formulations of DCOPF, is a natural next step. Second, applying the method to AC power flow models would enhance its applicability to real-world grids, though this introduces significant nonconvexity. Third, while our current pricing update relies on analytical computation of generalized gradients, a promising direction is to develop data-driven or learning-based approximations for the operator's update rule, especially in settings where exact DCOPF gradients are computationally expensive or unavailable in real time. Finally, although our current results focus on the single time-step case, extending the convergence and stability guarantees to multi-time-step scenarios is an important direction for future work, especially for dynamic and time-coupled systems.

REFERENCES

- [1] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A direct load control model for virtual power plant management," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 959–966, May 2009.
- [2] C. Chen, J. Wang, and S. Kishore, "A distributed direct load control approach for large-scale residential demand response," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2219–2228, Sept. 2014.
- [3] F. Rahimi and A. Iapkchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 82–88, June 2010.
- [4] P. You, Y. Jiang, E. Yeung, D. F. Gayme, and E. Mallada, "On the stability, economic efficiency and incentive compatibility of electricity market dynamics," *IEEE Transactions on Automatic Control*, pp. 1–16, 2025 (Early access).
- [5] M. Peplinski and K. T. Sanders, "Residential electricity demand on CAISO flex alert days: a case study of voluntary emergency demand response programs," *Environmental Research: Energy*, vol. 1, no. 1, p. 015002, Dec. 2023.
- [6] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, First Quarter 2014.
- [7] P. Li, H. Wang, and B. Zhang, "A distributed online pricing strategy for demand response programs," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 350–360, Jan. 2019.
- [8] K. Khezeli and E. Bitar, "Risk-sensitive learning and pricing for demand response," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6000–6007, Nov. 2018.
- [9] X. Kong, D. Kong, J. Yao, L. Bai, and J. Xiao, "Online pricing of demand response based on long short-term memory and reinforcement learning," *Applied energy*, vol. 271, p. 114945, Aug. 2020.
- [10] J. Li, M. Motoki, and B. Zhang, "Socially optimal energy usage via adaptive pricing," *Electric Power Systems Research*, vol. 235, p. 110640, Oct. 2024.
- [11] C. Maheshwari, K. Kulkarni, M. Wu, and S. S. Sastry, "Inducing social optimality in games via adaptive incentive design," in *Proc. of Conference on Decision and Control*, Dec. 2022, pp. 2864–2869.
- [12] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.
- [13] L. Zhang, Y. Chen, and B. Zhang, "A convex neural network solver for DCOPF with generalization guarantees," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 2, pp. 719–730, June 2022.
- [14] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*. Belmont, Mass. : Athena Scientific, 1997.
- [15] E. Mallada, C. Zhao, and S. Low, "Optimal load-side control for frequency regulation in smart grids," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6294–6309, Dec. 2017.
- [16] F. Dörfler and S. Grammatico, "Gather-and-broadcast frequency control in power systems," *Automatica*, vol. 79, pp. 296–305, May 2017.
- [17] Mathematics Department at Wayne State University, "Every convex function is locally lipschitz," *The American Mathematical Monthly*, vol. 79, no. 10, p. 1121–1124, Dec. 1972.
- [18] D. P. Bertsekas, *Convex Optimization Theory*. Athena Scientific, 2009.
- [19] S. J. Wright and B. Recht, *Optimization for Data Analysis*. Cambridge University Press, 2022.
- [20] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*. Springer-Verlag, 1998.
- [21] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [22] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. McGraw Hill, 2013.
- [23] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*. Springer, 1998.
- [24] J. Cortes, "Discontinuous dynamical systems," *IEEE Control Systems Magazine*, vol. 28, no. 3, pp. 36–73, June 2008.
- [25] P. D. Lax and M. S. Terrell, *Multivariable Calculus with Applications*, 1st ed. Springer, 2017.
- [26] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, 3rd ed. Elsevier, 2013.
- [27] M. Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*. Addison-Wesley Publishing Company, 1965.
- [28] C. Clason, *Nonsmooth Analysis and Optimization*. University of Graz, 2024.
- [29] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910–1914, Sept. 1994.
- [30] B. Zhang, R. Rajagopal, and D. Tse, "Network risk limiting dispatch: Optimal control and price of uncertainty," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2442–2456, Sept. 2014.